LinearRegresssionGenderScopus2minimalRobustLoop.R

MikeT

Tue Jun 11 19:58:52 2019

```
library(car) #vif
library(psych) # Geometric.mean

##
## Attaching package: 'psych'

## The following object is masked from 'package:car':
##
##     logit

library(robustbase) #lmrob for linear regression in the presence of
heteroscedasticity https://stackoverflow.com/questions/23893103/increase-
maximum-number-of-iterations-in-lmrob

##
## Attaching package: 'robustbase'

## The following object is masked from 'package:psych':
##
##     cushny

#library(plyr) #for count
library(expss) #for countif

## Loading required package: magrittr

## Loading required package: htmlTable

##
## Use 'expss_output_rnotebook()' to display tables inside R Notebooks.
##  To return to the console output, use 'expss_output_default()'.

##
## Attaching package: 'expss'

## The following object is masked from 'package:car':
##
##     recode

outlier_threshold = 2.5
options(digits=4)
#FirstCode <- 1000
#LastCode <- 3616 #3616
#row.names <- c(FirstCode:LastCode)
```

```
#FieldCount <- LastCode - FirstCode + 1

FieldCodeSet <-
c(1000,1100,1101,1102,1103,1104,1105,1106,1107,1108,1109,1110,1111,1200,1201,
1202,1203,1204,1205,1206,1207,1208,1209,1210,1211,1212,1213,1300,1301,1302,13
03,1304,1305,1306,1307,1308,1309,1310,1311,1312,1313,1314,1315,1400,1401,1402
,1403,1404,1405,1406,1407,1408,1409,1410,1500,1501,1502,1503,1504,1505,1506,1
507,1508,1600,1601,1602,1603,1604,1605,1606,1607,1700,1701,1702,1703,1704,170
5,1706,1707,1708,1709,1710,1711,1712,1800,1801,1802,1803,1804,1900,1901,1902,
1903,1904,1905,1906,1907,1908,1909,1910,1911,1912,1913,2000,2001,2002,2003,21
00,2101,2102,2103,2104,2105,2200,2201,2202,2203,2204,2205,2206,2207,2208,2209
,2210,2211,2212,2213,2214,2215,2216,2300,2301,2302,2303,2304,2305,2306,2307,2
308,2309,2310,2311,2312,2400,2401,2402,2403,2404,2405,2406,2500,2501,2502,250
3,2504,2505,2506,2507,2508,2600,2601,2602,2603,2604,2605,2606,2607,2608,2609,
2610,2611,2612,2613,2614,2700,2701,2702,2703,2704,2705,2706,2707,2708,2709,27
10,2711,2712,2713,2714,2715,2716,2717,2718,2719,2720,2721,2722,2723,2724,2725
,2726,2727,2728,2729,2730,2731,2732,2733,2734,2735,2736,2737,2738,2739,2740,2
741,2742,2743,2744,2745,2746,2747,2748,2800,2801,2802,2803,2804,2805,2806,280
7,2808,2809,2900,2901,2902,2903,2904,2905,2906,2907,2908,2909,2910,2911,2912,
2913,2914,2915,2916,2917,2918,2919,2920,2921,2922,2923,3000,3001,3002,3003,30
04,3005,3100,3101,3102,3103,3104,3105,3106,3107,3108,3109,3110,3200,3201,3202
,3203,3204,3205,3206,3207,3300,3301,3302,3303,3304,3305,3306,3307,3308,3309,3
310,3311,3312,3313,3314,3315,3316,3317,3318,3319,3320,3321,3322,3400,3401,340
2,3403,3404,3500,3501,3502,3503,3504,3505,3506,3600,3601,3602,3603,3604,3605,
3606,3607,3608,3609,3610,3611,3612,3613,3614,3615,3616)
#FieldCodeSet <- 2700:2748
FieldNeedsExtraTime <- FieldCodeSet
#FieldNeedsExtraTime <- c(1311, 3100, 2700, 2709, 2714, 2719,2722, 2734,
2744, 2747, 3104, 3107, 1701, 1706, 3302,3303,3309, 3320, 3322)
row.names <- FieldCodeSet
FieldCount <- length(FieldCodeSet)
FirstYearForCitationAnalysis <- 1996
LastYearForCitationAnalysis <- 2012
MaxAuthorsToCountInRegression <- 5;
MaxCountriesAllowed <- 1; #set to 0 to skip
#One country only unless comment below.
#RootFileName <- "E:\\data\\Scopus\\All fields regression\\All 1996-2018 332
fields was Jamaica\\cov3\\ScopusFind Jamaica"; Country <- "All"
#RootFileName <- "E:\\data\\Scopus\\All fields regression\\Australia 1996-
2018 331 fields\\cov3\\ScopusFind Australia"; Country <- "Aus"
RootFileName <- "E:\\data\\Scopus\\All fields regression\\Canada 1996-2018
331 fields\\cov3\\ScopusFind Canada"; Country <- "Ca"
#RootFileName <- "E:\\data\\Scopus\\All fields regression\\China 1996-2018
327 fields\\cov3\\ScopusFind China"; Country <- "CN"
#RootFileName <- "E:\\data\\Scopus\\All fields regression\\Germany 1996-2018
330 fields\\cov3\\ScopusFind Germany"; Country <- "De"
#RootFileName <- "E:\\data\\Scopus\\All fields regression\\Ireland 1996-2018
329 fields\\cov3\\ScopusFind Ireland"; Country <- "Ie"
#RootFileName <- "E:\\data\\Scopus\\All fields regression\\Spain 1996-2018
329 fields\\cov3\\ScopusFind Spain"; Country <- "Es"
```

```
#RootFileName <- "E:\\data\\Scopus\\All fields regression\\NZ 1996-2018 327
fields\\cov3\\ScopusFind New Zealand"; Country <- "Nz"
#RootFileName <- "E:\\data\\Scopus\\All fields regression\\UK gender 1996-
2018 330 fields\\cov3\\ScopusFind United Kingdom"; Country <- "UK"
RootFileName <- "E:\\data\\Scopus\\All fields regression\\USA 1996-2018 331
fields\\cov3\\United States "; Country <- "Usa"
column.names <- c("Articles",
"FirstF","FirstM","FirstP","LastF","LastM","LastP")
GenderTeamSize <- array(integer(7*FieldCount),dim = c(FieldCount, 7),
dimnames <- list(row.names, column.names))
column.names <- c("FFA1","FLA1","2","3","4","5+","FFA2",
"FLA2","FFA3","FLA4")
RegCoef <- array(integer(10*FieldCount),dim = c(FieldCount, 10), dimnames <-
list(row.names, column.names))
column.names <- c("FFA1p","FLA1p","2p","3p","4p","5+p","FFA2p",
"FLA2p","FFA3p","FLA4p")
RegP <- array(integer(10*FieldCount),dim = c(FieldCount, 10), dimnames <-
list(row.names, column.names))
RegStar <- array(character(10*FieldCount),dim = c(FieldCount, 10), dimnames
<- list(row.names, column.names))
column.names <- c("1Ctry","Gend 1st nth","Fem1 96","Fem1 18","Chg1","FemN
96","FemN 18", "ChgN")
BasicStats <- array(integer(8*FieldCount),dim = c(FieldCount, 8), dimnames <-
list(row.names, column.names))
column.names <- c("1CtyMFto14","Fem1All","Fem1 96","Fem1
14","Chg1","FemNAll","FemN96","FemN14", "ChgN", "1CtyMF96", "1CtyMF14")
BasicStats2 <- array(integer(11*FieldCount),dim = c(FieldCount, 11), dimnames
<- list(row.names, column.names))
column.names <- c("CitationSet","Female1st","FemaleLast")
GenderAnalysed <- array(integer(3*FieldCount),dim = c(FieldCount, 3),
dimnames <- list(row.names, column.names))

for (i in 1:FieldCount) {
  for (j in 1:7) {
    GenderTeamSize[i,j] <- NA; BasicStats[i,j]<-NA; BasicStats2[i,j]<-NA;
RegCoef[i,j]<-NA ; RegP[i,j]<-NA
  }
  BasicStats[i,8]<-NA; BasicStats2[i,8]<-NA; RegCoef[i,8]<-NA ; RegP[i,8]<-NA
  BasicStats2[i,9]<-NA; RegCoef[i,9]<-NA ; RegP[i,9]<-NA
  BasicStats2[i,10]<-NA;RegCoef[i,10]<-NA ; RegP[i,10]<-NA
  BasicStats2[i,11]<-NA
}

FieldCount = 0
#iCode <- 3302 #test
for (iCode in FieldCodeSet) {
  FieldCount <- FieldCount + 1 #
  SampleFile <- paste(RootFileName, iCode, "_cov.txt", sep="");
  BasicStats[FieldCount,1] <- 0;  BasicStats2[FieldCount,1] <- 0;
  if (file.exists(SampleFile)) {
```

```
    tryCatch({
      print(""); print("")
      print("##################################")
      print(paste("Analysis of AJSC",iCode))
      print("##################################")
      AllScopusData <-read.table(file=SampleFile, head=TRUE, sep = "\t")
      names(AllScopusData)[3] <- "NLCS"
      names(AllScopusData)[4] <- "Year"
      names(AllScopusData)[5] <- "OneField"
      names(AllScopusData)[6] <- "Fields"
      #AllScopusData <- AllScopusData[AllScopusData$Year!=2004,] #Uncomment
this to remove year with high outliers if code below reveals a problem
      if (MaxCountriesAllowed>0) AllScopusData <-
AllScopusData[AllScopusData$UniqueCountries <= MaxCountriesAllowed,] #Comment
out to include non-us authors after 1st
      AllScopusDataOlder <- AllScopusData[AllScopusData$Year <=
LastYearForCitationAnalysis,] #2015 Ensure citation window of at least 3
years 2013 for 5 years
      AllScopusDataOlder$Year <- factor(AllScopusDataOlder$Year) #Treat each
year separately
      AllScopusDataOlderFirstGendered <-
AllScopusDataOlder[AllScopusDataOlder$FirstAuthorFemale > -1,] #Female 1,
Male 0, Unknown -1
      AllScopusDataOlderFirstGendered$FirstAuthorFemale <-
factor(AllScopusDataOlderFirstGendered$FirstAuthorFemale)
      AllScopusDataOlderFirstLastGendered <-
AllScopusDataOlderFirstGendered[AllScopusDataOlderFirstGendered$LastAuthorFem
ale > -1,] #Female 1, Male 0, Unknown -1
      AllScopusDataOlderFirstLastGendered$LastAuthorFemale <-
factor(AllScopusDataOlderFirstLastGendered$LastAuthorFemale) #Female 1, Male
0, Unknown -1
      AllYearsBothGendered <-AllScopusData[AllScopusData$FirstAuthorFemale >
-1,]
      AllYearsBothGendered <-
AllYearsBothGendered[AllYearsBothGendered$LastAuthorFemale > -1,]
      AllYearsBothGendered1996 <-
AllYearsBothGendered[AllYearsBothGendered$Year == 1996,]
      AllYearsBothGenderedLast <-
AllYearsBothGendered[AllYearsBothGendered$Year ==
LastYearForCitationAnalysis,]
      AllYearsBothGendered2018 <-
AllYearsBothGendered[AllYearsBothGendered$Year == 2018,]
      BasicStats[FieldCount,1] <- nrow(AllScopusData)
      BasicStats[FieldCount,2] <- nrow(AllYearsBothGendered) /
nrow(AllScopusData) * 100 #Percentage
      BasicStats[FieldCount,3] <- count_if(1,
AllYearsBothGendered1996$FirstAuthorFemale)/ nrow(AllYearsBothGendered1996) *
100
      BasicStats[FieldCount,4] <- count_if(1,
AllYearsBothGendered2018$FirstAuthorFemale)/ nrow(AllYearsBothGendered2018) *
```

```
100
      BasicStats[FieldCount,5] <- BasicStats[FieldCount,4] -
BasicStats[FieldCount,3]
      BasicStats[FieldCount,6] <- count_if(1,
AllYearsBothGendered1996$LastAuthorFemale)/ nrow(AllYearsBothGendered1996) *
100
      BasicStats[FieldCount,7] <- count_if(1,
AllYearsBothGendered2018$LastAuthorFemale)/ nrow(AllYearsBothGendered2018) *
100
      BasicStats[FieldCount,8] <- BasicStats[FieldCount,7] -
BasicStats[FieldCount,6]
      ##c("US gendered96to14","Fem1 All","Fem1 96","Fem1 14","Chg","FemN
All","FemN 96","FemN 14", "Chg")
      BasicStats2[FieldCount,1] <- nrow(AllScopusDataOlderFirstLastGendered)
      BasicStats2[FieldCount,2] <- count_if(1,
AllScopusDataOlderFirstLastGendered$FirstAuthorFemale) /
nrow(AllScopusDataOlderFirstLastGendered) * 100 #Percentage
      BasicStats2[FieldCount,3] <- count_if(1,
AllYearsBothGendered1996$FirstAuthorFemale)/ nrow(AllYearsBothGendered1996) *
100
      BasicStats2[FieldCount,4] <- count_if(1,
AllYearsBothGenderedLast$FirstAuthorFemale)/ nrow(AllYearsBothGenderedLast) *
100
      BasicStats2[FieldCount,5] <- BasicStats2[FieldCount,4] -
BasicStats2[FieldCount,3]
      BasicStats2[FieldCount,6] <- count_if(1,
AllScopusDataOlderFirstLastGendered$LastAuthorFemale) /
nrow(AllScopusDataOlderFirstLastGendered) * 100
      BasicStats2[FieldCount,7] <- count_if(1,
AllYearsBothGendered1996$LastAuthorFemale)/ nrow(AllYearsBothGendered1996) *
100
      BasicStats2[FieldCount,8] <- count_if(1,
AllYearsBothGenderedLast$LastAuthorFemale)/ nrow(AllYearsBothGenderedLast) *
100
      BasicStats2[FieldCount,9] <- BasicStats2[FieldCount,8] -
BasicStats2[FieldCount,7]
      BasicStats2[FieldCount,10] <- nrow(AllYearsBothGendered1996)
      BasicStats2[FieldCount,11] <- nrow(AllYearsBothGenderedLast)
      #print("MNLCS for all years [All, first gendered, first & last
gendered], just to check nothing is odd")
      #print(tapply(AllScopusDataOlder$NLCS, AllScopusDataOlder$Year, mean))
      #print(tapply(AllScopusDataOlderFirstGendered$NLCS,
AllScopusDataOlderFirstGendered$Year, mean))
      #print(tapply(AllScopusDataOlderFirstLastGendered$NLCS,
AllScopusDataOlderFirstLastGendered$Year, mean))

      print("Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]")
      print(table(AllScopusDataOlder$Year))
      print(table(AllScopusDataOlderFirstGendered$Year))
```

```
      print(table(AllScopusDataOlderFirstLastGendered$Year))
    }, error = function(e) return("failed narrow field data entry and basic
processing"))
    tryCatch({
      ######################################################
      print("Heteroscedasticity checks, confirming that there are problems
with these")
      ######################################################
      #Check for outliers caused by field normalisation - should be evident
in heteroskedacity caused by individual years
      print(bartlett.test(NLCS~Year, data=AllScopusDataOlderFirstGendered))
#Homogeneity of Variances test. Big fail is OK because older years have
greater variability
      YearLm <- lm(NLCS~Year, data=AllScopusDataOlderFirstGendered)
      plot(YearLm, which = 1) #Residuals vs. fitted. Check that width of
residual bar is not huge for a few years
      AllScopusDataOlderFirstGendered$YMresiduals <- resid(YearLm)
      YearGenderLm <- lm(YMresiduals~FirstAuthorFemale,
data=AllScopusDataOlderFirstGendered)
      print(bartlett.test(YMresiduals~FirstAuthorFemale,
data=AllScopusDataOlderFirstGendered)) #Homogeneity of Variances test - this
is the key test - should not fail by much, but might becuase older years have
fewer females and higher variance
      plot(YearGenderLm, which = 1) #Residuals vs. fitted. Check that width
of residual bar does not vary too much for homoscedasticity; look out for
cone shape
      #If problems here, check the outliers in regression 1
    }, error = function(e) return("failed Heteroscedasticity checks"))
    tryCatch({
      ######################################################
      # Test for different team sizes for male and female authors - first and
last gendered
      ######################################################
      #First author gender 2018
      GenderAuthors2018 <-
split(AllYearsBothGendered2018$UniqueAuthors,AllYearsBothGendered2018$FirstAu
thorFemale)
      FemaleTeamSizes2018 <- as.numeric(GenderAuthors2018$'1')
      MaleTeamSizes2018 <- as.numeric(GenderAuthors2018$'0')
      GenderAnalysed[FieldCount,1] <- nrow(AllYearsBothGendered2018)
      GenderAnalysed[FieldCount,2] <- length(FemaleTeamSizes2018)
      print(paste("Female first author team size 2018 geometric mean:",
geometric.mean(FemaleTeamSizes2018)))
      print(paste("Male first author team size 2018 geometric mean:",
geometric.mean(MaleTeamSizes2018)))
      wilc<- wilcox.test(FemaleTeamSizes2018, MaleTeamSizes2018, alternative
= "two.sided")
      print(wilc)
      GenderTeamSize[FieldCount,1] <- nrow(AllYearsBothGendered2018)
      GenderTeamSize[FieldCount,2] <- geometric.mean(FemaleTeamSizes2018)
```

```
      GenderTeamSize[FieldCount,3] <- geometric.mean(MaleTeamSizes2018)
      GenderTeamSize[FieldCount,4] <- wilc$p.value
      #Last author gender 2018
      GenderAuthors2018 <-
split(AllYearsBothGendered2018$UniqueAuthors,AllYearsBothGendered2018$LastAut
horFemale)
      FemaleTeamSizes2018 <- as.numeric(GenderAuthors2018$'1')
      MaleTeamSizes2018 <- as.numeric(GenderAuthors2018$'0')
      GenderAnalysed[FieldCount,3] <- length(FemaleTeamSizes2018)
      print(paste("Female last author team size 2018 geometric mean:",
geometric.mean(FemaleTeamSizes2018)))
      print(paste("Male last author team size 2018 geometric mean:",
geometric.mean(MaleTeamSizes2018)))
      wilc<- wilcox.test(FemaleTeamSizes2018, MaleTeamSizes2018, alternative
= "two.sided")
      print(wilc)
      GenderTeamSize[FieldCount,5] <- geometric.mean(FemaleTeamSizes2018)
      GenderTeamSize[FieldCount,6] <- geometric.mean(MaleTeamSizes2018)
      GenderTeamSize[FieldCount,7] <- wilc$p.value
    }, error = function(e) return("Failed team size tests"))
    tryCatch({
      #Use 5 for 5+ authors
      for (i in 1:nrow(AllScopusDataOlderFirstLastGendered)) {
        AllScopusDataOlderFirstLastGendered$UniqueAuthors[i] <-
min(MaxAuthorsToCountInRegression,AllScopusDataOlderFirstLastGendered$UniqueA
uthors[i])
      }

AllScopusDataOlderFirstLastGendered$UniqueAuthors=factor(AllScopusDataOlderFi
rstLastGendered$UniqueAuthors) #Don't assume any team size influence formula
      ##########################################################
      print("Regression 1: First author gender, last author gender, team
size, Year as factors")
      ##########################################################
      if (iCode %in% FieldNeedsExtraTime) {
        FirstLastAuthorTeamLmrob <-
lmrob(NLCS~FirstAuthorFemale+LastAuthorFemale+UniqueAuthors+Year,
data=AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf, k.max = 1000))
      } else {
        FirstLastAuthorTeamLmrob <-
lmrob(NLCS~FirstAuthorFemale+LastAuthorFemale+UniqueAuthors+Year,
data=AllScopusDataOlderFirstLastGendered)
      }
      print(vif(FirstLastAuthorTeamLmrob)) # check under 5
      hist(resid(FirstLastAuthorTeamLmrob),main='Residuals from first and
last author and team size',xlab='Standardised Residuals',ylab='Frequency')
#Normality not needed but useful - check for outliers - if too many outside
+/- 1.96 then possible field normalisation problem
      AllScopusDataOlderFirstLastGendered$residuals <-
```

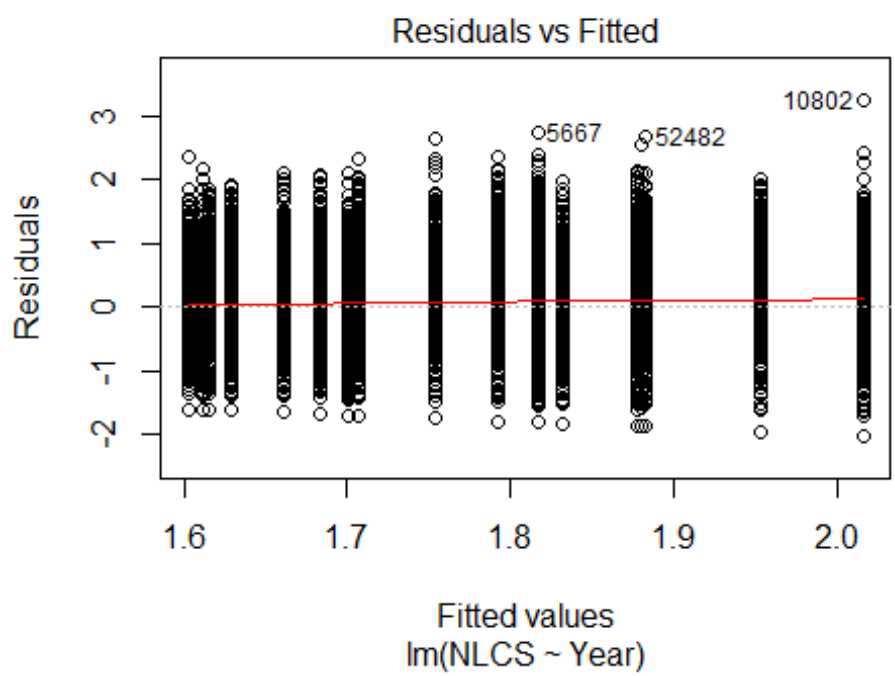
```
resid(FirstLastAuthorTeamLmrob)
      Outliers <- AllScopusDataOlderFirstLastGendered[
abs(AllScopusDataOlderFirstLastGendered$residuals) > outlier_threshold, ]
      Outliers <- Outliers[c(1,3,4,5,6,16)]
      Outliers$ScopusId <- gsub('scopus_id:', '', Outliers$ScopusId)
      print(paste("List of ", nrow(Outliers), "outliers with residuals above
", outlier_threshold))
      print(Outliers)
      sum <- summary(FirstLastAuthorTeamLmrob)
      print(sum)
      RegCoef[FieldCount,1] <- sum$coefficients[2]; RegP[FieldCount,1] <-
sum$coefficients[2,4];
      RegCoef[FieldCount,2] <- sum$coefficients[3]; RegP[FieldCount,2] <-
sum$coefficients[3,4];
      RegCoef[FieldCount,3] <- sum$coefficients[4]; RegP[FieldCount,3] <-
sum$coefficients[4,4];
      RegCoef[FieldCount,4] <- sum$coefficients[5]; RegP[FieldCount,4] <-
sum$coefficients[5,4];
      RegCoef[FieldCount,5] <- sum$coefficients[6]; RegP[FieldCount,5] <-
sum$coefficients[6,4];
      RegCoef[FieldCount,6] <- sum$coefficients[7]; RegP[FieldCount,6] <-
sum$coefficients[7,4];
    }, error = function(e) return("Failed regression 1"))
    tryCatch({
      ##################################################
      print("Regression 2: First author gender, Last author gender, Year as
factors")
      ##################################################
      #First author, last author regression
      if (iCode %in% FieldNeedsExtraTime) {
        FirstLastAuthorLmrob <-
lmrob(NLCS~FirstAuthorFemale+LastAuthorFemale+Year,
data=AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf, k.max = 1000))
      } else {
        FirstLastAuthorLmrob <-
lmrob(NLCS~FirstAuthorFemale+LastAuthorFemale+Year,
data=AllScopusDataOlderFirstLastGendered)
      }
      print(vif(FirstLastAuthorLmrob)) # check under 5
      hist(resid(FirstLastAuthorLmrob),main='Residuals from first and last
author',xlab='Standardised Residuals',ylab='Frequency') #Normality not needed
but useful - check for outliers - if too many outside +/- 1.96 then possible
field normalisation problem
      AllScopusDataOlderFirstLastGendered$residuals <-
resid(FirstLastAuthorLmrob)
      Outliers <- AllScopusDataOlderFirstLastGendered[
abs(AllScopusDataOlderFirstLastGendered$residuals) > outlier_threshold, ]
      Outliers <- Outliers[c(1,3,4,5,6,16)]
      Outliers$ScopusId <- gsub('scopus_id:', '', Outliers$ScopusId)
```

```
      print(paste("List of ", nrow(Outliers), "outliers with residuals above
", outlier_threshold))
      print(Outliers)
      sum <- summary(FirstLastAuthorLmrob)
      print(sum) #Robust to heteroscedacity
      RegCoef[FieldCount,7] <- sum$coefficients[2]; RegP[FieldCount,7] <-
sum$coefficients[2,4];
      RegCoef[FieldCount,8] <- sum$coefficients[3]; RegP[FieldCount,8] <-
sum$coefficients[3,4];
    }, error = function(e) return("Failed regression 4"))
    tryCatch({
      ########################################
      print("Regression 3: First author gender, Year as factors")
      ########################################
      #First author regression
      if (iCode %in% FieldNeedsExtraTime) {
        FirstAuthorLmrob <- lmrob(NLCS~FirstAuthorFemale+Year,
data=AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf, k.max = 1000))
      } else {
        FirstAuthorLmrob <- lmrob(NLCS~FirstAuthorFemale+Year,
data=AllScopusDataOlderFirstLastGendered)
      }
      print(vif(FirstAuthorLmrob)) # check under 5
      hist(resid(FirstAuthorLmrob),main='Residuals from first
author',xlab='Standardised Residuals',ylab='Frequency') #Normality not needed
but useful - check for outliers - if too many outside +/- 1.96 then possible
field normalisation problem
      AllScopusDataOlderFirstLastGendered$residuals <-
resid(FirstLastAuthorLmrob)
      Outliers <- AllScopusDataOlderFirstLastGendered[
abs(AllScopusDataOlderFirstLastGendered$residuals) > outlier_threshold, ]
      Outliers <- Outliers[c(1,3,4,5,6,16)]
      Outliers$ScopusId <- gsub('scopus_id:', '', Outliers$ScopusId)
      print(paste("List of ", nrow(Outliers), "outliers with residuals above
", outlier_threshold))
      print(Outliers)
      sum <- summary(FirstAuthorLmrob)
      print(sum) #Robust to heteroscedacity
      RegCoef[FieldCount,9] <- sum$coefficients[2]; RegP[FieldCount,9] <-
sum$coefficients[2,4];
    }, error = function(e) return("Failed regression 3"))
    tryCatch({
      #######################################
      print("Regression 4: Last author gender, Year as factors")
      #######################################
      if (iCode %in% FieldNeedsExtraTime) {
        LastAuthorLmrob <- lmrob(NLCS~LastAuthorFemale+Year,
data=AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf, k.max = 1000))
```

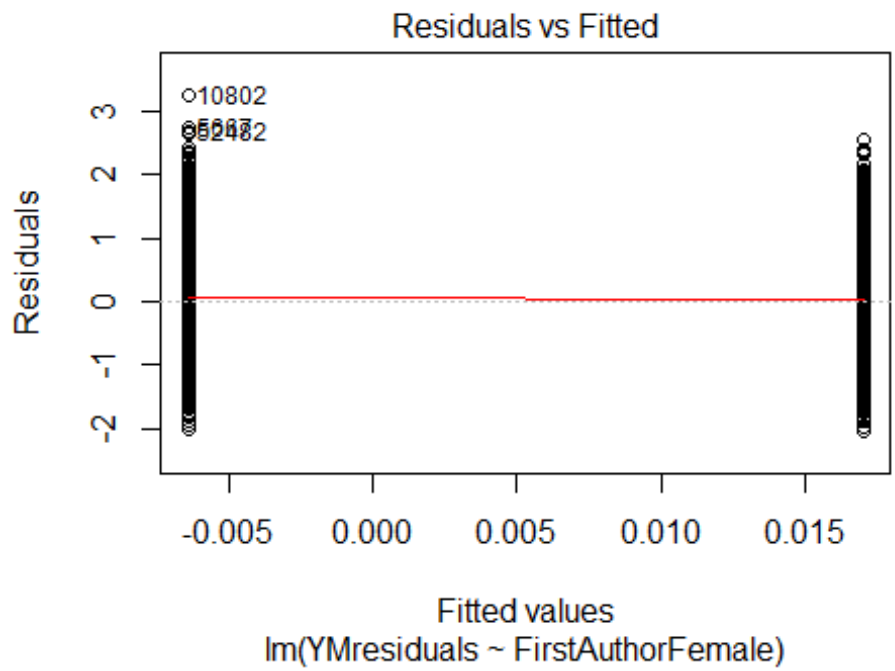
```
      } else {
        LastAuthorLmrob <- lmrob(NLCS~LastAuthorFemale+Year,
data=AllScopusDataOlderFirstLastGendered)
      }
      print(vif(LastAuthorLmrob)) # check under 5
      hist(resid(FirstLastAuthorLmrob),main='Residuals from last
author',xlab='Standardised Residuals',ylab='Frequency') #Normality not needed
but useful - check for outliers - if too many outside +/- 1.96 then possible
field normalisation problem
      AllScopusDataOlderFirstLastGendered$residuals <-
resid(FirstLastAuthorLmrob)
      Outliers <- AllScopusDataOlderFirstLastGendered[
abs(AllScopusDataOlderFirstLastGendered$residuals) > outlier_threshold, ]
      Outliers <- Outliers[c(1,3,4,5,6,16)]
      Outliers$ScopusId <- gsub('scopus_id:', '', Outliers$ScopusId)
      print(paste("List of ", nrow(Outliers), "outliers with residuals above
", outlier_threshold))
      print(Outliers)
      sum <- summary(LastAuthorLmrob) #Robust to heteroscedacity
      print(sum) #Robust to heteroscedacity
      #column.names <- c("FFA1","FLA1","2","3","4","5+","FFA2",
"FLA2","FFA3","FLA4")
      #RegCoef <- array(integer(11*FieldCount),dim = c(FieldCount, 11),
dimnames <- list(row.names, column.names))
      RegCoef[FieldCount,10] <- sum$coefficients[2]; RegP[FieldCount,10] <-
sum$coefficients[2,4];
    }, error = function(e) return("Failed regression 4"))
    tryCatch({
      print(paste("Sample size for the above analysis: ",
nrow(AllScopusDataOlderFirstLastGendered)))
    }, error = function(e) return("failed sample size"))
  }
}
```

```
## [1] ""
## [1] ""
## [1] "#####################################"
## [1] "Analysis of AJSC 1000"
## [1] "#####################################"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 2647 2638 2252 2378 2374 2605 2632 2414 2604 2644 2592 2561 2630 2671 2442
## 2011 2012
## 2517 2653
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 1544 1626 1395 1434 1356 1042 1747 1597 1713 1700 1744 1681 1750 1820 1622
## 2011 2012
```

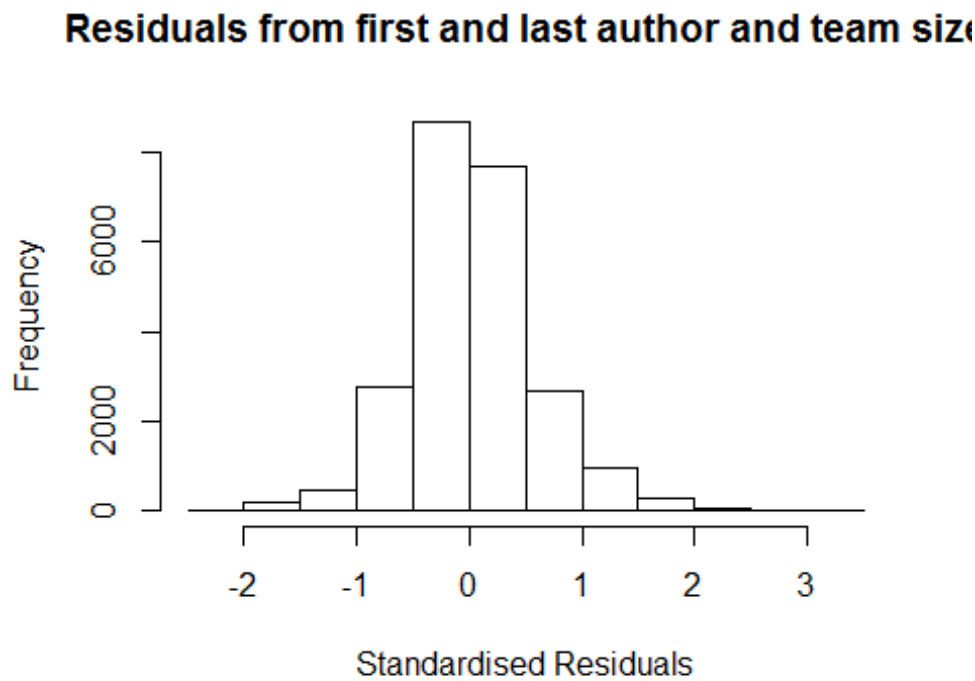
```
## 1703 1776
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 1360 1409 1238 1243 1174  937 1521 1374 1495 1459 1499 1446 1509 1590 1394
## 2011 2012
## 1475 1540
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
##  Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 860, df = 16, p-value <2e-16
```




```
##
##  Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 190, df = 1, p-value <2e-16
```



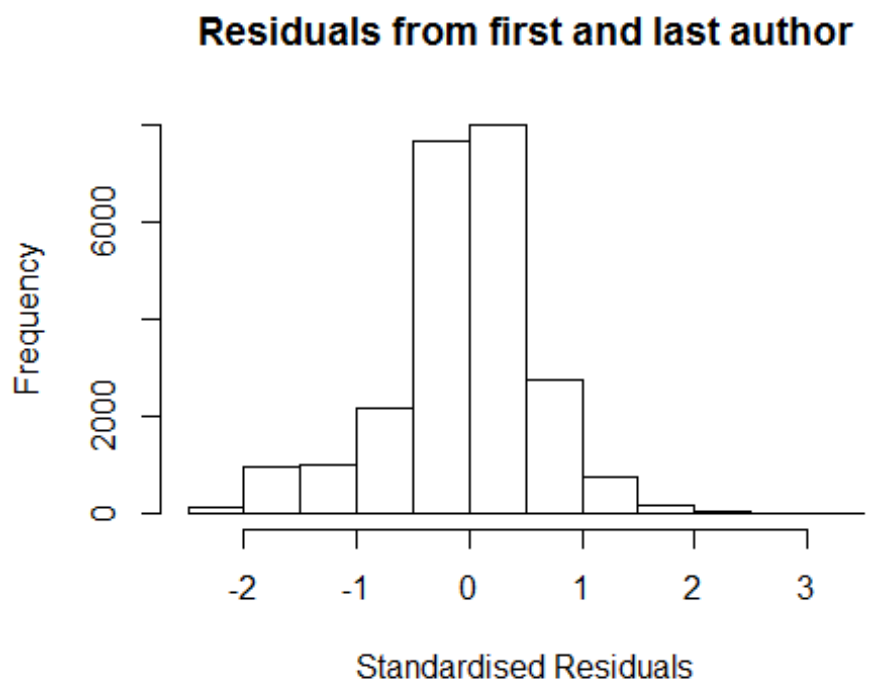

```
## [1] "Female first author team size 2018 geometric mean: 4.74762177470206"
## [1] "Male first author team size 2018 geometric mean: 4.56576746616828"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 670000, p-value = 0.2
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 4.60780152997279"
## [1] "Male last author team size 2018 geometric mean: 4.63282327310978"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 550000, p-value = 0.8
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##                     GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.013  1           1.007
## LastAuthorFemale  1.008  1           1.004
## UniqueAuthors     1.044  4           1.005
## Year              1.053 16           1.002
```




```
## [1] "List of  12 outliers with residuals above  2.5"
##          ScopusId  NLCS Year OneField Fields residuals
## 972    0029781508 4.403 1996     1000      2     2.529
## 3808   0030822582 3.451 1997     1000      2     2.690
## 3836   0030816449 3.472 1997     1000      2     2.711
## 5003   0030982247 3.862 1997     1000      2     3.101
## 6535   0032491416 3.662 1998     1000      1     2.863
## 7795   0032076192 3.426 1998     1000      2     2.627
## 10802  0033515827 5.255 1999     1000      2     3.014
## 12435  0343441557 3.320 2000     1000      2     2.696
## 14957  0035909372 4.438 2001     1000      2     2.517
## 26518  1842635159 2.989 2004     1000     11     2.554
## 28597 18744406314 3.417 2005     1000      2     2.886
## 52482 84858331964 4.546 2012     1000      2     2.553
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
UniqueAuthors +
##     Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##     k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -2.0791 -0.3154 -0.0119  0.3217  3.1010
##
## Coefficients:
```

```
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)          0.68737    0.02454   28.01  < 2e-16 ***
## FirstAuthorFemale1 -0.04349    0.00718   -6.06  1.4e-09 ***
## LastAuthorFemale1  -0.03418    0.00908   -3.76  0.00017 ***
## UniqueAuthors2      1.04239    0.02247   46.39  < 2e-16 ***
## UniqueAuthors3      1.13885    0.02156   52.81  < 2e-16 ***
## UniqueAuthors4      1.18687    0.02142   55.42  < 2e-16 ***
## UniqueAuthors5      1.30093    0.02040   63.78  < 2e-16 ***
## Year1997            0.07366    0.02083    3.54  0.00041 ***
## Year1998            0.11213    0.02144    5.23  1.7e-07 ***
## Year1999            0.25290    0.02275   11.11  < 2e-16 ***
## Year2000           -0.06299    0.02019   -3.12  0.00181 **
## Year2001            0.08991    0.02259    3.98  6.9e-05 ***
## Year2002           -0.17813    0.01861   -9.57  < 2e-16 ***
## Year2003           -0.19152    0.01891  -10.13  < 2e-16 ***
## Year2004           -0.17516    0.01839   -9.52  < 2e-16 ***
## Year2005           -0.15681    0.01876   -8.36  < 2e-16 ***
## Year2006           -0.16111    0.01842   -8.75  < 2e-16 ***
## Year2007           -0.12174    0.01935   -6.29  3.2e-10 ***
## Year2008           -0.09493    0.01918   -4.95  7.5e-07 ***
## Year2009            0.01907    0.01952    0.98  0.32865
## Year2010            0.05787    0.02060    2.81  0.00498 **
## Year2011            0.19779    0.02177    9.08  < 2e-16 ***
## Year2012            0.11869    0.02099    5.65  1.6e-08 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.464
## Multiple R-squared:  0.386,  Adjusted R-squared:  0.386
## Convergence in 18 IRWLS iterations
##
## Robustness weights:
##  38 observations
c(117,412,901,1306,1357,1738,1747,1848,2240,2302,2408,2597,2995,3557,3820,383
7,3992,4684,4916,5061,5564,6326,6607,8611,8795,9804,10206,11663,11743,12644,1
5782,16732,18501,18630,19620,20855,22456,23317)
##   are outliers with |weight| = 0 ( < 4.2e-06);
##  2039 weights are ~= 1. The remaining 21586 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0002  0.8510  0.9490  0.8770  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          4.23e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
```

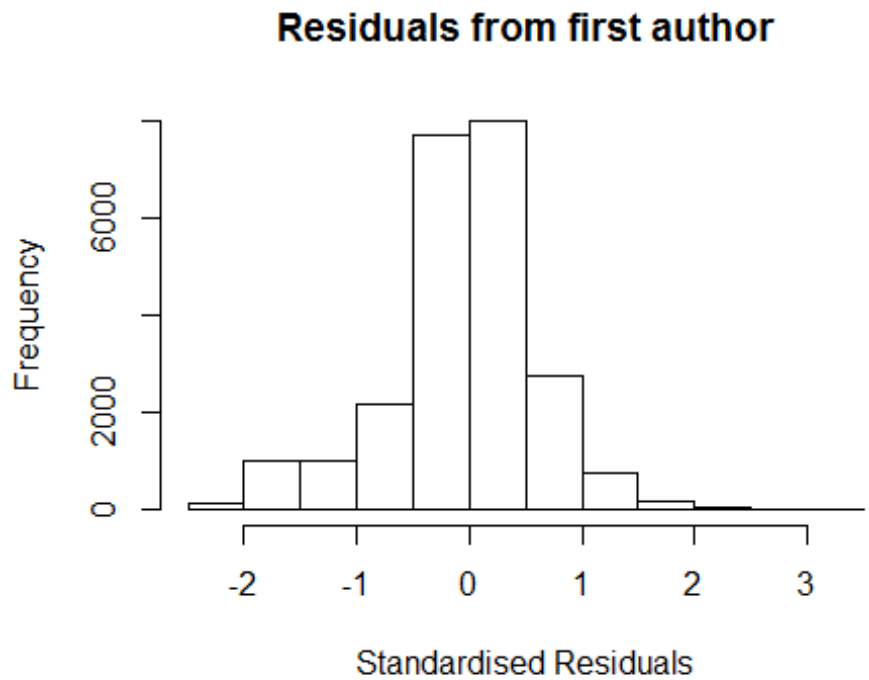
```
##           0        1000           0
##                   psi           subsampling                    cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.016  1           1.008
## LastAuthorFemale  1.011  1           1.005
## Year              1.019 16           1.001
```




```
## [1] "List of  5 outliers with residuals above  2.5"
##          ScopusId  NLCS Year OneField Fields residuals
## 972    0029781508 4.403 1996     1000      2     2.584
## 5667   0031019745 4.552 1997     1000      2     2.670
## 10802  0033515827 5.255 1999     1000      2     3.167
## 14957  0035909372 4.438 2001     1000      2     2.508
## 52482 84858331964 4.546 2012     1000      2     2.555
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
```

```
##      Min       1Q   Median       3Q      Max
## -2.09335 -0.34349 -0.00682  0.31884  3.16688
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.81945    0.01598  113.87  < 2e-16 ***
## FirstAuthorFemale1 -0.00363    0.00779   -0.47  0.64181
## LastAuthorFemale1  -0.04750    0.00993   -4.78  1.7e-06 ***
## Year1997            0.06254    0.02325    2.69  0.00714 **
## Year1998            0.11045    0.02333    4.73  2.2e-06 ***
## Year1999            0.26867    0.02447   10.98  < 2e-16 ***
## Year2000           -0.06595    0.02200   -3.00  0.00272 **
## Year2001            0.11402    0.02428    4.70  2.7e-06 ***
## Year2002           -0.19540    0.01978   -9.88  < 2e-16 ***
## Year2003           -0.19869    0.02046   -9.71  < 2e-16 ***
## Year2004           -0.16998    0.02026   -8.39  < 2e-16 ***
## Year2005           -0.14663    0.02075   -7.07  1.6e-12 ***
## Year2006           -0.13527    0.02004   -6.75  1.5e-11 ***
## Year2007           -0.10296    0.02142   -4.81  1.6e-06 ***
## Year2008           -0.07738    0.02164   -3.58  0.00035 ***
## Year2009            0.03619    0.02242    1.61  0.10649
## Year2010            0.09510    0.02278    4.17  3.0e-05 ***
## Year2011            0.27391    0.02455   11.15  < 2e-16 ***
## Year2012            0.17163    0.02389    7.18  7.0e-13 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.502
## Multiple R-squared:  0.0715, Adjusted R-squared:  0.0708
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
##  5 observations c(412,2597,4916,6607,23317)
##   are outliers with |weight| = 0 ( < 4.2e-06);
##  2183 weights are ~= 1. The remaining 21475 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.000   0.839   0.951   0.856   0.986   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          4.23e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
```

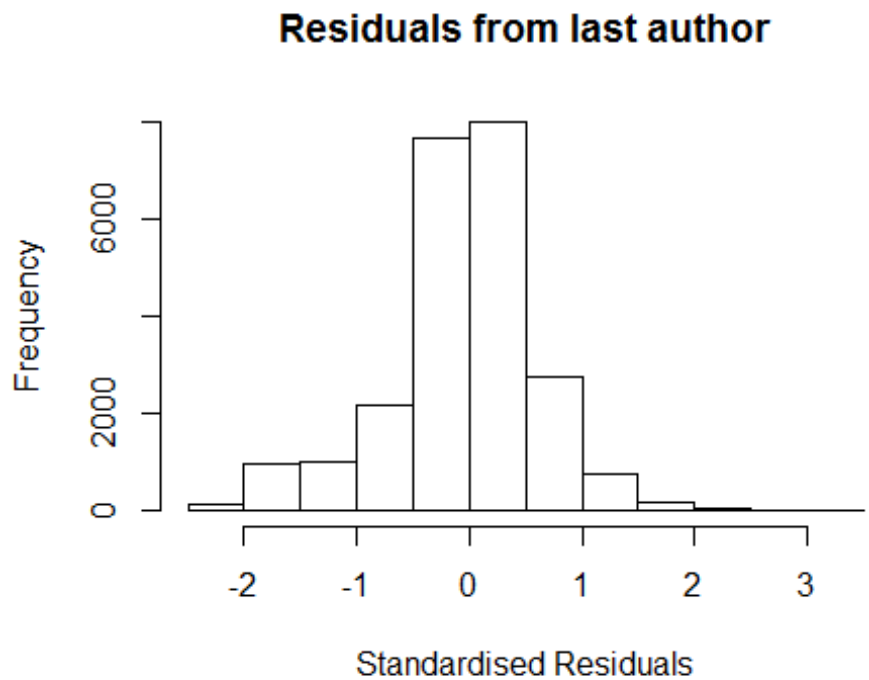
```
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.012  1           1.006
## Year              1.012 16           1.000
```




```
## [1] "List of  5 outliers with residuals above  2.5"
##          ScopusId  NLCS Year OneField Fields residuals
## 972    0029781508 4.403 1996     1000      2     2.584
## 5667   0031019745 4.552 1997     1000      2     2.670
## 10802  0033515827 5.255 1999     1000      2     3.167
## 14957  0035909372 4.438 2001     1000      2     2.508
## 52482 84858331964 4.546 2012     1000      2     2.555
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -2.08492 -0.34245 -0.00754  0.31949  3.17304
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
```

```
## (Intercept)         1.81445    0.01591  114.04  < 2e-16 ***
## FirstAuthorFemale1 -0.00640    0.00785   -0.82  0.41475
## Year1997            0.06220    0.02323    2.68  0.00742 **
## Year1998            0.10970    0.02331    4.71  2.5e-06 ***
## Year1999            0.26752    0.02447   10.93  < 2e-16 ***
## Year2000           -0.06784    0.02199   -3.08  0.00204 **
## Year2001            0.11225    0.02426    4.63  3.7e-06 ***
## Year2002           -0.19734    0.01977   -9.98  < 2e-16 ***
## Year2003           -0.20067    0.02043   -9.82  < 2e-16 ***
## Year2004           -0.17132    0.02024   -8.46  < 2e-16 ***
## Year2005           -0.14747    0.02073   -7.11  1.2e-12 ***
## Year2006           -0.13723    0.02002   -6.85  7.3e-12 ***
## Year2007           -0.10610    0.02138   -4.96  7.0e-07 ***
## Year2008           -0.07991    0.02160   -3.70  0.00022 ***
## Year2009            0.03300    0.02237    1.48  0.14009
## Year2010            0.09206    0.02275    4.05  5.2e-05 ***
## Year2011            0.27048    0.02448   11.05  < 2e-16 ***
## Year2012            0.16803    0.02387    7.04  2.0e-12 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.502
## Multiple R-squared:  0.0705, Adjusted R-squared:  0.0699
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
##  6 observations c(412,2597,4916,6607,10206,23317)
##   are outliers with |weight| = 0 ( < 4.2e-06);
##  2170 weights are ~= 1. The remaining 21487 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.000   0.839   0.951   0.856   0.986   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          4.23e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                    psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
```

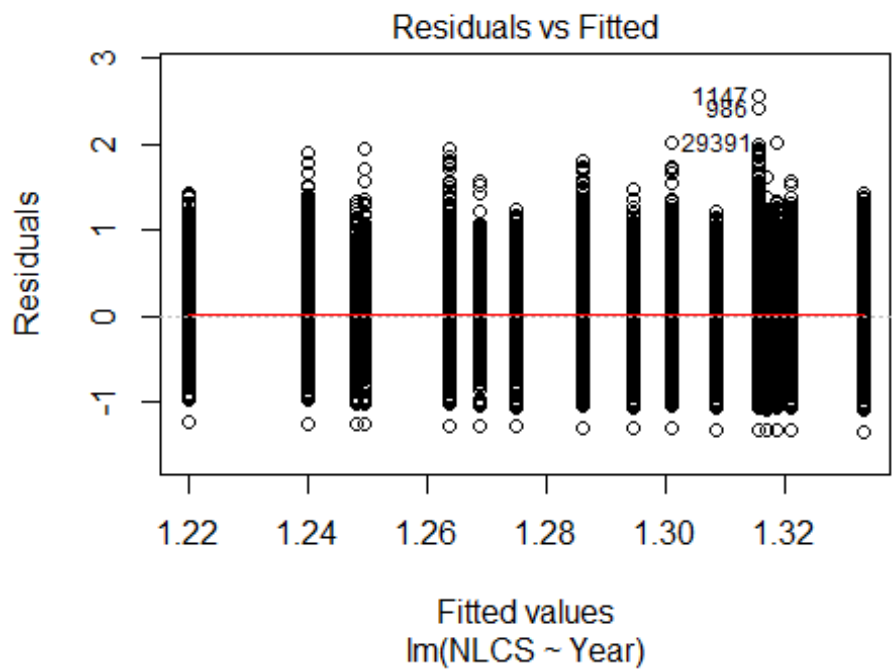
```
## LastAuthorFemale 1.007  1           1.004
## Year             1.007 16           1.000
```




```
## [1] "List of  5 outliers with residuals above  2.5"
##          ScopusId  NLCS Year OneField Fields residuals
## 972    0029781508 4.403 1996     1000      2     2.584
## 5667   0031019745 4.552 1997     1000      2     2.670
## 10802  0033515827 5.255 1999     1000      2     3.167
## 14957  0035909372 4.438 2001     1000      2     2.508
## 52482 84858331964 4.546 2012     1000      2     2.555
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -2.09192 -0.34235 -0.00664  0.31872  3.16796
##
## Coefficients:
##                   Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.8184     0.0158  115.14  < 2e-16 ***
## LastAuthorFemale1  -0.0479     0.0100   -4.78  1.7e-06 ***
## Year1997            0.0625     0.0233    2.69  0.00718 **
## Year1998            0.1105     0.0233    4.73  2.2e-06 ***
## Year1999            0.2686     0.0245   10.98  < 2e-16 ***
```

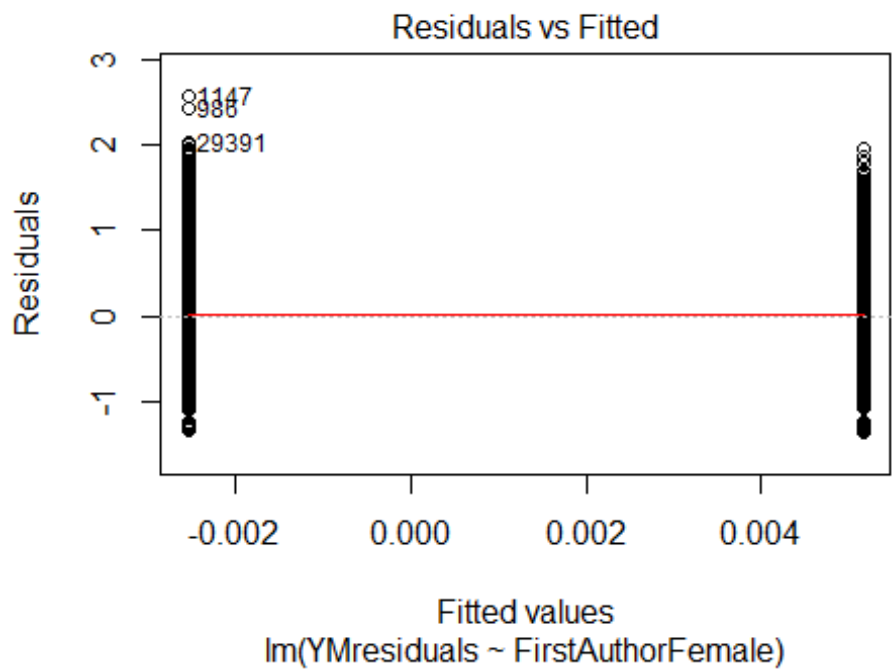
```
## Year2000           -0.0659     0.0220   -3.00  0.00274 **
## Year2001            0.1140     0.0243    4.69  2.7e-06 ***
## Year2002           -0.1953     0.0198   -9.87  < 2e-16 ***
## Year2003           -0.1987     0.0205   -9.71  < 2e-16 ***
## Year2004           -0.1700     0.0203   -8.39  < 2e-16 ***
## Year2005           -0.1467     0.0208   -7.07  1.6e-12 ***
## Year2006           -0.1354     0.0200   -6.75  1.5e-11 ***
## Year2007           -0.1031     0.0214   -4.81  1.5e-06 ***
## Year2008           -0.0775     0.0216   -3.58  0.00034 ***
## Year2009            0.0360     0.0224    1.61  0.10787
## Year2010            0.0949     0.0228    4.17  3.1e-05 ***
## Year2011            0.2735     0.0245   11.15  < 2e-16 ***
## Year2012            0.1712     0.0239    7.17  7.6e-13 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.502
## Multiple R-squared:  0.0713, Adjusted R-squared:  0.0706
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
##  5 observations c(412,2597,4916,6607,23317)
##   are outliers with |weight| = 0 ( < 4.2e-06);
##  2190 weights are ~= 1. The remaining 21468 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0001  0.8390  0.9510  0.8570  0.9860  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          4.23e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Sample size for the above analysis:  23663"
## [1] ""
## [1] ""
## [1] "####################################"
## [1] "Analysis of AJSC 1100"
## [1] "####################################"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
```

```
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 1411 1298 1341 1310 1268 1170 1220 1271 1363 1368 1477 1822 2211 2873 3595
## 2011 2012
## 4875 6610
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 1030  884  863  905  695  613  841  915  979  948 1056 1267 1533 2010 2566
## 2011 2012
## 3475 4644
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  925  788  789  823  609  538  766  804  851  852  936 1110 1295 1702 2202
## 2011 2012
## 2980 3972
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
##  Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 800, df = 16, p-value <2e-16
```



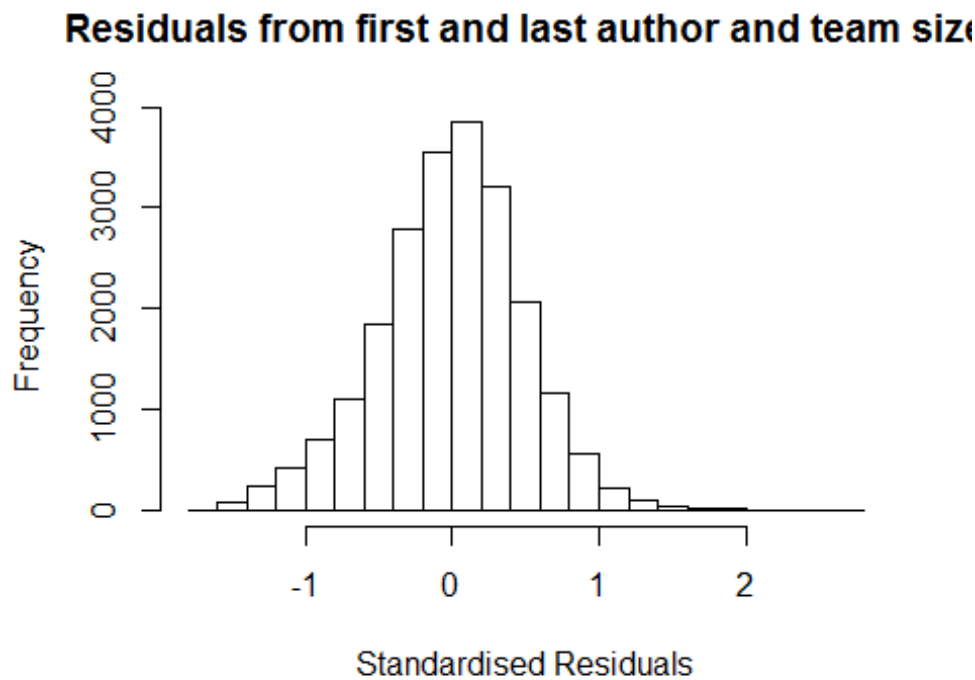

```
##
##  Bartlett test of homogeneity of variances
##
```

```
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 33, df = 1, p-value = 9e-09
```



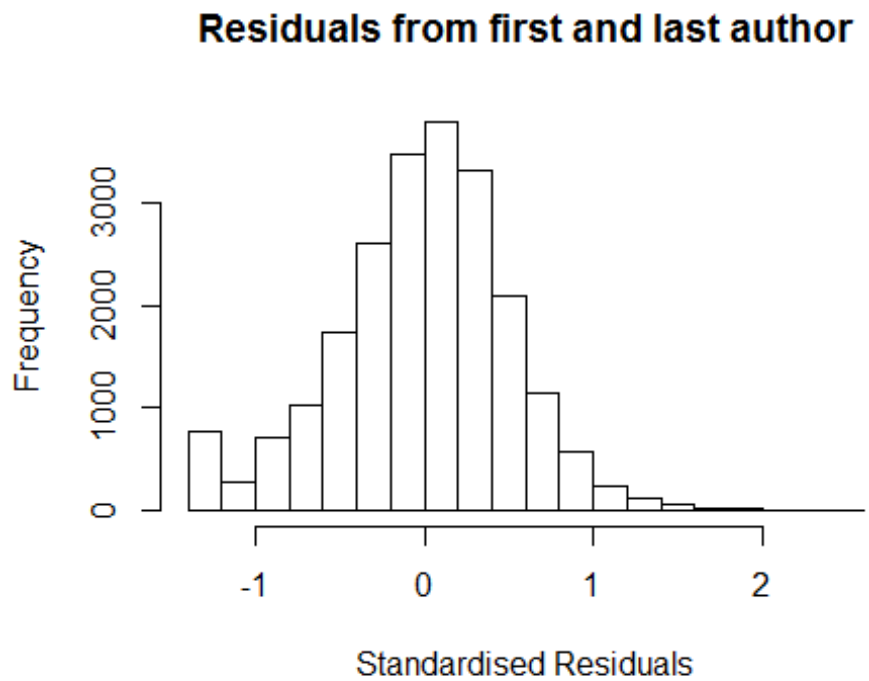

```
## [1] "Female first author team size 2018 geometric mean: 4.41833992635438"
## [1] "Male first author team size 2018 geometric mean: 3.91413491475533"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 1100000, p-value = 1e-05
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 4.20700675291648"
## [1] "Male last author team size 2018 geometric mean: 4.08870684276876"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 870000, p-value = 0.3
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.047  1           1.023
## LastAuthorFemale  1.016  1           1.008
```

```
## UniqueAuthors     1.249  4           1.028
## Year              1.236 16           1.007
```



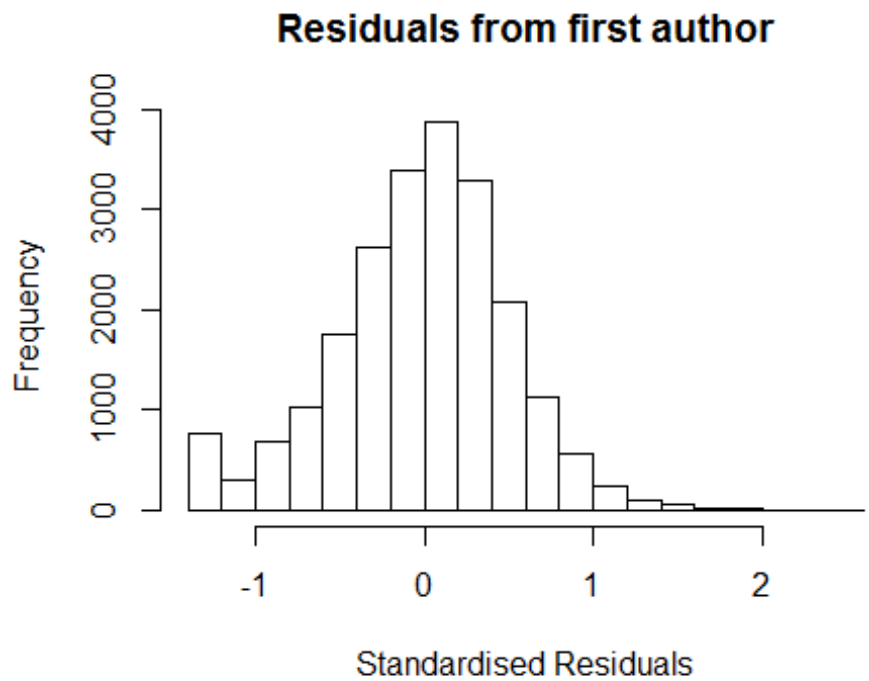

```
## [1] "List of  1 outliers with residuals above  2.5"
##        ScopusId  NLCS Year OneField Fields residuals
## 1147 0030606018 3.876 1996     1100      2     2.749
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
UniqueAuthors +
##     Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##     k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.6181 -0.3092  0.0134  0.3085  2.7492
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)          1.12685    0.02839   39.69  < 2e-16 ***
## FirstAuthorFemale1 -0.01195    0.00704   -1.70  0.08951 .
## LastAuthorFemale1  -0.03052    0.00812   -3.76  0.00017 ***
## UniqueAuthors2      0.26502    0.01297   20.44  < 2e-16 ***
## UniqueAuthors3      0.33166    0.01331   24.92  < 2e-16 ***
## UniqueAuthors4      0.39299    0.01378   28.52  < 2e-16 ***
## UniqueAuthors5      0.49125    0.01274   38.54  < 2e-16 ***
```

```
## Year1997           -0.08489    0.03723   -2.28  0.02260 *
## Year1998           -0.05927    0.03708   -1.60  0.10999
## Year1999           -0.11022    0.03410   -3.23  0.00123 **
## Year2000           -0.10043    0.03511   -2.86  0.00424 **
## Year2001           -0.05757    0.03535   -1.63  0.10340
## Year2002           -0.09464    0.03275   -2.89  0.00386 **
## Year2003           -0.13870    0.03214   -4.32  1.6e-05 ***
## Year2004           -0.08567    0.03160   -2.71  0.00672 **
## Year2005           -0.09743    0.03145   -3.10  0.00195 **
## Year2006           -0.08498    0.03156   -2.69  0.00710 **
## Year2007           -0.10154    0.03036   -3.34  0.00083 ***
## Year2008           -0.10376    0.03043   -3.41  0.00065 ***
## Year2009           -0.11161    0.02960   -3.77  0.00016 ***
## Year2010           -0.13315    0.02913   -4.57  4.9e-06 ***
## Year2011           -0.17541    0.02871   -6.11  1.0e-09 ***
## Year2012           -0.25255    0.02834   -8.91  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.459
## Multiple R-squared:  0.101,  Adjusted R-squared:   0.1
## Convergence in 16 IRWLS iterations
##
## Robustness weights:
##  2 observations c(630,753) are outliers with |weight| = 0 ( < 4.6e-06);
##  1838 weights are ~= 1. The remaining 20102 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0042  0.8630  0.9500  0.8950  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          4.56e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                  psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.024  1           1.012
## LastAuthorFemale  1.010  1           1.005
## Year              1.025 16           1.001
```



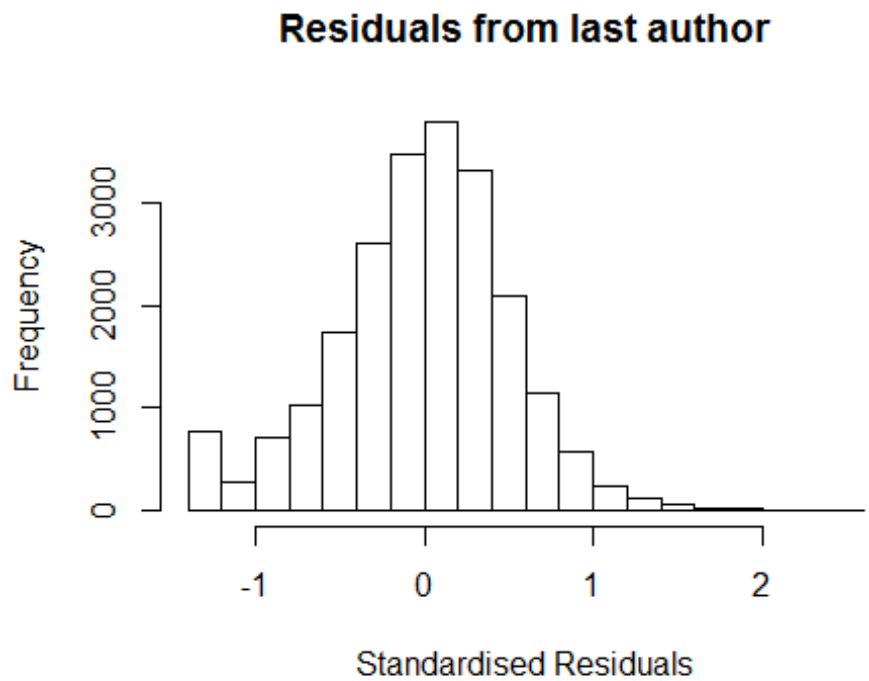

```
## [1] "List of  1 outliers with residuals above  2.5"
##        ScopusId  NLCS Year OneField Fields residuals
## 1147 0030606018 3.876 1996     1100      2     2.551
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.3704 -0.3172  0.0183  0.3173  2.5506
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)          1.32545    0.02579   51.39  < 2e-16 ***
## FirstAuthorFemale1  0.02111    0.00721    2.93  0.00344 **
## LastAuthorFemale1  -0.03239    0.00848   -3.82  0.00014 ***
## Year1997           -0.07856    0.03639   -2.16  0.03089 *
## Year1998           -0.04924    0.03613   -1.36  0.17293
## Year1999           -0.10081    0.03312   -3.04  0.00233 **
## Year2000           -0.06342    0.03470   -1.83  0.06761 .
## Year2001           -0.02310    0.03549   -0.65  0.51513
## Year2002           -0.03474    0.03174   -1.09  0.27385
## Year2003           -0.07174    0.03124   -2.30  0.02167 *
## Year2004           -0.01268    0.03070   -0.41  0.67968
## Year2005           -0.01772    0.03056   -0.58  0.56190
```

```
## Year2006           -0.00207    0.03061   -0.07  0.94619
## Year2007            0.00342    0.02924    0.12  0.90701
## Year2008            0.01602    0.02935    0.55  0.58520
## Year2009            0.02388    0.02826    0.84  0.39823
## Year2010            0.01768    0.02770    0.64  0.52331
## Year2011           -0.00873    0.02707   -0.32  0.74714
## Year2012           -0.07846    0.02663   -2.95  0.00322 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.47
## Multiple R-squared:  0.00739,    Adjusted R-squared:  0.00657
## Convergence in 16 IRWLS iterations
##
## Robustness weights:
##  2 observations c(630,753) are outliers with |weight| = 0 ( < 4.6e-06);
##  1902 weights are ~= 1. The remaining 20038 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0179  0.8610  0.9500  0.8900  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          4.56e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##                   GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.02  1           1.010
## Year              1.02 16           1.001
```




```
## [1] "List of  1 outliers with residuals above  2.5"
##        ScopusId  NLCS Year OneField Fields residuals
## 1147 0030606018 3.876 1996     1100      2     2.551
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.3604 -0.3174  0.0214  0.3184  2.5555
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)          1.32046    0.02577   51.24   <2e-16 ***
## FirstAuthorFemale1  0.01647    0.00724    2.27   0.0230 *
## Year1997           -0.07850    0.03637   -2.16   0.0309 *
## Year1998           -0.05079    0.03620   -1.40   0.1606
## Year1999           -0.10001    0.03311   -3.02   0.0025 **
## Year2000           -0.06474    0.03469   -1.87   0.0621 .
## Year2001           -0.02378    0.03554   -0.67   0.5034
## Year2002           -0.03466    0.03173   -1.09   0.2747
## Year2003           -0.07271    0.03123   -2.33   0.0199 *
## Year2004           -0.01347    0.03069   -0.44   0.6608
## Year2005           -0.01884    0.03056   -0.62   0.5376
## Year2006           -0.00321    0.03061   -0.10   0.9164
```

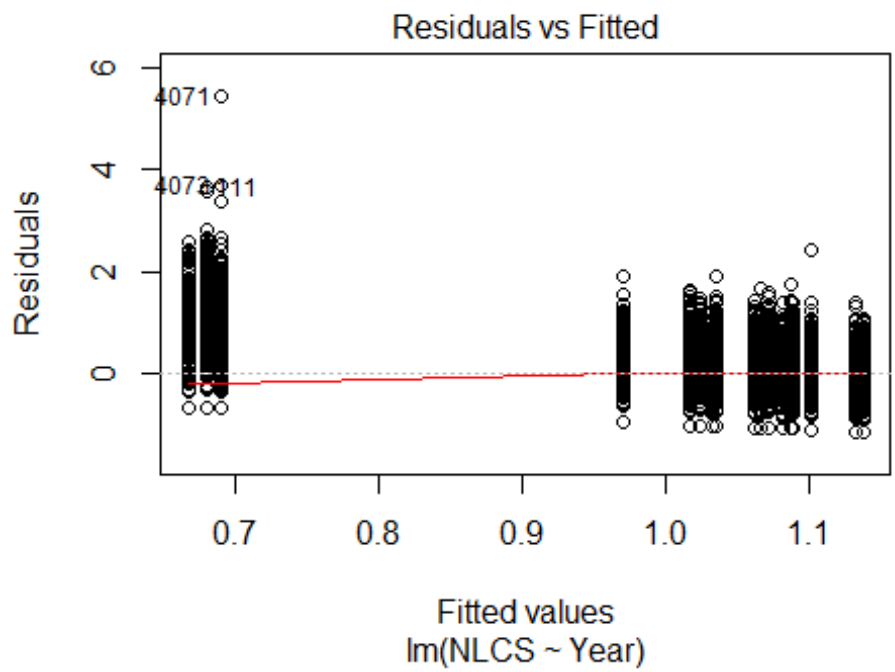
```
## Year2007            0.00336    0.02924    0.11   0.9086
## Year2008            0.01629    0.02935    0.55   0.5790
## Year2009            0.02347    0.02826    0.83   0.4064
## Year2010            0.01717    0.02770    0.62   0.5354
## Year2011           -0.00985    0.02706   -0.36   0.7158
## Year2012           -0.07953    0.02663   -2.99   0.0028 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.47
## Multiple R-squared:  0.00669,    Adjusted R-squared:  0.00592
## Convergence in 16 IRWLS iterations
##
## Robustness weights:
##  2 observations c(630,753) are outliers with |weight| = 0 ( < 4.6e-06);
##  1846 weights are ~= 1. The remaining 20094 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0242  0.8610  0.9500  0.8900  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          4.56e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##                   GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.006  1           1.003
## Year             1.006 16           1.000
```



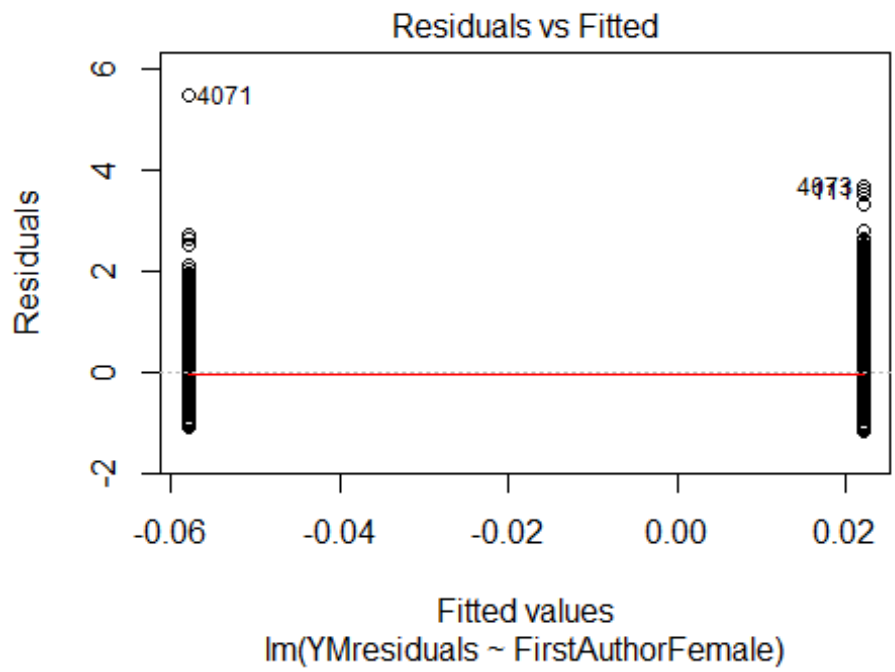

```
## [1] "List of  1 outliers with residuals above  2.5"
##        ScopusId  NLCS Year OneField Fields residuals
## 1147 0030606018 3.876 1996     1100      2     2.551
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.3554 -0.3198  0.0186  0.3190  2.5460
##
## Coefficients:
##                   Estimate Std. Error t value Pr(>|t|)
## (Intercept)        1.32999    0.02571   51.73  < 2e-16 ***
## LastAuthorFemale1 -0.02852    0.00849   -3.36  0.00078 ***
## Year1997          -0.07928    0.03638   -2.18  0.02935 *
## Year1998          -0.04926    0.03617   -1.36  0.17328
## Year1999          -0.10121    0.03313   -3.05  0.00225 **
## Year2000          -0.06298    0.03474   -1.81  0.06986 .
## Year2001          -0.02316    0.03550   -0.65  0.51413
## Year2002          -0.03389    0.03177   -1.07  0.28598
## Year2003          -0.07067    0.03127   -2.26  0.02380 *
## Year2004          -0.01158    0.03071   -0.38  0.70617
## Year2005          -0.01684    0.03057   -0.55  0.58167
## Year2006          -0.00133    0.03062   -0.04  0.96538
```

```
## Year2007           0.00484    0.02925    0.17  0.86867
## Year2008           0.01749    0.02937    0.60  0.55147
## Year2009           0.02539    0.02827    0.90  0.36926
## Year2010           0.02041    0.02771    0.74  0.46143
## Year2011          -0.00599    0.02707   -0.22  0.82481
## Year2012          -0.07527    0.02664   -2.83  0.00472 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.47
## Multiple R-squared:  0.00705,    Adjusted R-squared:  0.00628
## Convergence in 16 IRWLS iterations
##
## Robustness weights:
##  2 observations c(630,753) are outliers with |weight| = 0 ( < 4.6e-06);
##  1888 weights are ~= 1. The remaining 20052 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0207  0.8600  0.9490  0.8900  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          4.56e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                    psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Sample size for the above analysis:  21942"
## [1] ""
## [1] ""
## [1] "####################################"
## [1] "Analysis of AJSC 1101"
## [1] "####################################"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 3950 2873 2021  656  697  789  702  676  770  844  708  666  702  396  403
## 2011 2012
##  372  393
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  764  843  627  345  313  445  492  446  422  487  523  508  527  287  297
```

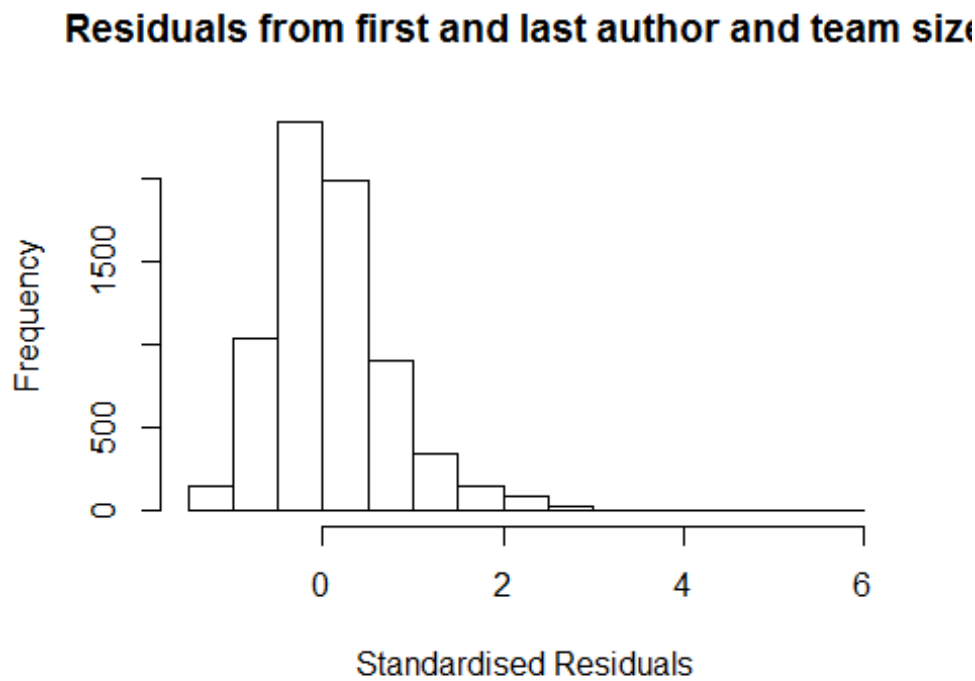
```
## 2011 2012
##  272  292
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  642  736  554  301  288  399  426  407  378  435  473  452  463  257  268
## 2011 2012
##  245  264
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
##  Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 1000, df = 16, p-value <2e-16
```




```
##
##  Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 7.6, df = 1, p-value = 0.006
```



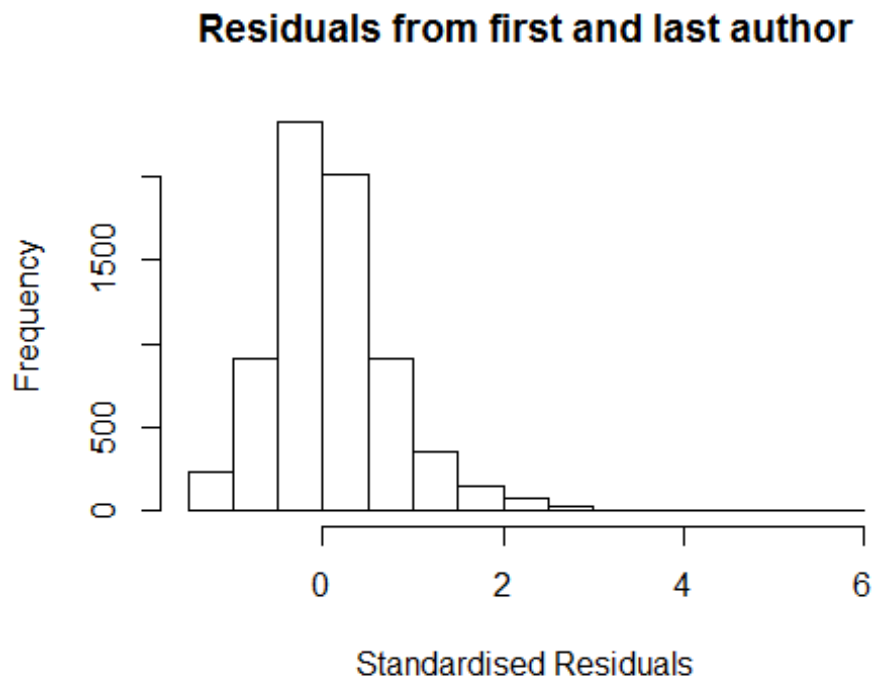

```
## [1] "Female first author team size 2018 geometric mean: 2.70869044428199"
## [1] "Male first author team size 2018 geometric mean: 2.28118364601821"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 8900, p-value = 0.02
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 2.7386994244164"
## [1] "Male last author team size 2018 geometric mean: 2.28760672129686"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 8200, p-value = 0.07
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.164  1           1.079
## LastAuthorFemale  1.168  1           1.081
## UniqueAuthors     1.150  4           1.018
## Year              1.211 16           1.006
```




```
## [1] "List of  31 outliers with residuals above  2.5"
##         ScopusId  NLCS Year OneField Fields residuals
## 75    0030285171 3.319 1996     1101      1     3.110
## 104   0030324448 3.125 1996     1101      1     2.727
## 111   0030431384 4.318 1996     1101      1     3.891
## 174   0001119747 3.118 1996     1101      3     2.700
## 3664  0029783016 3.087 1996     1100      2     2.660
## 3665  0029823855 3.316 1996     1100      2     2.889
## 3803  0002054008 4.235 1996     1101      1     4.026
## 3805  0030319283 3.487 1996     1101      1     3.278
## 3806  0030378061 3.125 1996     1101      1     2.916
## 3822  0000557865 3.053 1996     1101      2     2.626
## 3824  0000600958 2.770 1996     1101      2     2.561
## 3825  0000878057 3.223 1996     1100      2     2.796
## 3827  0001414635 2.759 1996     1100      2     2.550
## 3860  0030320143 2.722 1996     1101      2     2.513
## 3866  0030390658 2.948 1996     1101      2     2.521
## 3920  0029679856 3.073 1996     1101      3     2.601
## 3999  0000297493 3.185 1996     1100      6     2.976
## 4003  0000445638 3.022 1996     1100      6     2.813
## 4009  0000904732 2.942 1996     1100      6     2.515
## 4067  0031393667 3.337 1997     1101      1     2.936
## 4071  0031419628 6.119 1997     1101      1     5.927
## 4073  0031428001 4.380 1997     1101      1     3.866
## 4075  0031429035 3.337 1997     1101      1     3.014
## 4078 16944367296 4.052 1997     1101      1     3.363
## 6685  0030893266 3.212 1997     1100      5     2.671
```

```
## 6755  0030898875 2.970 1997     1100      2     2.647
## 6757  0030913496 3.112 1997     1100      2     2.581
## 6800  0031411844 3.117 1997     1101      2     2.576
## 7064  0032464077 2.915 1998     1101      1     2.526
## 8838  0032579440 3.265 1998     1101      6     2.539
## 9004  0032160863 3.247 1998     1101      3     2.668
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
UniqueAuthors +
##     Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##     k.max = 1000))
##  \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.33884 -0.38911 -0.00306  0.41485  5.92714
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)          0.2088     0.0381    5.48  4.3e-08 ***
## FirstAuthorFemale1  -0.0286     0.0169   -1.69  0.09071 .
## LastAuthorFemale1   -0.1023     0.0184   -5.55  2.9e-08 ***
## UniqueAuthors2       0.2182     0.0200   10.91  < 2e-16 ***
## UniqueAuthors3       0.1908     0.0213    8.97  < 2e-16 ***
## UniqueAuthors4       0.2087     0.0264    7.90  3.3e-15 ***
## UniqueAuthors5       0.3659     0.0263   13.91  < 2e-16 ***
## Year1997             0.1139     0.0463    2.46  0.01387 *
## Year1998             0.1803     0.0497    3.63  0.00028 ***
## Year1999             0.6787     0.0501   13.56  < 2e-16 ***
## Year2000             0.7164     0.0486   14.74  < 2e-16 ***
## Year2001             0.7408     0.0466   15.91  < 2e-16 ***
## Year2002             0.7046     0.0448   15.73  < 2e-16 ***
## Year2003             0.7346     0.0438   16.79  < 2e-16 ***
## Year2004             0.6607     0.0453   14.60  < 2e-16 ***
## Year2005             0.7642     0.0421   18.17  < 2e-16 ***
## Year2006             0.7619     0.0425   17.95  < 2e-16 ***
## Year2007             0.7429     0.0439   16.93  < 2e-16 ***
## Year2008             0.7463     0.0445   16.76  < 2e-16 ***
## Year2009             0.7388     0.0540   13.69  < 2e-16 ***
## Year2010             0.7221     0.0513   14.06  < 2e-16 ***
## Year2011             0.6893     0.0548   12.59  < 2e-16 ***
## Year2012             0.6092     0.0550   11.07  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.504
## Multiple R-squared:  0.238,  Adjusted R-squared:  0.235
## Convergence in 16 IRWLS iterations
##
```

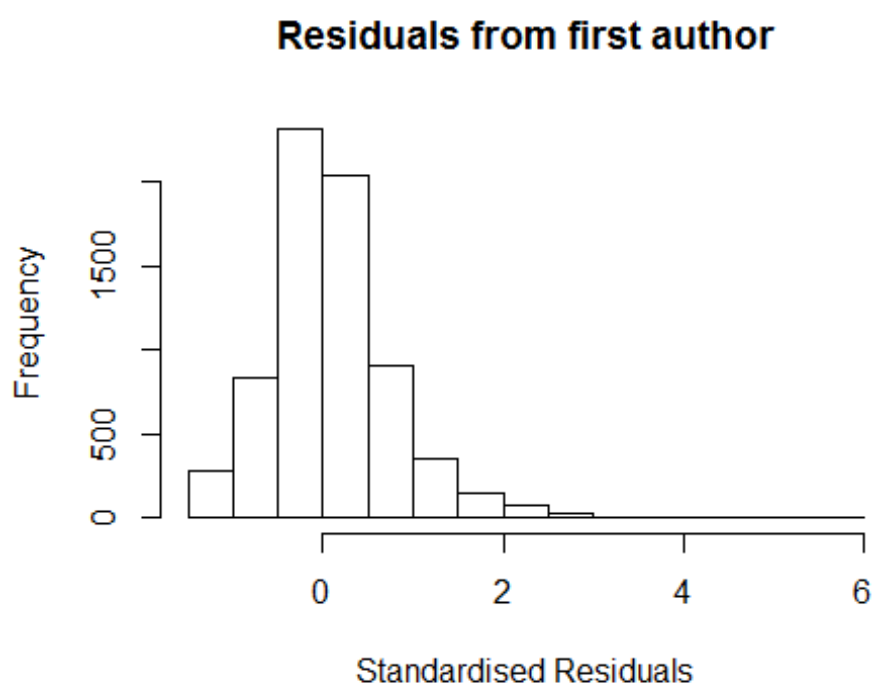
```
## Robustness weights:
##  46 observations
c(4,5,6,7,11,20,424,425,452,502,503,504,506,515,516,517,518,519,538,540,542,5
45,549,569,617,619,620,622,629,645,649,651,652,654,1145,1166,1203,1204,1225,1
261,1384,1758,1768,1834,1895,1905)
##   are outliers with |weight| = 0 ( < 1.4e-05);
##  450 weights are ~= 1. The remaining 6492 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.000   0.848   0.940   0.872   0.980   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.43e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.118  1           1.057
## LastAuthorFemale  1.136  1           1.066
## Year              1.070 16           1.002
```




```
## [1] "List of  30 outliers with residuals above  2.5"
##         ScopusId  NLCS Year OneField Fields residuals
## 18    0008215563 2.763 1996     1101      1     2.511
## 75    0030285171 3.319 1996     1101      1     2.949
## 104   0030324448 3.125 1996     1101      1     2.764
## 111   0030431384 4.318 1996     1101      1     3.948
## 174   0001119747 3.118 1996     1101      3     2.748
## 3664  0029783016 3.087 1996     1100      2     2.717
## 3665  0029823855 3.316 1996     1100      2     2.946
## 3803  0002054008 4.235 1996     1101      1     3.865
## 3805  0030319283 3.487 1996     1101      1     3.117
## 3806  0030378061 3.125 1996     1101      1     2.755
## 3822  0000557865 3.053 1996     1101      2     2.683
## 3825  0000878057 3.223 1996     1100      2     2.853
## 3866  0030390658 2.948 1996     1101      2     2.578
## 3920  0029679856 3.073 1996     1101      3     2.821
## 3999  0000297493 3.185 1996     1100      6     2.815
## 4003  0000445638 3.022 1996     1100      6     2.652
## 4009  0000904732 2.942 1996     1100      6     2.572
## 4067  0031393667 3.337 1997     1101      1     3.012
## 4071  0031419628 6.119 1997     1101      1     5.794
## 4073  0031428001 4.380 1997     1101      1     3.928
## 4075  0031429035 3.337 1997     1101      1     2.885
## 4078 16944367296 4.052 1997     1101      1     3.600
## 6685  0030893266 3.212 1997     1100      5     2.760
## 6755  0030898875 2.970 1997     1100      2     2.518
## 6757  0030913496 3.112 1997     1100      2     2.660
```

```
## 6800  0031411844 3.117 1997     1101      2     2.665
## 8838  0032579440 3.265 1998     1101      6     2.746
## 9004  0032160863 3.247 1998     1101      3     2.728
## 9085  0000911137 3.031 1998     1100      6     2.503
## 9101  0002799511 3.102 1998     1100      6     2.574
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.17052 -0.37688  0.00704  0.42767  5.79367
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         0.36954    0.03564   10.37  < 2e-16 ***
## FirstAuthorFemale1 -0.00897    0.01700   -0.53   0.5975
## LastAuthorFemale1  -0.11740    0.01876   -6.26  4.1e-10 ***
## Year1997            0.08217    0.04480    1.83   0.0667 .
## Year1998            0.15813    0.04841    3.27   0.0011 **
## Year1999            0.66214    0.05044   13.13  < 2e-16 ***
## Year2000            0.68657    0.04861   14.12  < 2e-16 ***
## Year2001            0.74005    0.04663   15.87  < 2e-16 ***
## Year2002            0.71462    0.04440   16.10  < 2e-16 ***
## Year2003            0.74064    0.04327   17.12  < 2e-16 ***
## Year2004            0.68413    0.04508   15.18  < 2e-16 ***
## Year2005            0.79016    0.04104   19.25  < 2e-16 ***
## Year2006            0.80098    0.04105   19.51  < 2e-16 ***
## Year2007            0.75858    0.04278   17.73  < 2e-16 ***
## Year2008            0.75934    0.04382   17.33  < 2e-16 ***
## Year2009            0.73211    0.05441   13.46  < 2e-16 ***
## Year2010            0.70802    0.05264   13.45  < 2e-16 ***
## Year2011            0.69094    0.05718   12.08  < 2e-16 ***
## Year2012            0.61011    0.05691   10.72  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.509
## Multiple R-squared:  0.209,  Adjusted R-squared:  0.207
## Convergence in 16 IRWLS iterations
##
## Robustness weights:
##  43 observations
c(4,5,7,11,20,403,424,425,452,502,503,504,515,516,517,518,519,541,545,549,569
,617,619,622,629,645,649,651,652,654,1145,1166,1203,1204,1209,1225,1228,1384,
1758,1834,1895,1905,5242)
##   are outliers with |weight| <= 1.3e-05 ( < 1.4e-05);
##  458 weights are ~= 1. The remaining 6487 ones are summarized as
```

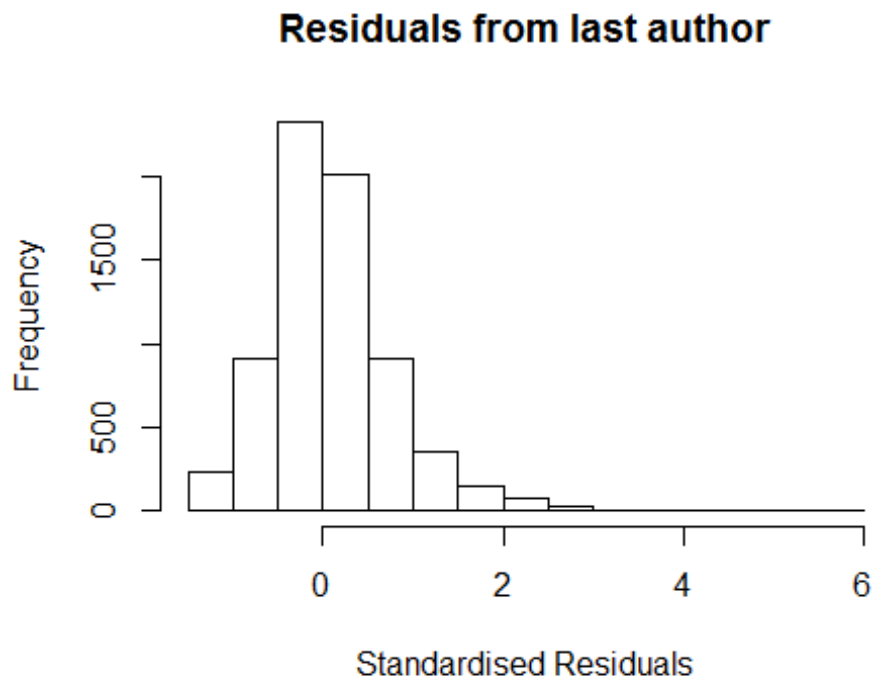
```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0004  0.8460  0.9370  0.8700  0.9780  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.43e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.036  1           1.018
## Year              1.036 16           1.001
```




```
## [1] "List of  30 outliers with residuals above  2.5"
##         ScopusId  NLCS Year OneField Fields residuals
## 18    0008215563 2.763 1996     1101      1     2.511
## 75    0030285171 3.319 1996     1101      1     2.949
```

```
## 104   0030324448 3.125 1996     1101      1     2.764
## 111   0030431384 4.318 1996     1101      1     3.948
## 174   0001119747 3.118 1996     1101      3     2.748
## 3664  0029783016 3.087 1996     1100      2     2.717
## 3665  0029823855 3.316 1996     1100      2     2.946
## 3803  0002054008 4.235 1996     1101      1     3.865
## 3805  0030319283 3.487 1996     1101      1     3.117
## 3806  0030378061 3.125 1996     1101      1     2.755
## 3822  0000557865 3.053 1996     1101      2     2.683
## 3825  0000878057 3.223 1996     1100      2     2.853
## 3866  0030390658 2.948 1996     1101      2     2.578
## 3920  0029679856 3.073 1996     1101      3     2.821
## 3999  0000297493 3.185 1996     1100      6     2.815
## 4003  0000445638 3.022 1996     1100      6     2.652
## 4009  0000904732 2.942 1996     1100      6     2.572
## 4067  0031393667 3.337 1997     1101      1     3.012
## 4071  0031419628 6.119 1997     1101      1     5.794
## 4073  0031428001 4.380 1997     1101      1     3.928
## 4075  0031429035 3.337 1997     1101      1     2.885
## 4078 16944367296 4.052 1997     1101      1     3.600
## 6685  0030893266 3.212 1997     1100      5     2.760
## 6755  0030898875 2.970 1997     1100      2     2.518
## 6757  0030913496 3.112 1997     1100      2     2.660
## 6800  0031411844 3.117 1997     1101      2     2.665
## 8838  0032579440 3.265 1998     1101      6     2.746
## 9004  0032160863 3.247 1998     1101      3     2.728
## 9085  0000911137 3.031 1998     1100      6     2.503
## 9101  0002799511 3.102 1998     1100      6     2.574
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.1536 -0.3830  0.0113  0.4205  5.7360
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)          0.3567     0.0354   10.07   <2e-16 ***
## FirstAuthorFemale1  -0.0461     0.0166   -2.78   0.0054 **
## Year1997             0.0724     0.0444    1.63   0.1030
## Year1998             0.1480     0.0481    3.08   0.0021 **
## Year1999             0.6625     0.0509   13.00   <2e-16 ***
## Year2000             0.6813     0.0489   13.93   <2e-16 ***
## Year2001             0.7381     0.0466   15.83   <2e-16 ***
## Year2002             0.7148     0.0444   16.09   <2e-16 ***
## Year2003             0.7416     0.0431   17.20   <2e-16 ***
## Year2004             0.6793     0.0452   15.02   <2e-16 ***
```

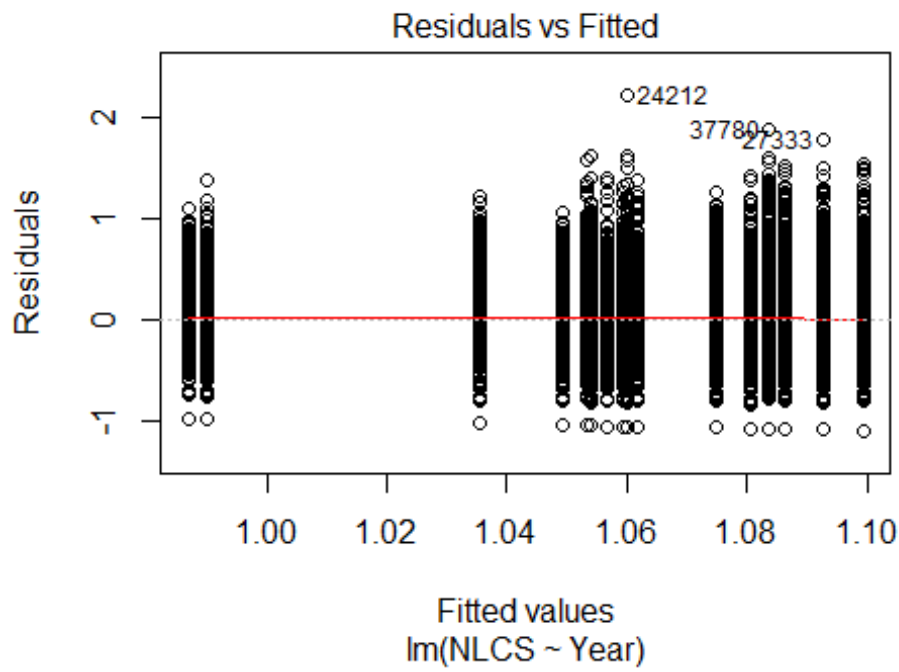
```
## Year2005             0.7893     0.0409   19.28   <2e-16 ***
## Year2006             0.7969     0.0411   19.39   <2e-16 ***
## Year2007             0.7582     0.0427   17.77   <2e-16 ***
## Year2008             0.7551     0.0438   17.24   <2e-16 ***
## Year2009             0.7298     0.0548   13.33   <2e-16 ***
## Year2010             0.7075     0.0530   13.34   <2e-16 ***
## Year2011             0.6886     0.0576   11.96   <2e-16 ***
## Year2012             0.6059     0.0578   10.49   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.509
## Multiple R-squared:  0.205,  Adjusted R-squared:  0.204
## Convergence in 16 IRWLS iterations
##
## Robustness weights:
##  45 observations
c(4,5,7,11,20,403,424,425,452,502,503,504,515,516,517,518,519,541,545,549,569
,617,619,622,629,645,646,649,651,652,654,1145,1166,1203,1204,1209,1225,1228,1
384,1758,1830,1834,1895,1905,5242)
##   are outliers with |weight| <= 2.8e-08 ( < 1.4e-05);
##  436 weights are ~= 1. The remaining 6507 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0001  0.8430  0.9360  0.8700  0.9760  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.43e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##                   GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.049  1           1.024
## Year             1.049 16           1.002
```



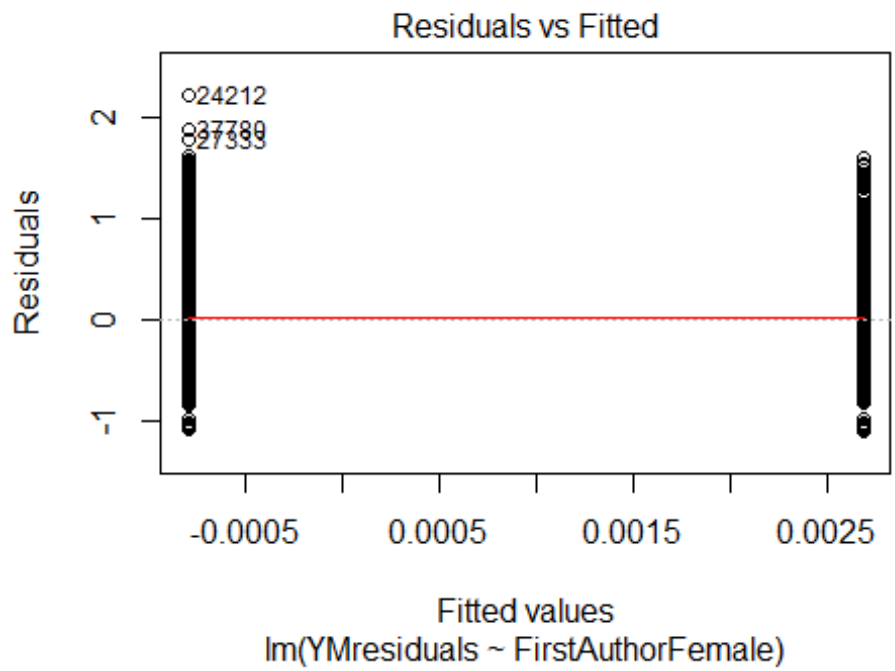

```
## [1] "List of  30 outliers with residuals above  2.5"
##         ScopusId  NLCS Year OneField Fields residuals
## 18    0008215563 2.763 1996     1101      1     2.511
## 75    0030285171 3.319 1996     1101      1     2.949
## 104   0030324448 3.125 1996     1101      1     2.764
## 111   0030431384 4.318 1996     1101      1     3.948
## 174   0001119747 3.118 1996     1101      3     2.748
## 3664  0029783016 3.087 1996     1100      2     2.717
## 3665  0029823855 3.316 1996     1100      2     2.946
## 3803  0002054008 4.235 1996     1101      1     3.865
## 3805  0030319283 3.487 1996     1101      1     3.117
## 3806  0030378061 3.125 1996     1101      1     2.755
## 3822  0000557865 3.053 1996     1101      2     2.683
## 3825  0000878057 3.223 1996     1100      2     2.853
## 3866  0030390658 2.948 1996     1101      2     2.578
## 3920  0029679856 3.073 1996     1101      3     2.821
## 3999  0000297493 3.185 1996     1100      6     2.815
## 4003  0000445638 3.022 1996     1100      6     2.652
## 4009  0000904732 2.942 1996     1100      6     2.572
## 4067  0031393667 3.337 1997     1101      1     3.012
## 4071  0031419628 6.119 1997     1101      1     5.794
## 4073  0031428001 4.380 1997     1101      1     3.928
## 4075  0031429035 3.337 1997     1101      1     2.885
## 4078 16944367296 4.052 1997     1101      1     3.600
## 6685  0030893266 3.212 1997     1100      5     2.760
## 6755  0030898875 2.970 1997     1100      2     2.518
## 6757  0030913496 3.112 1997     1100      2     2.660
```

```
## 6800  0031411844 3.117 1997     1101      2     2.665
## 8838  0032579440 3.265 1998     1101      6     2.746
## 9004  0032160863 3.247 1998     1101      3     2.728
## 9085  0000911137 3.031 1998     1100      6     2.503
## 9101  0002799511 3.102 1998     1100      6     2.574
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.16892 -0.37734  0.00452  0.42751  5.78991
##
## Coefficients:
##                   Estimate Std. Error t value Pr(>|t|)
## (Intercept)         0.3673     0.0351   10.46  < 2e-16 ***
## LastAuthorFemale1  -0.1207     0.0181   -6.69  2.5e-11 ***
## Year1997            0.0825     0.0448    1.84   0.0657 .
## Year1998            0.1585     0.0484    3.28   0.0011 **
## Year1999            0.6620     0.0504   13.13  < 2e-16 ***
## Year2000            0.6866     0.0486   14.12  < 2e-16 ***
## Year2001            0.7410     0.0465   15.94  < 2e-16 ***
## Year2002            0.7152     0.0443   16.13  < 2e-16 ***
## Year2003            0.7412     0.0432   17.15  < 2e-16 ***
## Year2004            0.6848     0.0450   15.20  < 2e-16 ***
## Year2005            0.7906     0.0410   19.29  < 2e-16 ***
## Year2006            0.8016     0.0410   19.56  < 2e-16 ***
## Year2007            0.7588     0.0427   17.75  < 2e-16 ***
## Year2008            0.7599     0.0437   17.37  < 2e-16 ***
## Year2009            0.7327     0.0543   13.48  < 2e-16 ***
## Year2010            0.7084     0.0526   13.46  < 2e-16 ***
## Year2011            0.6908     0.0572   12.07  < 2e-16 ***
## Year2012            0.6102     0.0569   10.72  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.509
## Multiple R-squared:  0.209,  Adjusted R-squared:  0.207
## Convergence in 16 IRWLS iterations
##
## Robustness weights:
##  43 observations
c(4,5,7,11,20,403,424,425,452,502,503,504,515,516,517,518,519,541,545,549,569
,617,619,622,629,645,649,651,652,654,1145,1166,1203,1204,1209,1225,1228,1384,
1758,1834,1895,1905,5242)
##   are outliers with |weight| <= 2.9e-06 ( < 1.4e-05);
##  464 weights are ~= 1. The remaining 6481 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
```

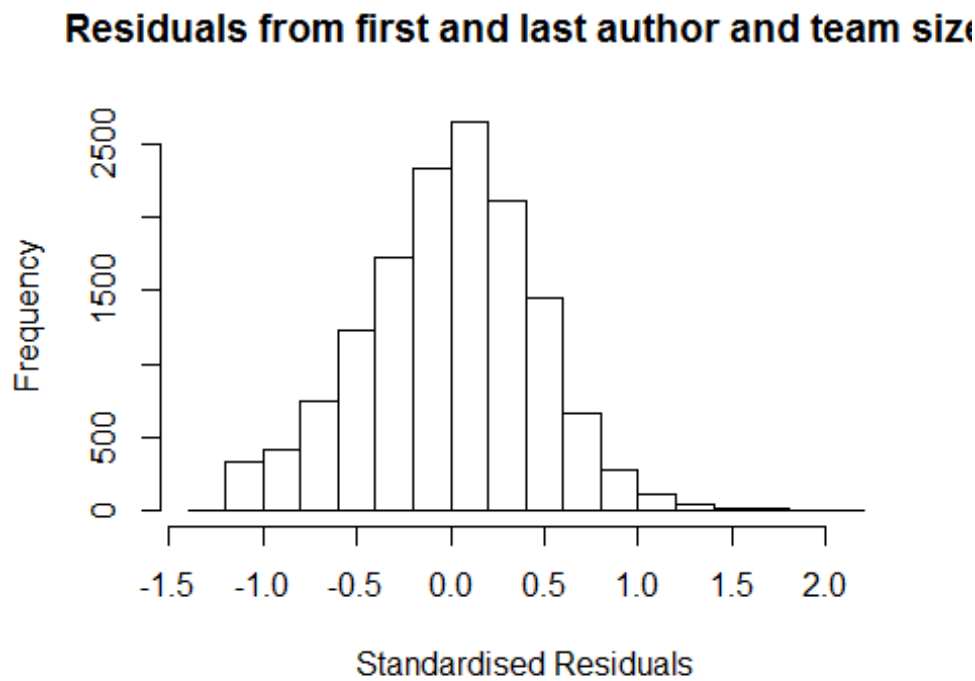
```
##  0.0003  0.8460  0.9370  0.8700  0.9780  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.43e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Sample size for the above analysis:  6988"
## [1] ""
## [1] ""
## [1] "##################################"
## [1] "Analysis of AJSC 1102"
## [1] "##################################"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 2068 1924 1797 1650 1787 1728 1745 1753 1880 1811 2077 1982 2138 2170 2140
## 2011 2012
## 2301 2318
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  952  864  808  811  671  578  859  807  963  881 1115 1150 1223 1193 1183
## 2011 2012
## 1254 1305
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  841  751  703  684  573  492  716  689  808  750  943  957 1020 1009 1002
## 2011 2012
## 1075 1091
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
##  Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 130, df = 16, p-value <2e-16
```




```
##
##  Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 12, df = 1, p-value = 6e-04
```

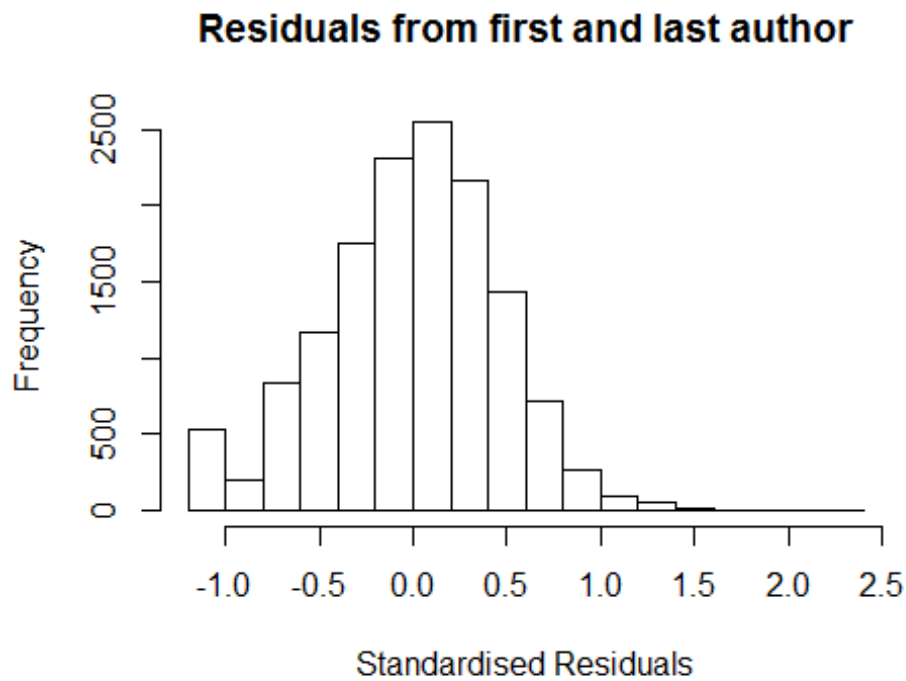



```
## [1] "Female first author team size 2018 geometric mean: 3.65588735191911"
## [1] "Male first author team size 2018 geometric mean: 3.60273785334978"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 150000, p-value = 0.7
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 3.41963867535791"
## [1] "Male last author team size 2018 geometric mean: 3.68402752661567"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 110000, p-value = 0.2
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.031  1           1.015
## LastAuthorFemale  1.028  1           1.014
## UniqueAuthors     1.071  4           1.009
## Year              1.085 16           1.003
```



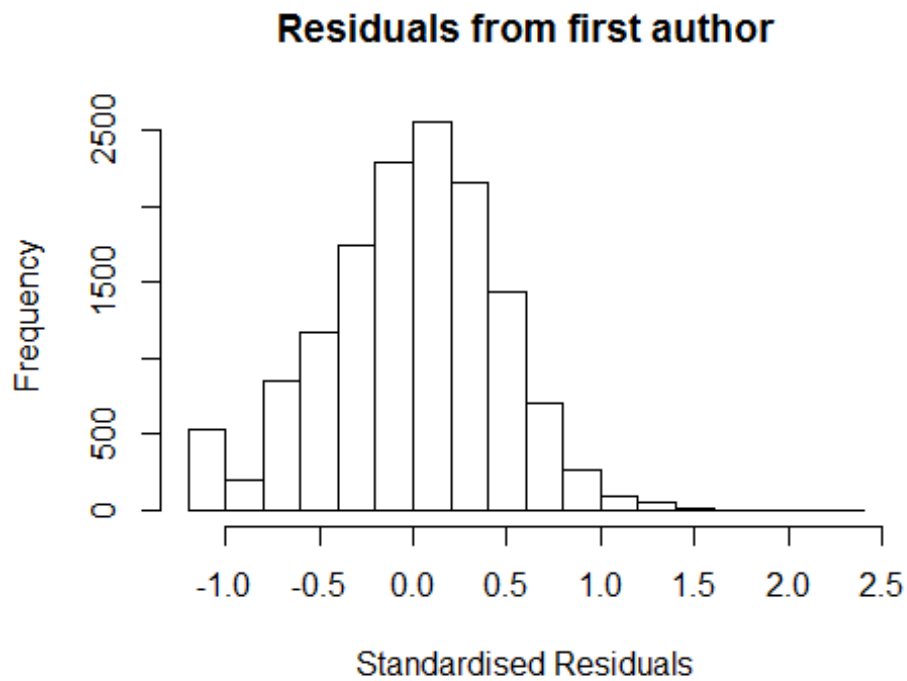

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
UniqueAuthors +
##     Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##     k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.2166 -0.2991  0.0174  0.3004  2.1000
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)         0.912340   0.020590   44.31  < 2e-16 ***
## FirstAuthorFemale1 -0.000386   0.009825   -0.04  0.96864
## LastAuthorFemale1   0.009924   0.010891    0.91  0.36218
## UniqueAuthors2      0.142615   0.015227    9.37  < 2e-16 ***
## UniqueAuthors3      0.149434   0.015141    9.87  < 2e-16 ***
## UniqueAuthors4      0.213987   0.016033   13.35  < 2e-16 ***
## UniqueAuthors5      0.283081   0.015862   17.85  < 2e-16 ***
## Year1997            0.006687   0.023322    0.29  0.77433
## Year1998            0.011633   0.023509    0.49  0.62072
## Year1999            0.010086   0.024307    0.41  0.67817
```

```
## Year2000            0.011209   0.026097    0.43  0.66756
## Year2001           -0.005511   0.026139   -0.21  0.83302
## Year2002           -0.024997   0.023530   -1.06  0.28811
## Year2003           -0.087125   0.023688   -3.68  0.00024 ***
## Year2004           -0.079535   0.022523   -3.53  0.00041 ***
## Year2005           -0.012123   0.022702   -0.53  0.59333
## Year2006           -0.029781   0.022170   -1.34  0.17919
## Year2007           -0.018386   0.022313   -0.82  0.40996
## Year2008            0.003348   0.022706    0.15  0.88279
## Year2009           -0.007368   0.022877   -0.32  0.74741
## Year2010           -0.011965   0.022808   -0.52  0.59987
## Year2011           -0.001697   0.022864   -0.07  0.94084
## Year2012           -0.013095   0.023621   -0.55  0.57932
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.442
## Multiple R-squared:  0.0363, Adjusted R-squared:  0.0347
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
##  observation 8511 is an outlier with |weight| = 0 ( < 7.1e-06);
##  1237 weights are ~= 1. The remaining 12866 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0095  0.8640  0.9490  0.9010  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          7.09e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.025  1           1.012
## LastAuthorFemale  1.027  1           1.013
## Year              1.021 16           1.001
```



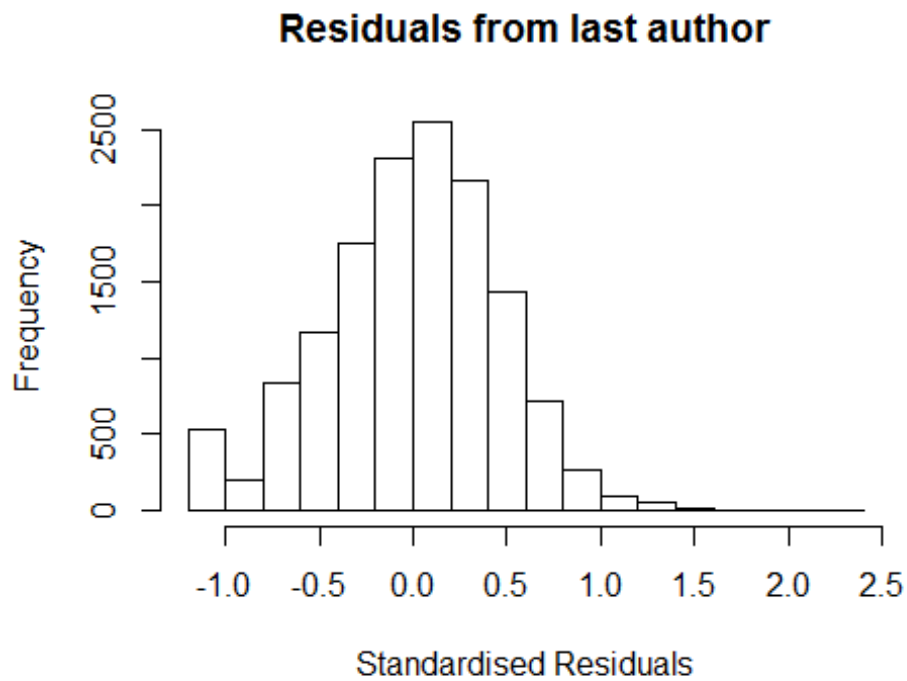

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.1137 -0.3045  0.0184  0.3056  2.2172
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.055970   0.017231   61.28   <2e-16 ***
## FirstAuthorFemale1  0.009150   0.009966    0.92   0.3586
## LastAuthorFemale1   0.008441   0.011070    0.76   0.4457
## Year1997            0.008495   0.023756    0.36   0.7207
## Year1998            0.013628   0.023843    0.57   0.5676
## Year1999            0.012366   0.024593    0.50   0.6151
## Year2000            0.016986   0.025998    0.65   0.5135
## Year2001            0.004076   0.026284    0.16   0.8768
## Year2002           -0.011246   0.023967   -0.47   0.6389
## Year2003           -0.074657   0.024214   -3.08   0.0021 **
## Year2004           -0.064853   0.023071   -2.81   0.0049 **
## Year2005            0.011939   0.023024    0.52   0.6041
```

```
## Year2006            0.000755   0.022522    0.03   0.9733
## Year2007            0.005349   0.022697    0.24   0.8137
## Year2008            0.032734   0.023179    1.41   0.1579
## Year2009            0.021323   0.023314    0.91   0.3604
## Year2010            0.016810   0.023269    0.72   0.4701
## Year2011            0.040168   0.023249    1.73   0.0841 .
## Year2012            0.021874   0.024102    0.91   0.3641
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.448
## Multiple R-squared:  0.00404,    Adjusted R-squared:  0.00277
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  observation 8511 is an outlier with |weight| = 0 ( < 7.1e-06);
##  1268 weights are ~= 1. The remaining 12835 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0725  0.8660  0.9490  0.9010  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          7.09e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##                   GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.01  1           1.005
## Year              1.01 16           1.000
```




```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.1078 -0.3055  0.0187  0.3055  2.2246
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.057006   0.017192   61.48   <2e-16 ***
## FirstAuthorFemale1  0.010516   0.009905    1.06    0.288
## Year1997            0.008639   0.023754    0.36    0.716
## Year1998            0.013578   0.023846    0.57    0.569
## Year1999            0.012143   0.024594    0.49    0.621
## Year2000            0.016992   0.025994    0.65    0.513
## Year2001            0.004164   0.026275    0.16    0.874
## Year2002           -0.011054   0.023959   -0.46    0.645
## Year2003           -0.074708   0.024206   -3.09    0.002 **
## Year2004           -0.064708   0.023071   -2.80    0.005 **
## Year2005            0.011624   0.023023    0.50    0.614
## Year2006            0.000686   0.022518    0.03    0.976
```

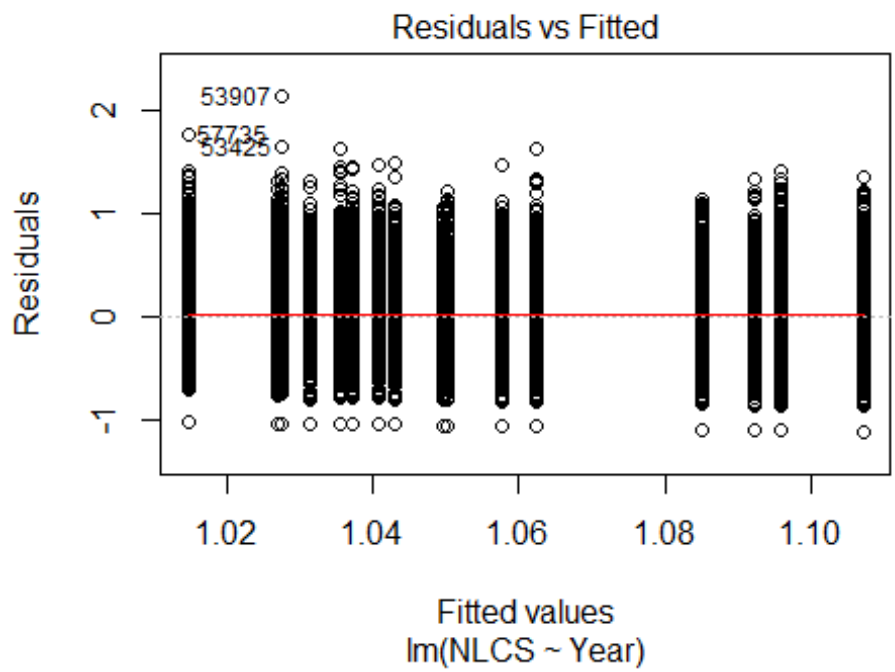
```
## Year2007            0.005417   0.022697    0.24    0.811
## Year2008            0.032822   0.023175    1.42    0.157
## Year2009            0.021579   0.023303    0.93    0.354
## Year2010            0.017247   0.023246    0.74    0.458
## Year2011            0.040302   0.023252    1.73    0.083 .
## Year2012            0.022111   0.024097    0.92    0.359
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.448
## Multiple R-squared:  0.004,  Adjusted R-squared:  0.0028
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  observation 8511 is an outlier with |weight| = 0 ( < 7.1e-06);
##  1275 weights are ~= 1. The remaining 12828 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0737  0.8660  0.9490  0.9010  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          7.09e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##                   GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.011  1           1.006
## Year             1.011 16           1.000
```



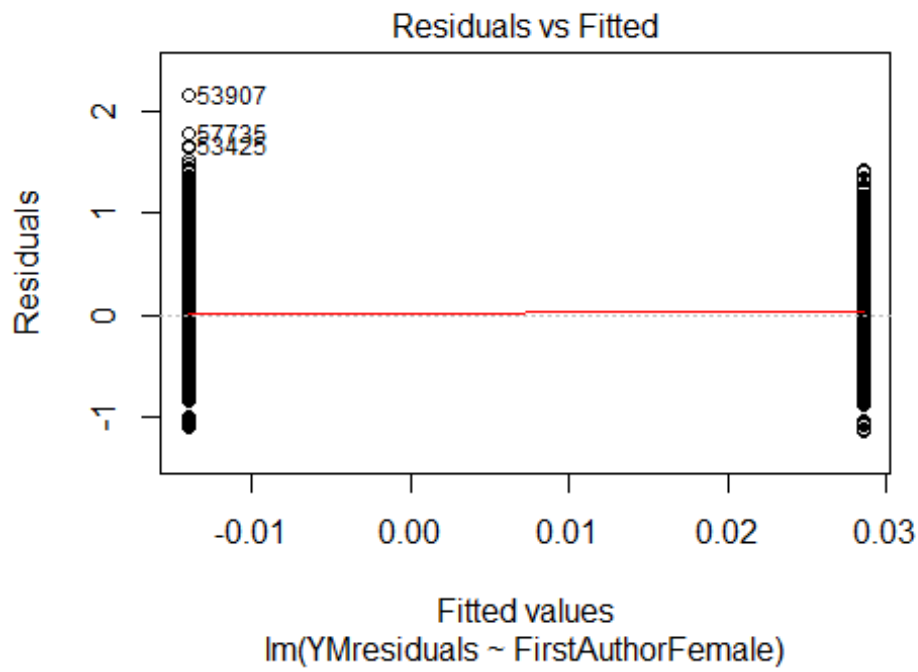

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.1084 -0.3058  0.0187  0.3057  2.2134
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.05744    0.01720   61.49   <2e-16 ***
## LastAuthorFemale1   0.01026    0.01100    0.93   0.3510
## Year1997            0.00838    0.02376    0.35   0.7243
## Year1998            0.01344    0.02383    0.56   0.5727
## Year1999            0.01229    0.02460    0.50   0.6173
## Year2000            0.01689    0.02599    0.65   0.5158
## Year2001            0.00433    0.02628    0.16   0.8690
## Year2002           -0.01111    0.02396   -0.46   0.6430
## Year2003           -0.07428    0.02421   -3.07   0.0022 **
## Year2004           -0.06463    0.02307   -2.80   0.0051 **
## Year2005            0.01232    0.02301    0.54   0.5922
## Year2006            0.00090    0.02252    0.04   0.9681
```

```
## Year2007           0.00591    0.02268    0.26   0.7943
## Year2008           0.03304    0.02318    1.43   0.1541
## Year2009           0.02144    0.02331    0.92   0.3576
## Year2010           0.01739    0.02326    0.75   0.4546
## Year2011           0.04073    0.02323    1.75   0.0795 .
## Year2012           0.02248    0.02408    0.93   0.3505
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.448
## Multiple R-squared:  0.00397,    Adjusted R-squared:  0.00277
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  observation 8511 is an outlier with |weight| = 0 ( < 7.1e-06);
##  1283 weights are ~= 1. The remaining 12820 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0734  0.8660  0.9490  0.9000  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          7.09e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Sample size for the above analysis:  14104"
## [1] ""
## [1] ""
## [1] "####################################"
## [1] "Analysis of AJSC 1103"
## [1] "####################################"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 3041 3067 3038 2923 2873 2859 2821 2680 2703 2805 3033 3224 3281 3165 3143
## 2011 2012
## 3043 3140
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 1600 1594 1695 1696 1131 1032 1511 1451 1459 1539 1722 1879 1949 1776 1891
```

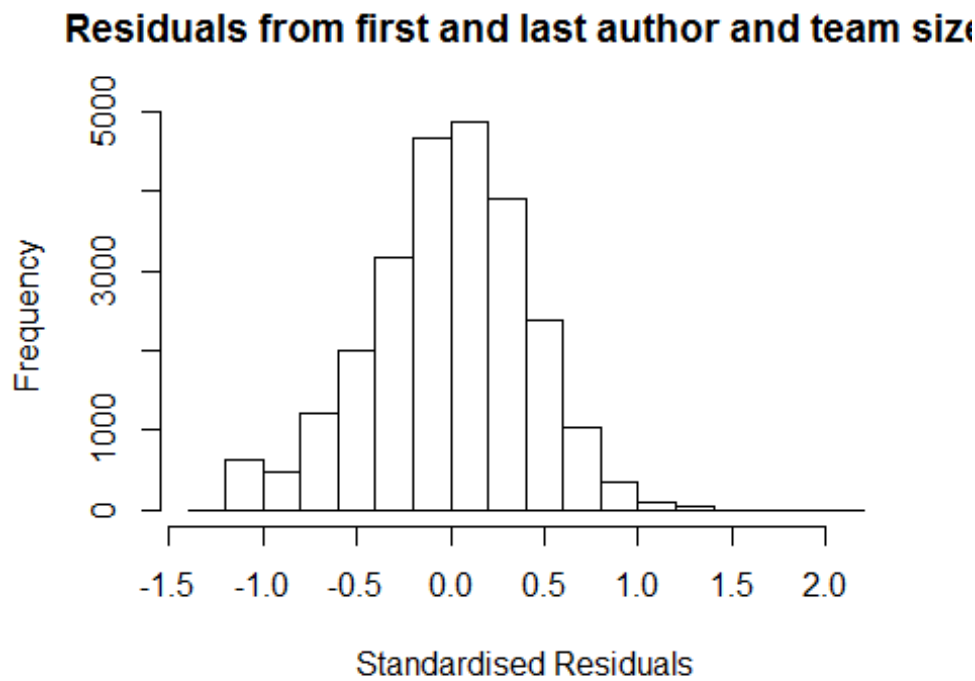
```
## 2011 2012
## 1850 1810
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 1453 1449 1535 1502 1014  925 1331 1278 1325 1403 1555 1708 1759 1595 1692
## 2011 2012
## 1688 1648
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
##  Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 160, df = 16, p-value <2e-16
```




```
##
##  Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 21, df = 1, p-value = 5e-06
```

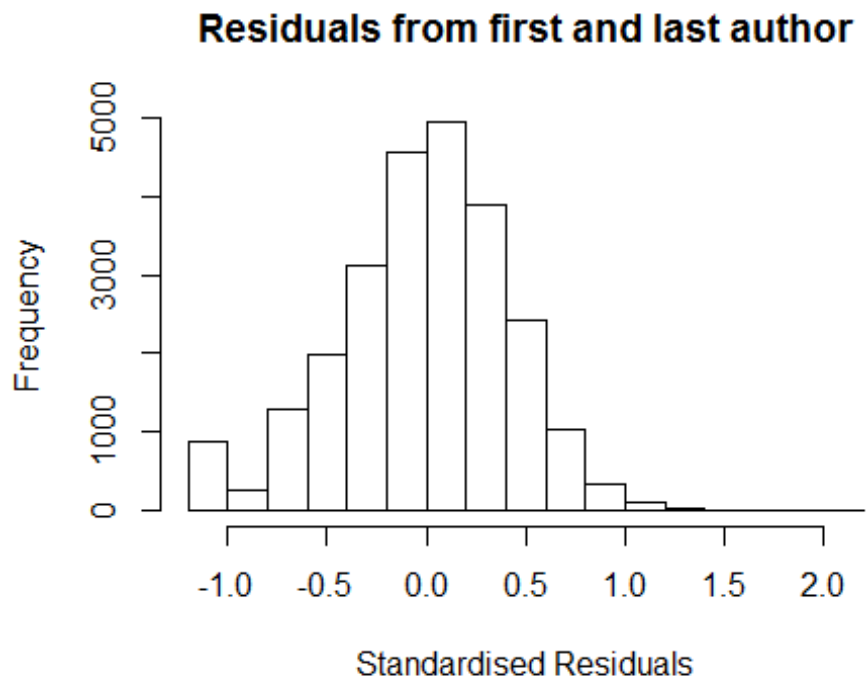



```
## [1] "Female first author team size 2018 geometric mean: 3.41122452563452"
## [1] "Male first author team size 2018 geometric mean: 3.01966113053655"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 410000, p-value = 6e-05
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 3.14745682141362"
## [1] "Male last author team size 2018 geometric mean: 3.23578283281117"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 310000, p-value = 0.5
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##                     GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.062  1           1.030
## LastAuthorFemale  1.052  1           1.026
## UniqueAuthors     1.047  4           1.006
## Year              1.054 16           1.002
```



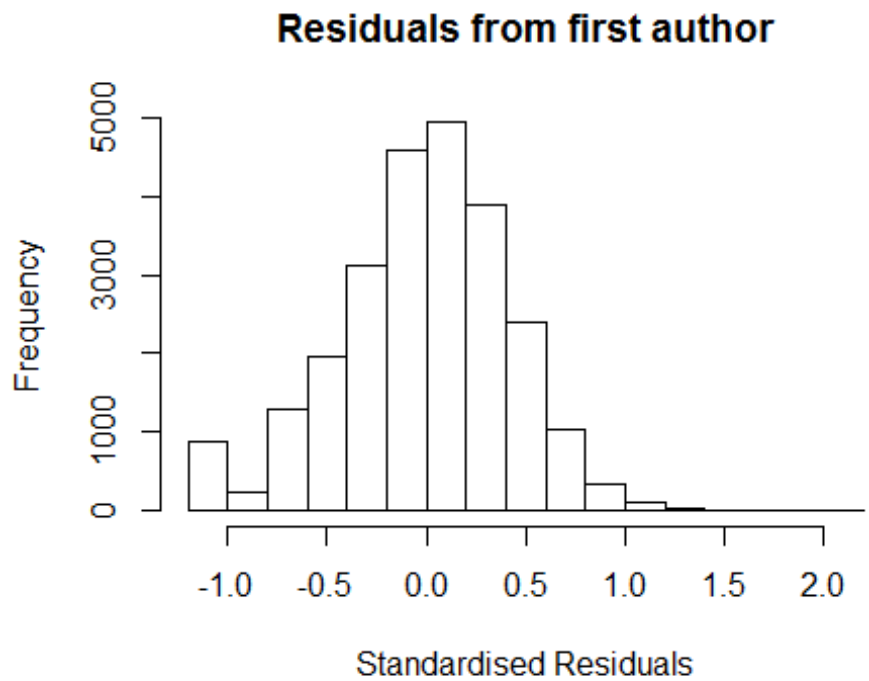

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
UniqueAuthors +
##     Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##     k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.2305 -0.2705  0.0119  0.2736  2.1700
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.04442    0.01347   77.53  < 2e-16 ***
## FirstAuthorFemale1  0.03831    0.00587    6.53  6.7e-11 ***
## LastAuthorFemale1   0.02113    0.00645    3.27   0.0011 **
## UniqueAuthors2      0.08441    0.00833   10.13  < 2e-16 ***
## UniqueAuthors3      0.11269    0.00885   12.73  < 2e-16 ***
## UniqueAuthors4      0.12531    0.00987   12.70  < 2e-16 ***
## UniqueAuthors5      0.14781    0.00994   14.87  < 2e-16 ***
## Year1997           -0.02530    0.01713   -1.48   0.1398
## Year1998           -0.03034    0.01658   -1.83   0.0674 .
## Year1999           -0.06492    0.01655   -3.92  8.8e-05 ***
```

```
## Year2000           -0.02166    0.01807   -1.20   0.2307
## Year2001           -0.02919    0.01849   -1.58   0.1145
## Year2002           -0.08730    0.01668   -5.23  1.7e-07 ***
## Year2003           -0.08687    0.01643   -5.29  1.3e-07 ***
## Year2004           -0.07839    0.01622   -4.83  1.4e-06 ***
## Year2005           -0.10236    0.01608   -6.37  2.0e-10 ***
## Year2006           -0.10104    0.01626   -6.21  5.2e-10 ***
## Year2007           -0.11564    0.01577   -7.33  2.3e-13 ***
## Year2008           -0.10431    0.01568   -6.65  2.9e-11 ***
## Year2009           -0.10814    0.01623   -6.66  2.8e-11 ***
## Year2010           -0.12479    0.01603   -7.78  7.2e-15 ***
## Year2011           -0.12878    0.01656   -7.78  7.6e-15 ***
## Year2012           -0.13652    0.01754   -7.79  7.2e-15 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.404
## Multiple R-squared:  0.0229, Adjusted R-squared:  0.022
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  observation 22904 is an outlier with |weight| = 0 ( < 4e-06);
##  2102 weights are ~= 1. The remaining 22757 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0156  0.8630  0.9500  0.8980  0.9860  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          4.02e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                    psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.058  1           1.029
## LastAuthorFemale  1.052  1           1.026
## Year              1.017 16           1.001
```



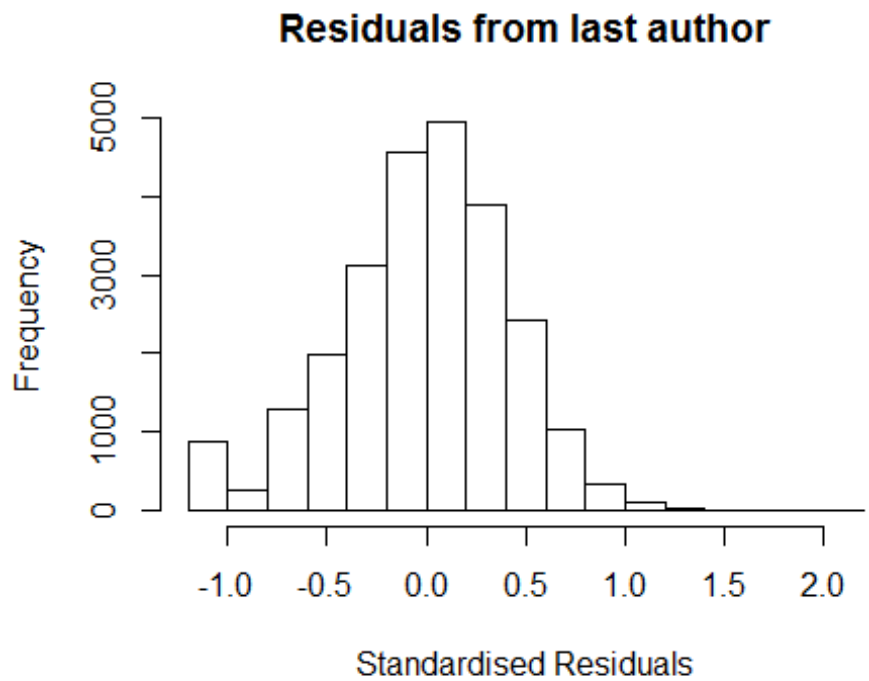

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.1822 -0.2745  0.0148  0.2729  2.1609
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)          1.11585    0.01219   91.54  < 2e-16 ***
## FirstAuthorFemale1   0.04753    0.00588    8.09  6.2e-16 ***
## LastAuthorFemale1    0.01882    0.00647    2.91  0.00361 **
## Year1997            -0.02124    0.01703   -1.25  0.21239
## Year1998            -0.02573    0.01653   -1.56  0.11941
## Year1999            -0.05973    0.01643   -3.64  0.00028 ***
## Year2000            -0.01694    0.01796   -0.94  0.34584
## Year2001            -0.02135    0.01855   -1.15  0.24983
## Year2002            -0.07564    0.01653   -4.57  4.8e-06 ***
## Year2003            -0.07247    0.01637   -4.43  9.6e-06 ***
## Year2004            -0.06135    0.01614   -3.80  0.00014 ***
## Year2005            -0.08640    0.01600   -5.40  6.8e-08 ***
```

```
## Year2006           -0.08534    0.01619   -5.27  1.4e-07 ***
## Year2007           -0.09898    0.01570   -6.30  3.0e-10 ***
## Year2008           -0.08766    0.01555   -5.64  1.8e-08 ***
## Year2009           -0.09188    0.01624   -5.66  1.5e-08 ***
## Year2010           -0.10529    0.01596   -6.60  4.3e-11 ***
## Year2011           -0.10675    0.01652   -6.46  1.1e-10 ***
## Year2012           -0.11059    0.01748   -6.33  2.6e-10 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.405
## Multiple R-squared:  0.00958,    Adjusted R-squared:  0.00886
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  observation 22904 is an outlier with |weight| = 0 ( < 4e-06);
##  2105 weights are ~= 1. The remaining 22754 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0132  0.8640  0.9500  0.8970  0.9860  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          4.02e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.013  1           1.006
## Year              1.013 16           1.000
```




```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##    Min     1Q Median     3Q    Max
## -1.171 -0.274  0.014  0.273  2.157
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)          1.11880    0.01214   92.13  < 2e-16 ***
## FirstAuthorFemale1   0.05185    0.00575    9.02  < 2e-16 ***
## Year1997            -0.02095    0.01703   -1.23  0.21873
## Year1998            -0.02572    0.01653   -1.56  0.11968
## Year1999            -0.05978    0.01643   -3.64  0.00027 ***
## Year2000            -0.01711    0.01796   -0.95  0.34092
## Year2001            -0.02086    0.01855   -1.13  0.26059
## Year2002            -0.07485    0.01653   -4.53  6.0e-06 ***
## Year2003            -0.07225    0.01637   -4.41  1.0e-05 ***
## Year2004            -0.06121    0.01614   -3.79  0.00015 ***
## Year2005            -0.08643    0.01600   -5.40  6.7e-08 ***
## Year2006            -0.08468    0.01619   -5.23  1.7e-07 ***
```

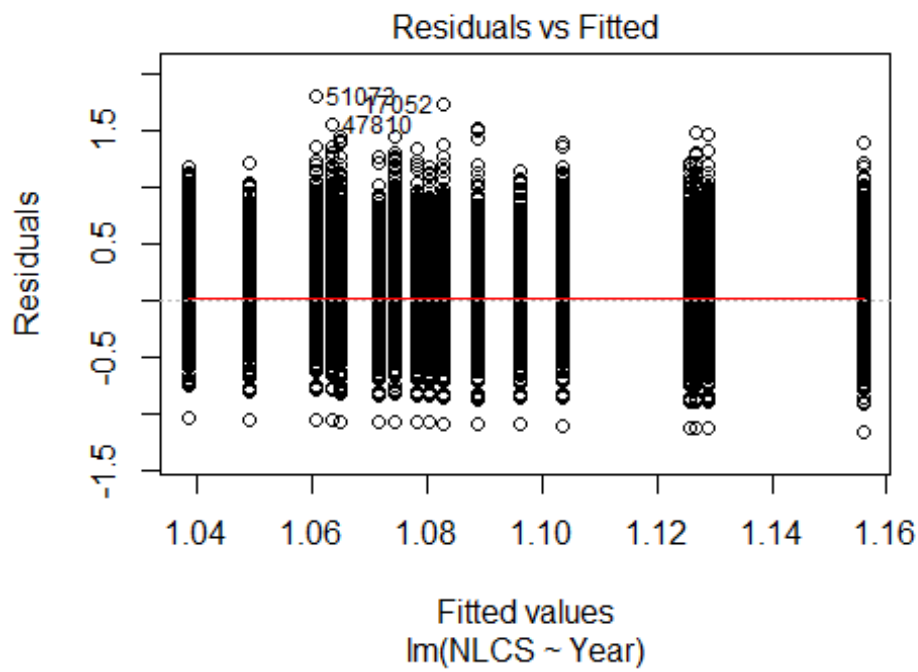
```
## Year2007           -0.09839    0.01571   -6.26  3.8e-10 ***
## Year2008           -0.08744    0.01555   -5.62  1.9e-08 ***
## Year2009           -0.09161    0.01623   -5.64  1.7e-08 ***
## Year2010           -0.10515    0.01595   -6.59  4.4e-11 ***
## Year2011           -0.10630    0.01653   -6.43  1.3e-10 ***
## Year2012           -0.11006    0.01748   -6.30  3.1e-10 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.405
## Multiple R-squared:  0.00921,    Adjusted R-squared:  0.00853
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  observation 22904 is an outlier with |weight| = 0 ( < 4e-06);
##  2117 weights are ~= 1. The remaining 22742 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0141  0.8640  0.9500  0.8970  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          4.02e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##                   GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.007  1           1.003
## Year             1.007 16           1.000
```



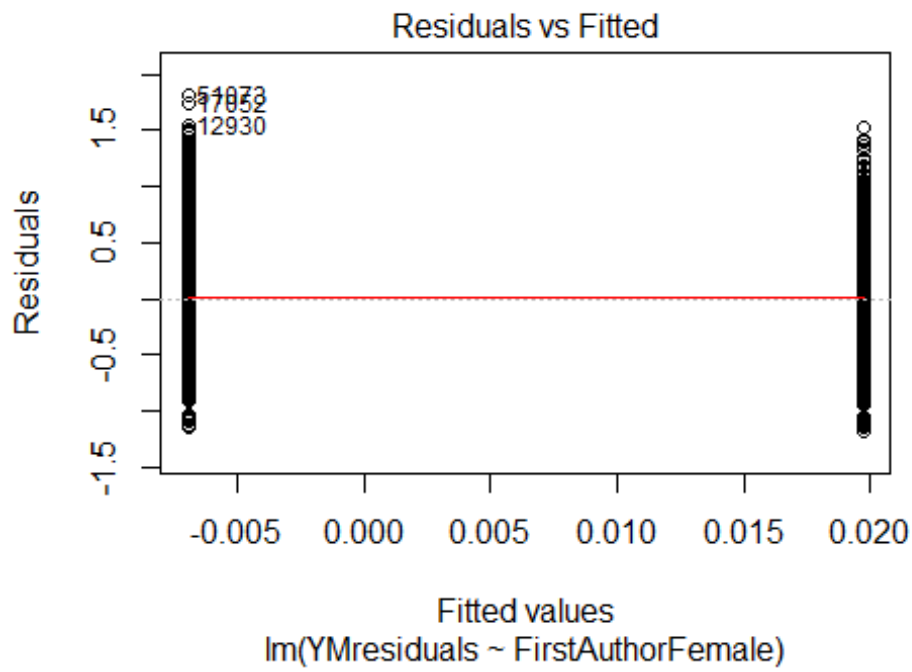

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.1575 -0.2743  0.0142  0.2735  2.1473
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.12570    0.01210   93.03  < 2e-16 ***
## LastAuthorFemale1   0.03176    0.00633    5.02  5.3e-07 ***
## Year1997           -0.02115    0.01701   -1.24  0.21376
## Year1998           -0.02563    0.01650   -1.55  0.12035
## Year1999           -0.05970    0.01644   -3.63  0.00028 ***
## Year2000           -0.01534    0.01796   -0.85  0.39326
## Year2001           -0.02026    0.01852   -1.09  0.27401
## Year2002           -0.07324    0.01652   -4.43  9.3e-06 ***
## Year2003           -0.07057    0.01637   -4.31  1.6e-05 ***
## Year2004           -0.05779    0.01613   -3.58  0.00034 ***
## Year2005           -0.08369    0.01600   -5.23  1.7e-07 ***
## Year2006           -0.08220    0.01618   -5.08  3.8e-07 ***
```

```
## Year2007          -0.09423    0.01568   -6.01  1.9e-09 ***
## Year2008          -0.08311    0.01554   -5.35  9.0e-08 ***
## Year2009          -0.08620    0.01623   -5.31  1.1e-07 ***
## Year2010          -0.10122    0.01596   -6.34  2.3e-10 ***
## Year2011          -0.10301    0.01652   -6.24  4.6e-10 ***
## Year2012          -0.10513    0.01746   -6.02  1.7e-09 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.405
## Multiple R-squared:  0.00693,    Adjusted R-squared:  0.00625
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  observation 22904 is an outlier with |weight| = 0 ( < 4e-06);
##  2088 weights are ~= 1. The remaining 22771 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0177  0.8630  0.9500  0.8980  0.9860  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          4.02e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Sample size for the above analysis:  24860"
## [1] ""
## [1] ""
## [1] "####################################"
## [1] "Analysis of AJSC 1104"
## [1] "####################################"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 2621 2743 2649 2777 2639 3011 2923 2882 2661 2851 3041 3113 3207 3056 2987
## 2011 2012
## 2938 3106
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 1727 1728 1714 1758 1444 1468 2042 1896 1858 2021 2136 2191 2231 2156 2028
```

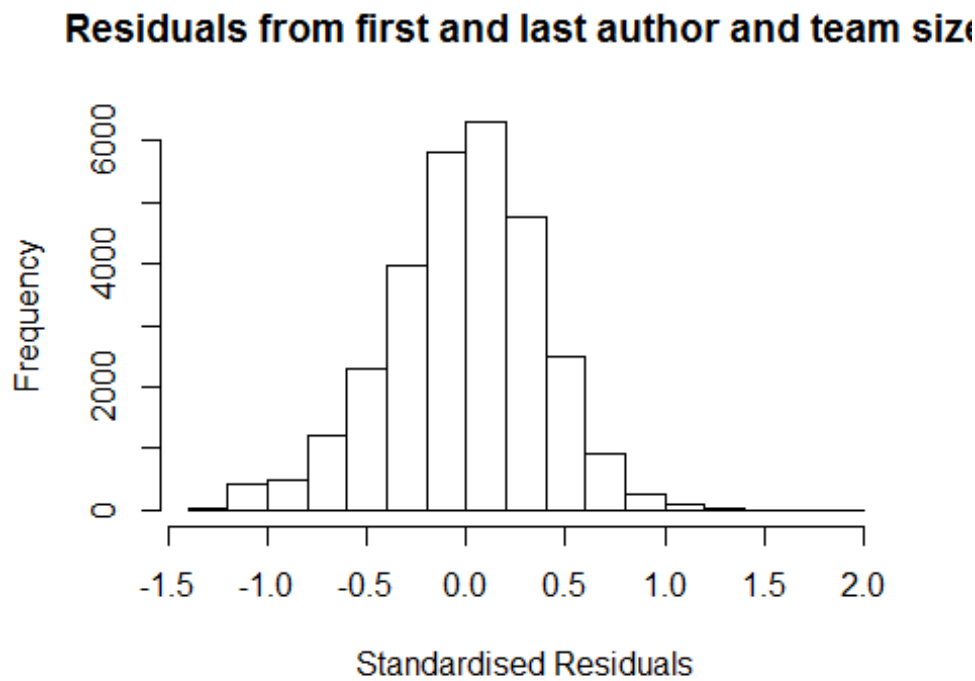
```
## 2011 2012
## 2072 2270
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 1561 1569 1554 1577 1292 1289 1772 1673 1661 1806 1871 1948 1963 1924 1804
## 2011 2012
## 1826 2036
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
##  Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 210, df = 16, p-value <2e-16
```




```
##
##  Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 71, df = 1, p-value <2e-16
```

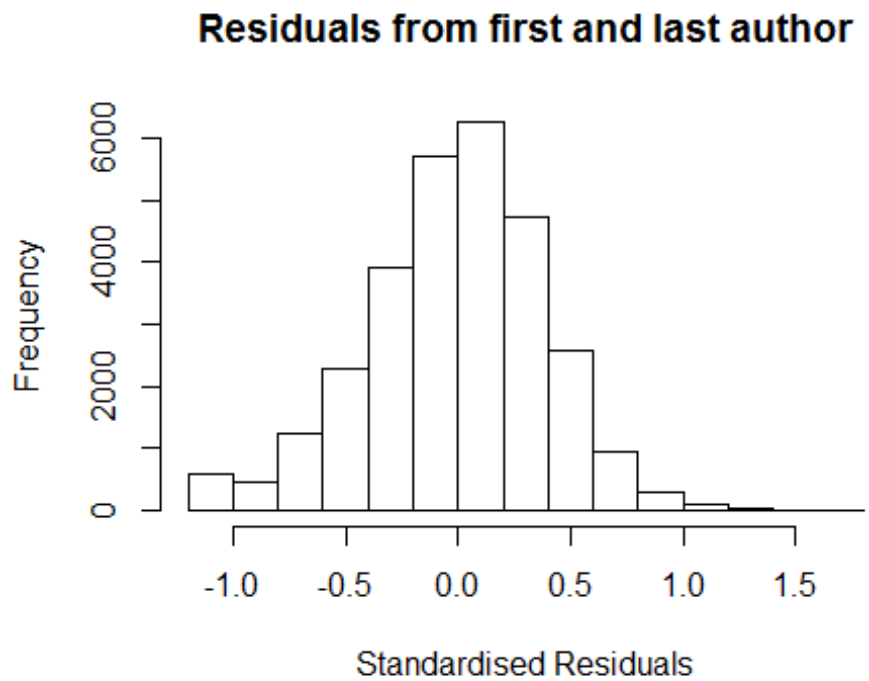



```
## [1] "Female first author team size 2018 geometric mean: 3.73679762306509"
## [1] "Male first author team size 2018 geometric mean: 3.49857667726197"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 460000, p-value = 0.2
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 3.73408706865361"
## [1] "Male last author team size 2018 geometric mean: 3.53720490142217"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 390000, p-value = 0.03
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.040  1           1.020
## LastAuthorFemale  1.025  1           1.012
## UniqueAuthors     1.049  4           1.006
## Year              1.062 16           1.002
```



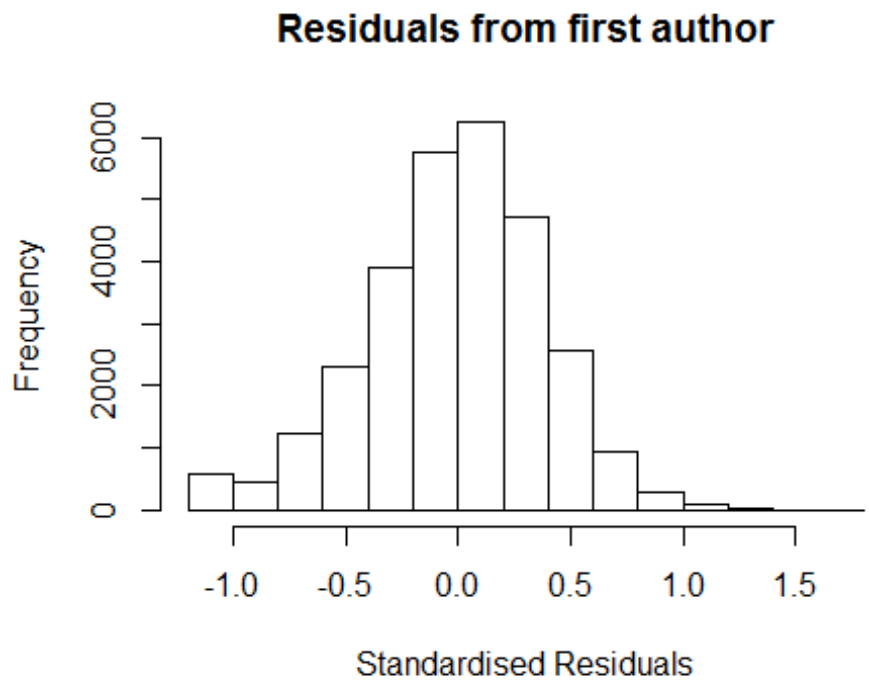

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
UniqueAuthors +
##     Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##     k.max = 1000))
##  \--> method = "MM"
## Residuals:
##    Min     1Q Median     3Q    Max
## -1.279 -0.249  0.011  0.248  1.838
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)          1.06571    0.01254   85.00  < 2e-16 ***
## FirstAuthorFemale1   0.02372    0.00501    4.73  2.2e-06 ***
## LastAuthorFemale1    0.00640    0.00564    1.14   0.2562
## UniqueAuthors2       0.09535    0.00826   11.54  < 2e-16 ***
## UniqueAuthors3       0.12138    0.00851   14.26  < 2e-16 ***
## UniqueAuthors4       0.15419    0.00917   16.81  < 2e-16 ***
## UniqueAuthors5       0.24427    0.00896   27.28  < 2e-16 ***
## Year1997            -0.03141    0.01494   -2.10   0.0356 *
## Year1998            -0.03540    0.01501   -2.36   0.0183 *
## Year1999            -0.04481    0.01458   -3.07   0.0021 **
```

```
## Year2000           -0.08663    0.01518   -5.71  1.2e-08 ***
## Year2001           -0.08894    0.01583   -5.62  1.9e-08 ***
## Year2002           -0.07285    0.01444   -5.05  4.5e-07 ***
## Year2003           -0.09308    0.01373   -6.78  1.3e-11 ***
## Year2004           -0.11744    0.01387   -8.47  < 2e-16 ***
## Year2005           -0.13985    0.01365  -10.24  < 2e-16 ***
## Year2006           -0.11812    0.01358   -8.70  < 2e-16 ***
## Year2007           -0.12043    0.01348   -8.94  < 2e-16 ***
## Year2008           -0.13435    0.01365   -9.84  < 2e-16 ***
## Year2009           -0.12784    0.01382   -9.25  < 2e-16 ***
## Year2010           -0.13644    0.01452   -9.40  < 2e-16 ***
## Year2011           -0.13185    0.01425   -9.25  < 2e-16 ***
## Year2012           -0.16295    0.01440  -11.32  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.366
## Multiple R-squared:  0.0396, Adjusted R-squared:  0.0389
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  2 observations c(8264,25673) are outliers with |weight| = 0 ( < 3.4e-06);
##  2476 weights are ~= 1. The remaining 26648 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0363  0.8670  0.9500  0.8980  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          3.43e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.035  1           1.018
## LastAuthorFemale  1.023  1           1.012
## Year              1.017 16           1.001
```



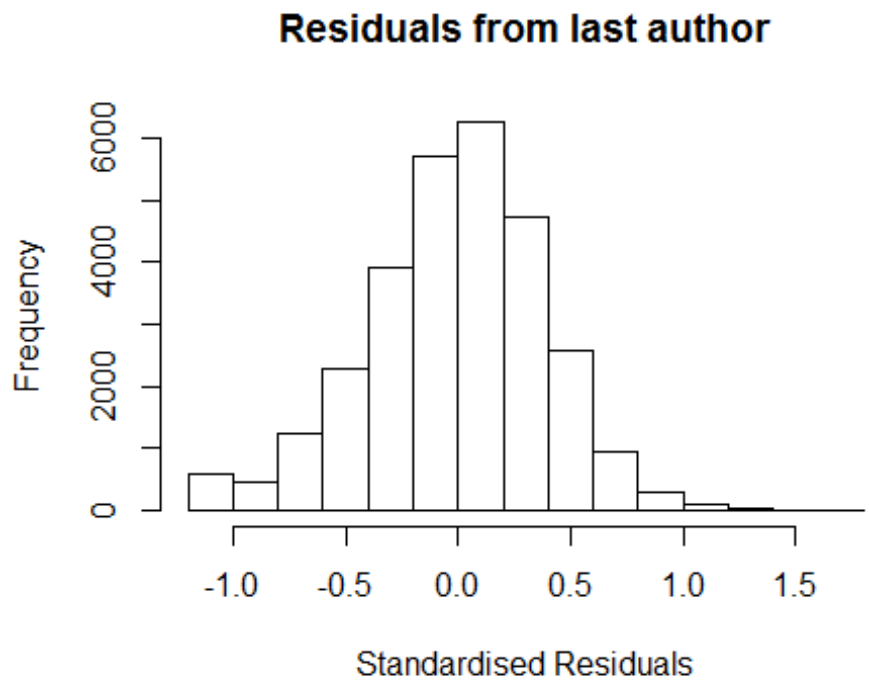

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.1994 -0.2507  0.0114  0.2501  1.7949
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)          1.16571    0.01084  107.51  < 2e-16 ***
## FirstAuthorFemale1   0.02778    0.00508    5.47  4.4e-08 ***
## LastAuthorFemale1    0.00588    0.00573    1.03    0.305
## Year1997            -0.03506    0.01505   -2.33    0.020 *
## Year1998            -0.02955    0.01504   -1.96    0.050 *
## Year1999            -0.03664    0.01465   -2.50    0.012 *
## Year2000            -0.07827    0.01523   -5.14  2.8e-07 ***
## Year2001            -0.07393    0.01606   -4.60  4.2e-06 ***
## Year2002            -0.05858    0.01459   -4.01  6.0e-05 ***
## Year2003            -0.07344    0.01383   -5.31  1.1e-07 ***
## Year2004            -0.09401    0.01395   -6.74  1.6e-11 ***
## Year2005            -0.11727    0.01377   -8.52  < 2e-16 ***
```

```
## Year2006           -0.08960    0.01362   -6.58  4.8e-11 ***
## Year2007           -0.09482    0.01355   -7.00  2.7e-12 ***
## Year2008           -0.10302    0.01374   -7.50  6.7e-14 ***
## Year2009           -0.09229    0.01390   -6.64  3.2e-11 ***
## Year2010           -0.10033    0.01455   -6.90  5.5e-12 ***
## Year2011           -0.09459    0.01427   -6.63  3.5e-11 ***
## Year2012           -0.12276    0.01446   -8.49  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.371
## Multiple R-squared:  0.00768,    Adjusted R-squared:  0.00707
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  observation 25673 is an outlier with |weight| = 0 ( < 3.4e-06);
##  2410 weights are ~= 1. The remaining 26715 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0004  0.8670  0.9500  0.8980  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          3.43e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.015  1           1.007
## Year              1.015 16           1.000
```




```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.1952 -0.2504  0.0113  0.2506  1.7939
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)          1.16661    0.01079  108.07  < 2e-16 ***
## FirstAuthorFemale1  0.02863    0.00503    5.70  1.2e-08 ***
## Year1997           -0.03519    0.01505   -2.34    0.019 *
## Year1998           -0.02959    0.01504   -1.97    0.049 *
## Year1999           -0.03666    0.01465   -2.50    0.012 *
## Year2000           -0.07834    0.01522   -5.15  2.7e-07 ***
## Year2001           -0.07395    0.01606   -4.61  4.1e-06 ***
## Year2002           -0.05858    0.01459   -4.01  6.0e-05 ***
## Year2003           -0.07344    0.01383   -5.31  1.1e-07 ***
## Year2004           -0.09386    0.01395   -6.73  1.8e-11 ***
## Year2005           -0.11720    0.01377   -8.51  < 2e-16 ***
## Year2006           -0.08955    0.01362   -6.57  5.0e-11 ***
```

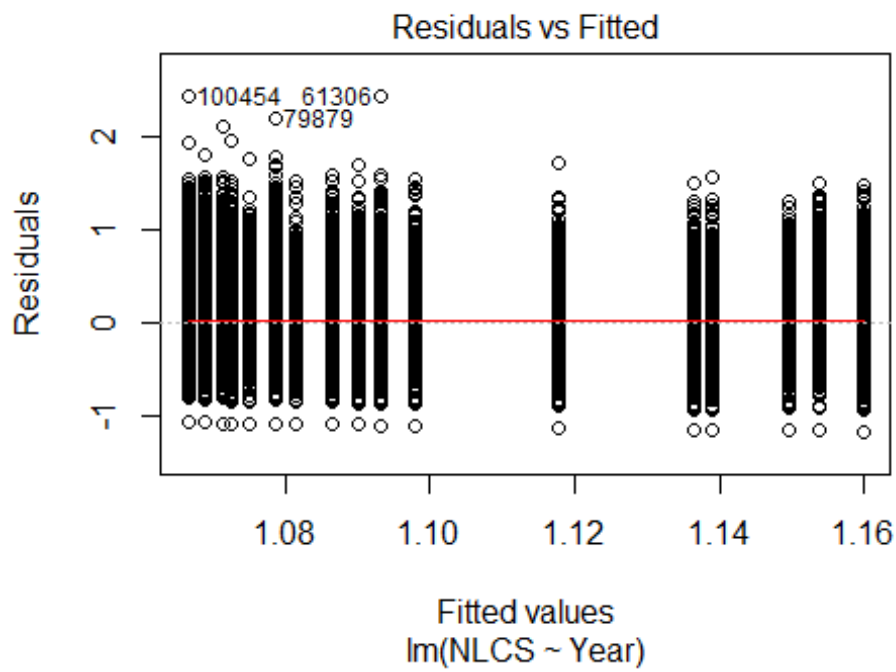
```
## Year2007           -0.09470    0.01355   -6.99  2.9e-12 ***
## Year2008           -0.10289    0.01374   -7.49  7.3e-14 ***
## Year2009           -0.09221    0.01390   -6.64  3.3e-11 ***
## Year2010           -0.10025    0.01455   -6.89  5.7e-12 ***
## Year2011           -0.09454    0.01427   -6.63  3.5e-11 ***
## Year2012           -0.12263    0.01447   -8.48  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.371
## Multiple R-squared:  0.00764,    Adjusted R-squared:  0.00706
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  observation 25673 is an outlier with |weight| = 0 ( < 3.4e-06);
##  2398 weights are ~= 1. The remaining 26727 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0004  0.8670  0.9500  0.8980  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          3.43e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##                   GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.003  1           1.001
## Year             1.003 16           1.000
```



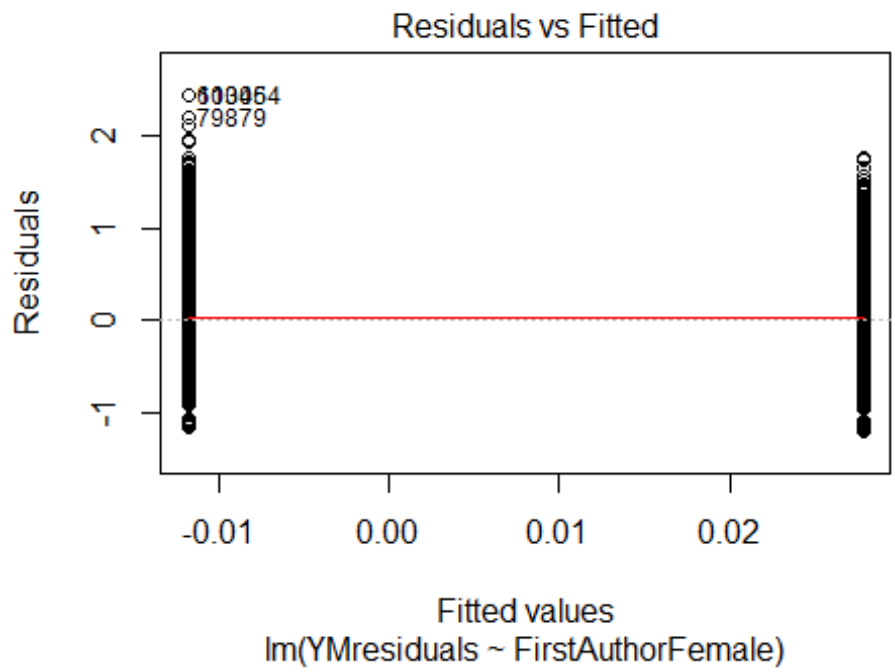

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##    Min     1Q Median     3Q    Max
## -1.181 -0.251  0.012  0.250  1.786
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.17008    0.01081  108.27  < 2e-16 ***
## LastAuthorFemale1  0.01078    0.00567    1.90    0.057 .
## Year1997           -0.03487    0.01506   -2.31    0.021 *
## Year1998           -0.02879    0.01504   -1.91    0.056 .
## Year1999           -0.03582    0.01465   -2.45    0.014 *
## Year2000           -0.07656    0.01522   -5.03  4.9e-07 ***
## Year2001           -0.07283    0.01605   -4.54  5.7e-06 ***
## Year2002           -0.05745    0.01460   -3.93  8.3e-05 ***
## Year2003           -0.07169    0.01383   -5.18  2.2e-07 ***
## Year2004           -0.09201    0.01396   -6.59  4.5e-11 ***
## Year2005           -0.11503    0.01377   -8.36  < 2e-16 ***
## Year2006           -0.08667    0.01361   -6.37  1.9e-10 ***
```

```
## Year2007          -0.09207    0.01355   -6.79  1.1e-11 ***
## Year2008          -0.10003    0.01373   -7.28  3.3e-13 ***
## Year2009          -0.08914    0.01389   -6.42  1.4e-10 ***
## Year2010          -0.09693    0.01453   -6.67  2.6e-11 ***
## Year2011          -0.09023    0.01424   -6.34  2.4e-10 ***
## Year2012          -0.11859    0.01444   -8.21  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.371
## Multiple R-squared:  0.00668,    Adjusted R-squared:  0.0061
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  observation 25673 is an outlier with |weight| = 0 ( < 3.4e-06);
##  2433 weights are ~= 1. The remaining 26692 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0007  0.8670  0.9500  0.8970  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          3.43e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Sample size for the above analysis:  29126"
## [1] ""
## [1] ""
## [1] "####################################"
## [1] "Analysis of AJSC 1105"
## [1] "####################################"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 4546 4332 4377 4420 4742 4955 4880 4458 4531 4961 5117 5672 6453 7213 7337
## 2011 2012
## 6926 7100
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 3271 3060 3201 3266 2768 2284 3854 3498 3527 3847 3881 4407 5026 5646 5838
```

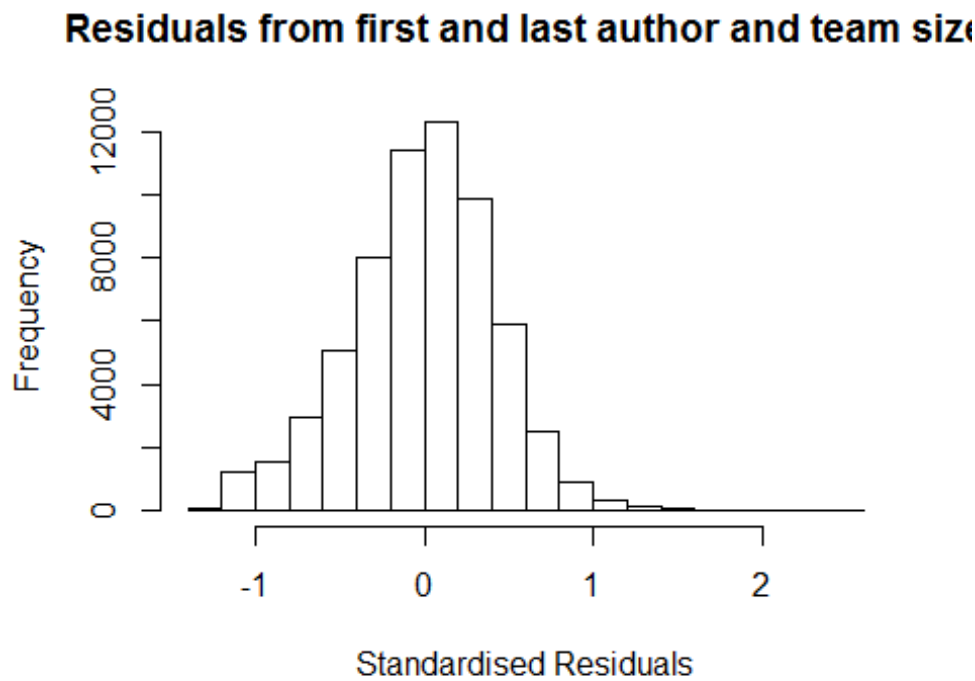
```
## 2011 2012
## 5561 5622
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 2998 2818 2919 2987 2543 2042 3478 3194 3226 3500 3506 3967 4554 5117 5276
## 2011 2012
## 5049 5113
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
##  Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 360, df = 16, p-value <2e-16
```




```
##
##  Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 190, df = 1, p-value <2e-16
```

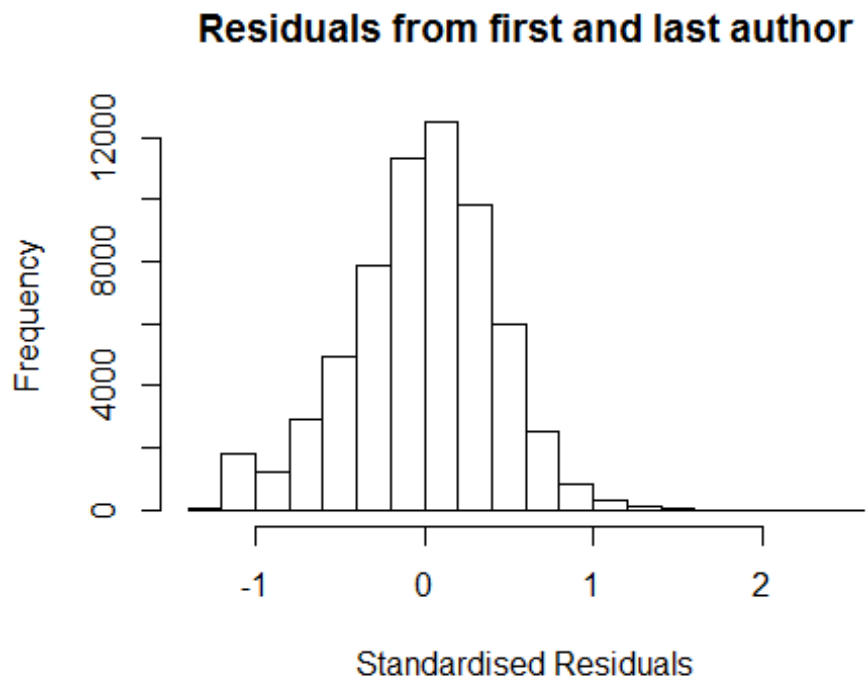



```
## [1] "Female first author team size 2018 geometric mean: 3.47658892094234"
## [1] "Male first author team size 2018 geometric mean: 3.17097772012575"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 3e+06, p-value = 2e-06
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 3.37237151723"
## [1] "Male last author team size 2018 geometric mean: 3.25764111110504"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 2400000, p-value = 0.1
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.039  1           1.019
## LastAuthorFemale  1.030  1           1.015
## UniqueAuthors     1.057  4           1.007
## Year              1.064 16           1.002
```



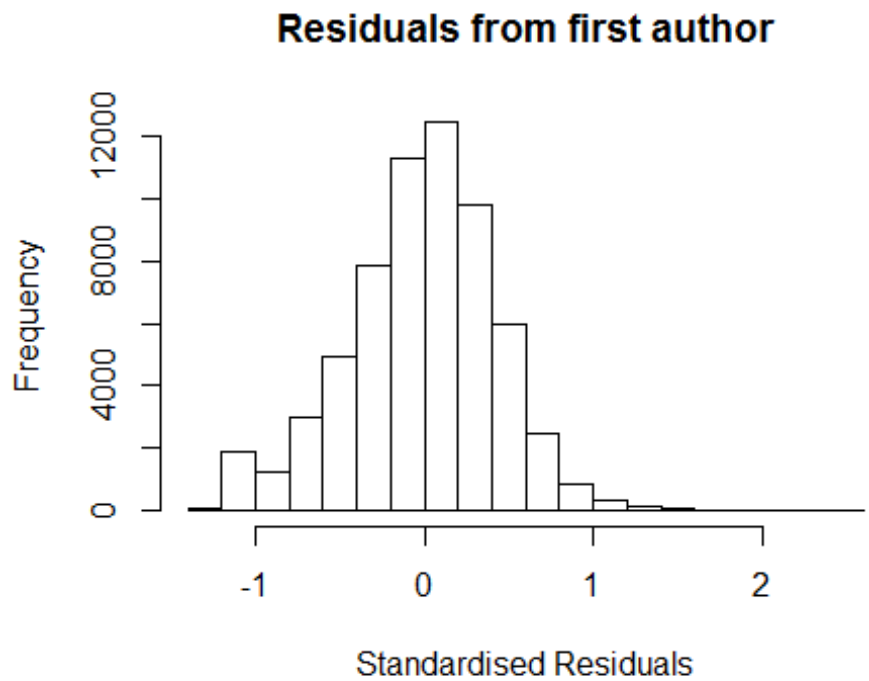

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
UniqueAuthors +
##     Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##     k.max = 1000))
##  \--> method = "MM"
## Residuals:
##    Min     1Q Median     3Q    Max
## -1.349 -0.270  0.014  0.274  2.430
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.07318    0.00951  112.79  < 2e-16 ***
## FirstAuthorFemale1  0.03335    0.00369    9.04  < 2e-16 ***
## LastAuthorFemale1   0.02143    0.00408    5.25  1.5e-07 ***
## UniqueAuthors2      0.10039    0.00565   17.76  < 2e-16 ***
## UniqueAuthors3      0.12801    0.00590   21.68  < 2e-16 ***
## UniqueAuthors4      0.15553    0.00661   23.54  < 2e-16 ***
## UniqueAuthors5      0.24033    0.00661   36.34  < 2e-16 ***
## Year1997           -0.00884    0.01188   -0.74   0.4569
## Year1998           -0.01920    0.01164   -1.65   0.0989 .
## Year1999           -0.05380    0.01139   -4.72  2.3e-06 ***
```

```
## Year2000           -0.03633    0.01164   -3.12   0.0018 **
## Year2001           -0.02945    0.01252   -2.35   0.0187 *
## Year2002           -0.09021    0.01118   -8.07  7.1e-16 ***
## Year2003           -0.09748    0.01117   -8.73  < 2e-16 ***
## Year2004           -0.10544    0.01127   -9.35  < 2e-16 ***
## Year2005           -0.10116    0.01095   -9.24  < 2e-16 ***
## Year2006           -0.09584    0.01090   -8.79  < 2e-16 ***
## Year2007           -0.10766    0.01061  -10.15  < 2e-16 ***
## Year2008           -0.12717    0.01079  -11.79  < 2e-16 ***
## Year2009           -0.13032    0.01065  -12.23  < 2e-16 ***
## Year2010           -0.13224    0.01059  -12.49  < 2e-16 ***
## Year2011           -0.13771    0.01072  -12.84  < 2e-16 ***
## Year2012           -0.14678    0.01089  -13.48  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.403
## Multiple R-squared:  0.0346, Adjusted R-squared:  0.0342
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
##  6 observations c(34698,40204,45508,56206,57543,60717)
##   are outliers with |weight| = 0 ( < 1.6e-06);
##  5191 weights are ~= 1. The remaining 57090 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0103  0.8650  0.9500  0.8980  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.61e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.039  1           1.019
## LastAuthorFemale  1.030  1           1.015
## Year              1.012 16           1.000
```



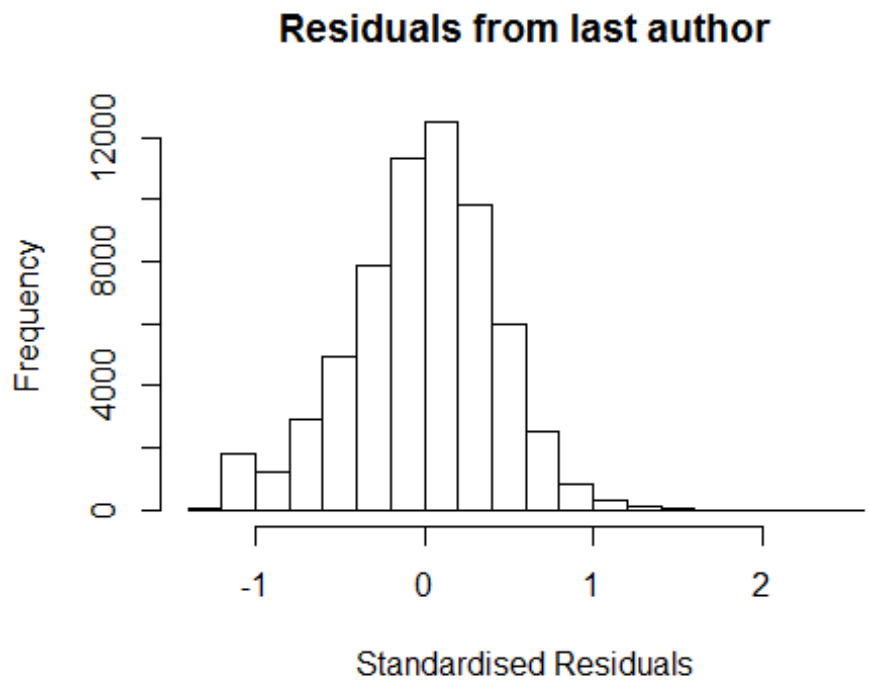

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.2214 -0.2732  0.0151  0.2749  2.4271
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)          1.16210    0.00849  136.86  < 2e-16 ***
## FirstAuthorFemale1   0.04114    0.00371   11.08  < 2e-16 ***
## LastAuthorFemale1    0.01818    0.00411    4.42  9.8e-06 ***
## Year1997            -0.00788    0.01174   -0.67    0.502
## Year1998            -0.01290    0.01153   -1.12    0.264
## Year1999            -0.04766    0.01122   -4.25  2.2e-05 ***
## Year2000            -0.02722    0.01152   -2.36    0.018 *
## Year2001            -0.01881    0.01248   -1.51    0.132
## Year2002            -0.07512    0.01112   -6.76  1.4e-11 ***
## Year2003            -0.08138    0.01111   -7.33  2.4e-13 ***
## Year2004            -0.08495    0.01120   -7.59  3.3e-14 ***
## Year2005            -0.07633    0.01089   -7.01  2.5e-12 ***
```

```
## Year2006           -0.06621    0.01081   -6.13  9.1e-10 ***
## Year2007           -0.07723    0.01051   -7.35  2.0e-13 ***
## Year2008           -0.09281    0.01070   -8.67  < 2e-16 ***
## Year2009           -0.09576    0.01058   -9.05  < 2e-16 ***
## Year2010           -0.09650    0.01049   -9.20  < 2e-16 ***
## Year2011           -0.09761    0.01066   -9.16  < 2e-16 ***
## Year2012           -0.10303    0.01083   -9.52  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.406
## Multiple R-squared:  0.00812,    Adjusted R-squared:  0.00783
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  6 observations c(34698,40204,45508,56206,57543,60717)
##   are outliers with |weight| = 0 ( < 1.6e-06);
##  5098 weights are ~= 1. The remaining 57183 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.025   0.864   0.950   0.896   0.986   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.61e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.011  1           1.005
## Year              1.011 16           1.000
```




```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.2095 -0.2735  0.0159  0.2753  2.4240
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)          1.16494    0.00846  137.67  < 2e-16 ***
## FirstAuthorFemale1   0.04461    0.00367   12.16  < 2e-16 ***
## Year1997            -0.00798    0.01174   -0.68    0.497
## Year1998            -0.01294    0.01154   -1.12    0.262
## Year1999            -0.04759    0.01123   -4.24  2.3e-05 ***
## Year2000            -0.02709    0.01152   -2.35    0.019 *
## Year2001            -0.01854    0.01248   -1.49    0.137
## Year2002            -0.07495    0.01112   -6.74  1.6e-11 ***
## Year2003            -0.08141    0.01111   -7.33  2.4e-13 ***
## Year2004            -0.08483    0.01120   -7.57  3.7e-14 ***
## Year2005            -0.07612    0.01090   -6.98  2.9e-12 ***
## Year2006            -0.06573    0.01081   -6.08  1.2e-09 ***
```

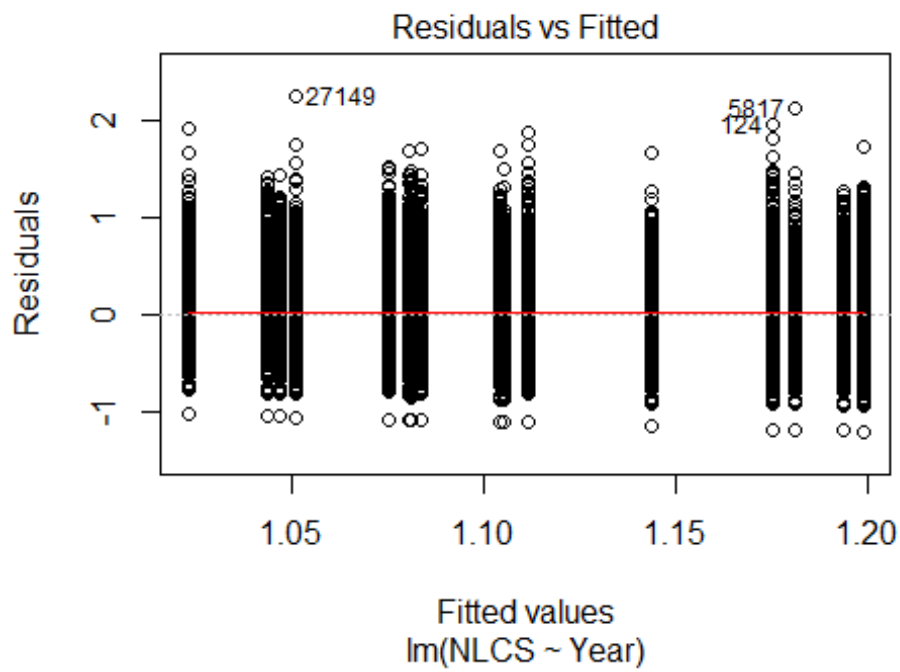
```
## Year2007           -0.07691    0.01051   -7.31  2.6e-13 ***
## Year2008           -0.09259    0.01070   -8.65  < 2e-16 ***
## Year2009           -0.09550    0.01059   -9.02  < 2e-16 ***
## Year2010           -0.09619    0.01049   -9.17  < 2e-16 ***
## Year2011           -0.09726    0.01066   -9.12  < 2e-16 ***
## Year2012           -0.10267    0.01083   -9.48  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.406
## Multiple R-squared:  0.00778,    Adjusted R-squared:  0.00751
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  6 observations c(34698,40204,45508,56206,57543,60717)
##   are outliers with |weight| = 0 ( < 1.6e-06);
##  5097 weights are ~= 1. The remaining 57184 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0239  0.8640  0.9500  0.8970  0.9860  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.61e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##                   GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.002  1           1.001
## Year             1.002 16           1.000
```



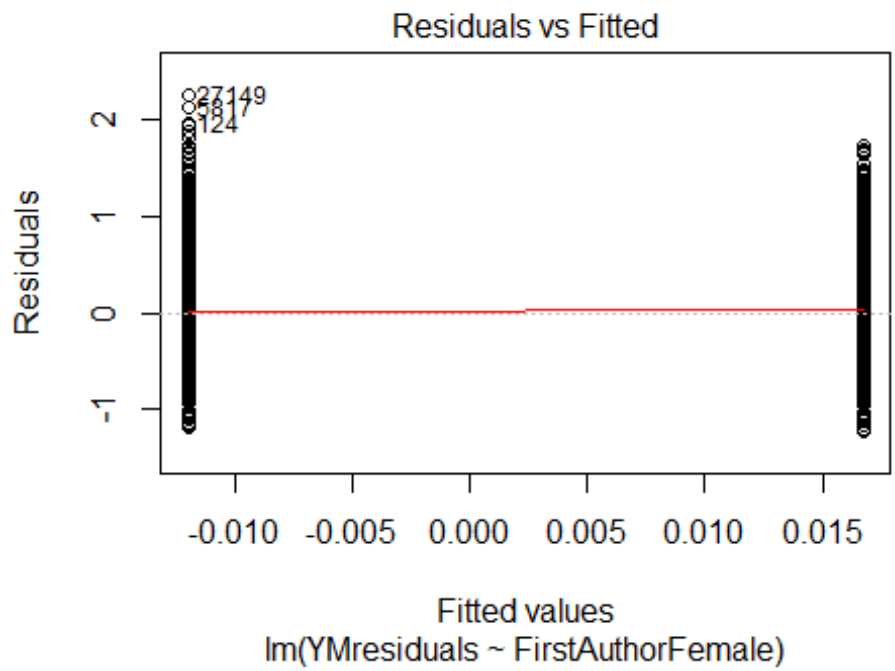

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##    Min     1Q Median     3Q    Max
## -1.197 -0.273  0.016  0.275  2.416
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.16951    0.00847  138.13  < 2e-16 ***
## LastAuthorFemale1   0.02766    0.00406    6.81  9.5e-12 ***
## Year1997           -0.00733    0.01174   -0.62    0.532
## Year1998           -0.01173    0.01153   -1.02    0.309
## Year1999           -0.04648    0.01124   -4.14  3.5e-05 ***
## Year2000           -0.02524    0.01153   -2.19    0.029 *
## Year2001           -0.01637    0.01249   -1.31    0.190
## Year2002           -0.07330    0.01112   -6.59  4.4e-11 ***
## Year2003           -0.07851    0.01112   -7.06  1.7e-12 ***
## Year2004           -0.08186    0.01120   -7.31  2.8e-13 ***
## Year2005           -0.07378    0.01090   -6.77  1.3e-11 ***
## Year2006           -0.06283    0.01082   -5.81  6.3e-09 ***
```

```
## Year2007          -0.07395    0.01052   -7.03  2.1e-12 ***
## Year2008          -0.08890    0.01071   -8.30  < 2e-16 ***
## Year2009          -0.09081    0.01058   -8.58  < 2e-16 ***
## Year2010          -0.09200    0.01049   -8.77  < 2e-16 ***
## Year2011          -0.09296    0.01066   -8.72  < 2e-16 ***
## Year2012          -0.09792    0.01082   -9.05  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.406
## Multiple R-squared:  0.00617,    Adjusted R-squared:  0.0059
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  6 observations c(34698,40204,45508,56206,57543,60717)
##   are outliers with |weight| = 0 ( < 1.6e-06);
##  5184 weights are ~= 1. The remaining 57097 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0173  0.8640  0.9500  0.8960  0.9860  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.61e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                  psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Sample size for the above analysis:  62287"
## [1] ""
## [1] ""
## [1] "##################################"
## [1] "Analysis of AJSC 1106"
## [1] "##################################"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 2028 2190 1944 1975 2152 2125 2256 2137 2167 2084 2173 2442 2327 2313 2301
## 2011 2012
## 2405 2577
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
```

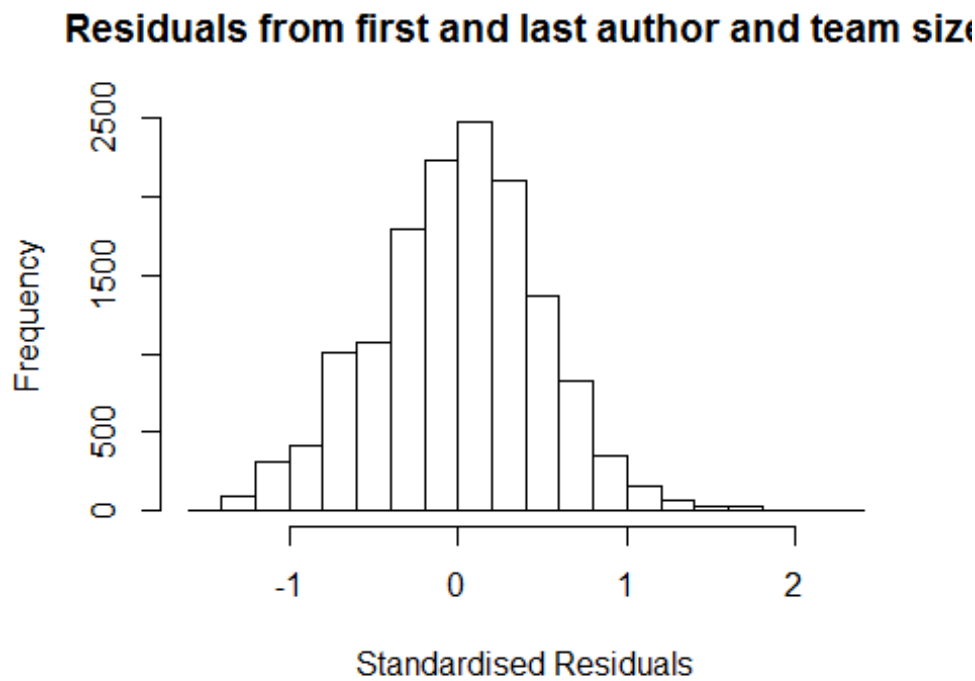
```
##  898  943  829  820  789  713  941  916  950  991  992 1160 1049 1031 1128
## 2011 2012
## 1198 1264
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  790  825  734  722  696  621  808  801  820  849  866  986  885  881  945
## 2011 2012
## 1017 1093
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
##  Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 180, df = 16, p-value <2e-16
```




```
##
##  Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 0.8, df = 1, p-value = 0.4
```

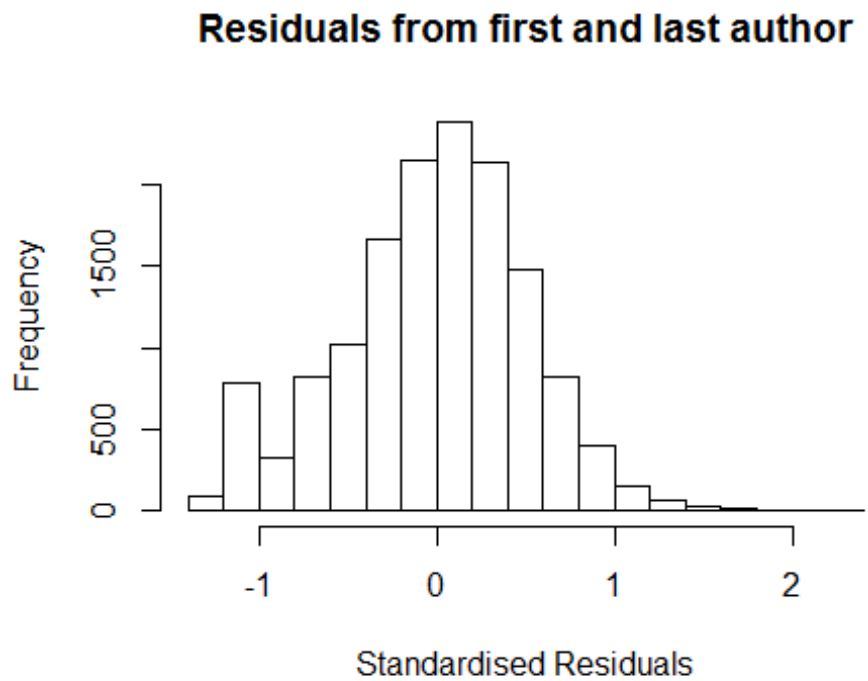



```
## [1] "Female first author team size 2018 geometric mean: 3.77729523591286"
## [1] "Male first author team size 2018 geometric mean: 3.55957890949966"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 150000, p-value = 0.2
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 3.61407851995296"
## [1] "Male last author team size 2018 geometric mean: 3.70088378067904"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 130000, p-value = 0.7
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.070  1           1.035
## LastAuthorFemale  1.058  1           1.029
## UniqueAuthors     1.063  4           1.008
## Year              1.062 16           1.002
```



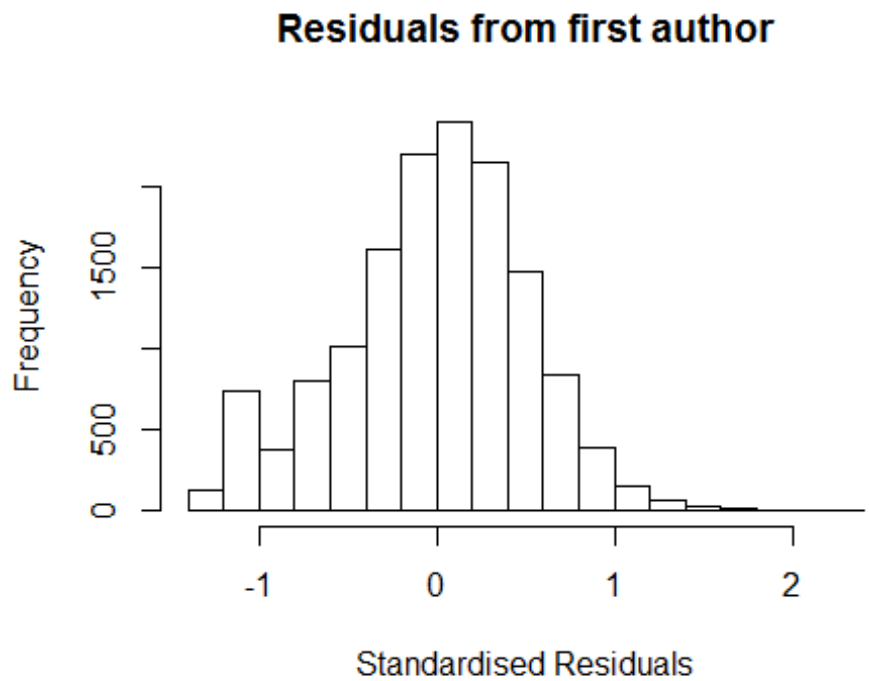

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
UniqueAuthors +
##     Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##     k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.4428 -0.3179  0.0181  0.3232  2.2561
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         0.85347    0.02458   34.73  < 2e-16 ***
## FirstAuthorFemale1  0.01773    0.00865    2.05   0.0404 *
## LastAuthorFemale1  -0.03010    0.00954   -3.16   0.0016 **
## UniqueAuthors2      0.36273    0.01624   22.33  < 2e-16 ***
## UniqueAuthors3      0.39597    0.01630   24.29  < 2e-16 ***
## UniqueAuthors4      0.46695    0.01690   27.63  < 2e-16 ***
## UniqueAuthors5      0.54387    0.01565   34.74  < 2e-16 ***
## Year1997            0.02777    0.02915    0.95   0.3408
## Year1998           -0.01732    0.02889   -0.60   0.5489
## Year1999           -0.00384    0.02816   -0.14   0.8917
```

```
## Year2000           -0.14094    0.02998   -4.70  2.6e-06 ***
## Year2001           -0.15223    0.03017   -5.05  4.6e-07 ***
## Year2002           -0.11649    0.02763   -4.22  2.5e-05 ***
## Year2003           -0.15236    0.02701   -5.64  1.7e-08 ***
## Year2004           -0.12049    0.02667   -4.52  6.3e-06 ***
## Year2005           -0.10063    0.02506   -4.02  5.9e-05 ***
## Year2006           -0.15657    0.02603   -6.02  1.8e-09 ***
## Year2007           -0.19166    0.02583   -7.42  1.2e-13 ***
## Year2008           -0.17394    0.02645   -6.58  5.0e-11 ***
## Year2009           -0.21680    0.02737   -7.92  2.5e-15 ***
## Year2010           -0.15214    0.02671   -5.70  1.2e-08 ***
## Year2011           -0.17819    0.02659   -6.70  2.1e-11 ***
## Year2012           -0.14156    0.02603   -5.44  5.5e-08 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.474
## Multiple R-squared:  0.126,  Adjusted R-squared:  0.125
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
##  observation 14272 is an outlier with |weight| = 0 ( < 7e-06);
##  1184 weights are ~= 1. The remaining 13154 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0034  0.8600  0.9500  0.9020  0.9860  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          6.97e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.043  1           1.021
## LastAuthorFemale  1.042  1           1.021
## Year              1.020 16           1.001
```



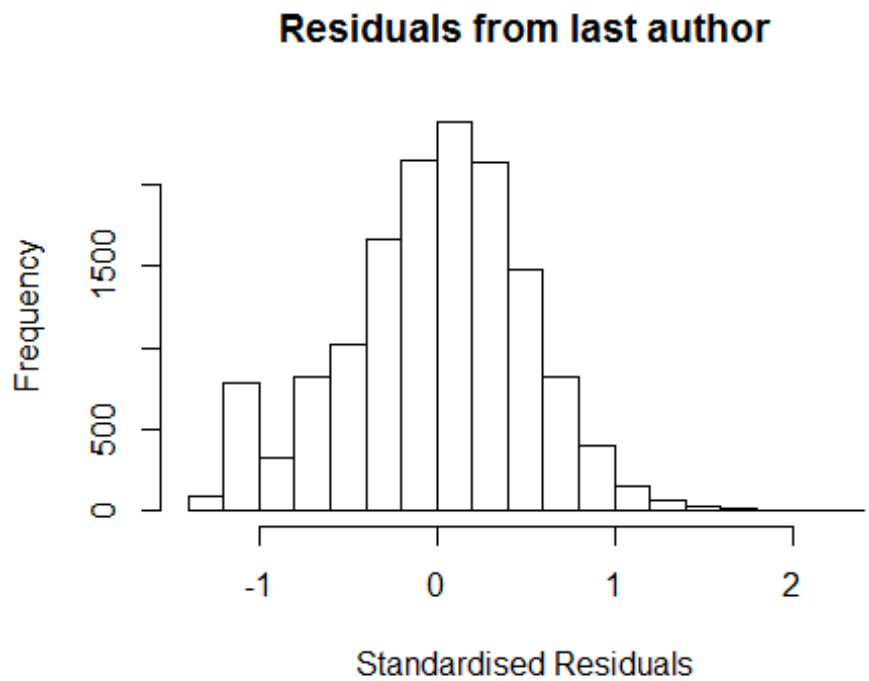

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##    Min     1Q Median     3Q    Max
## -1.282 -0.326  0.025  0.334  2.288
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)          1.17791    0.02270   51.89  < 2e-16 ***
## FirstAuthorFemale1   0.05289    0.00904    5.85  4.9e-09 ***
## LastAuthorFemale1   -0.04656    0.01014   -4.59  4.4e-06 ***
## Year1997             0.05087    0.03021    1.68  0.09227 .
## Year1998             0.00898    0.03033    0.30  0.76714
## Year1999             0.02942    0.02941    1.00  0.31725
## Year2000            -0.12861    0.03243   -3.97  7.3e-05 ***
## Year2001            -0.13088    0.03264   -4.01  6.1e-05 ***
## Year2002            -0.07821    0.02948   -2.65  0.00799 **
## Year2003            -0.10131    0.02884   -3.51  0.00045 ***
## Year2004            -0.06533    0.02853   -2.29  0.02206 *
## Year2005            -0.02444    0.02635   -0.93  0.35376
```

```
## Year2006           -0.09181    0.02733   -3.36  0.00078 ***
## Year2007           -0.13194    0.02707   -4.87  1.1e-06 ***
## Year2008           -0.10412    0.02832   -3.68  0.00024 ***
## Year2009           -0.17155    0.02927   -5.86  4.7e-09 ***
## Year2010           -0.09445    0.02821   -3.35  0.00082 ***
## Year2011           -0.11409    0.02808   -4.06  4.9e-05 ***
## Year2012           -0.07012    0.02773   -2.53  0.01145 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.488
## Multiple R-squared:  0.0165, Adjusted R-squared:  0.0152
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
##  observation 9111 is an outlier with |weight| = 0 ( < 7e-06);
##  1196 weights are ~= 1. The remaining 13142 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0783  0.8590  0.9500  0.8980  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          6.97e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.012  1           1.006
## Year              1.012 16           1.000
```




```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.2605 -0.3228  0.0258  0.3320  2.2497
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)          1.16989    0.02269   51.55  < 2e-16 ***
## FirstAuthorFemale1   0.04121    0.00897    4.60  4.4e-06 ***
## Year1997             0.04943    0.03030    1.63  0.10287
## Year1998             0.00990    0.03042    0.33  0.74488
## Year1999             0.03163    0.02946    1.07  0.28291
## Year2000            -0.13083    0.03251   -4.02  5.7e-05 ***
## Year2001            -0.13497    0.03277   -4.12  3.8e-05 ***
## Year2002            -0.07889    0.02954   -2.67  0.00758 **
## Year2003            -0.10240    0.02892   -3.54  0.00040 ***
## Year2004            -0.06585    0.02857   -2.31  0.02118 *
## Year2005            -0.02496    0.02639   -0.95  0.34425
## Year2006            -0.09300    0.02736   -3.40  0.00068 ***
```

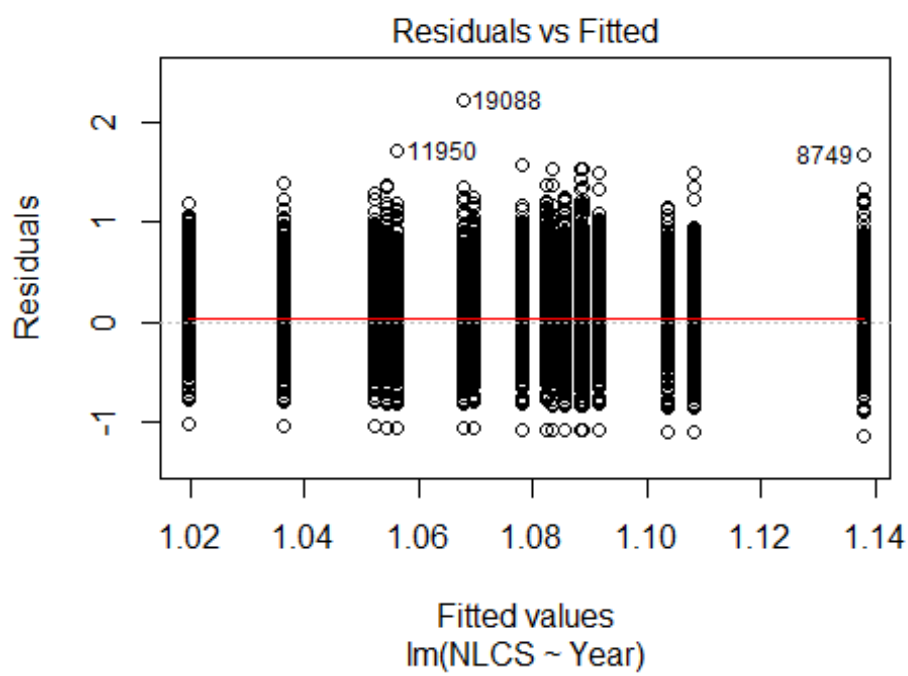
```
## Year2007           -0.13257    0.02714   -4.88  1.0e-06 ***
## Year2008           -0.10410    0.02836   -3.67  0.00024 ***
## Year2009           -0.17114    0.02933   -5.83  5.5e-09 ***
## Year2010           -0.09605    0.02831   -3.39  0.00069 ***
## Year2011           -0.11618    0.02816   -4.13  3.7e-05 ***
## Year2012           -0.07298    0.02779   -2.63  0.00865 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.488
## Multiple R-squared:  0.015,  Adjusted R-squared:  0.0138
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
##  1200 weights are ~= 1. The remaining 13139 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.001   0.859   0.950   0.897   0.985   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          6.97e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##                   GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.012  1           1.006
## Year             1.012 16           1.000
```



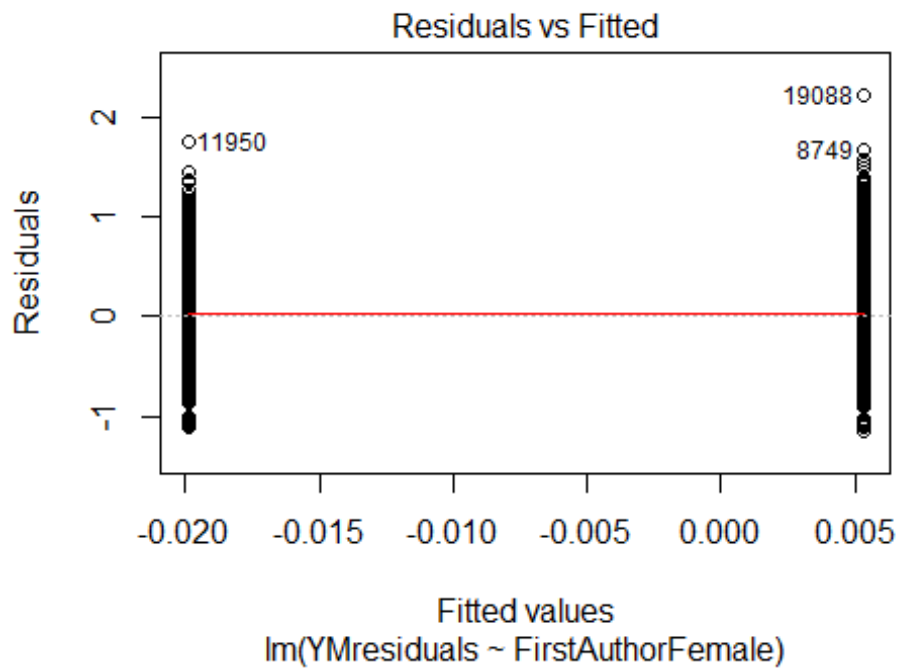

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.2434 -0.3268  0.0242  0.3335  2.2551
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.19276    0.02255   52.89  < 2e-16 ***
## LastAuthorFemale1 -0.03106    0.00999   -3.11  0.00188 **
## Year1997            0.05067    0.03020    1.68  0.09342 .
## Year1998            0.01047    0.03035    0.34  0.73026
## Year1999            0.03102    0.02947    1.05  0.29253
## Year2000           -0.12579    0.03251   -3.87  0.00011 ***
## Year2001           -0.12732    0.03265   -3.90  9.7e-05 ***
## Year2002           -0.07485    0.02950   -2.54  0.01119 *
## Year2003           -0.09886    0.02887   -3.42  0.00062 ***
## Year2004           -0.06109    0.02859   -2.14  0.03264 *
## Year2005           -0.02041    0.02639   -0.77  0.43933
## Year2006           -0.09078    0.02734   -3.32  0.00090 ***
```

```
## Year2007          -0.12980    0.02714   -4.78  1.7e-06 ***
## Year2008          -0.10205    0.02831   -3.60  0.00031 ***
## Year2009          -0.16823    0.02930   -5.74  9.6e-09 ***
## Year2010          -0.09040    0.02822   -3.20  0.00136 **
## Year2011          -0.10925    0.02813   -3.88  0.00010 ***
## Year2012          -0.06324    0.02770   -2.28  0.02244 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.488
## Multiple R-squared:  0.0142, Adjusted R-squared:  0.013
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
##  1214 weights are ~= 1. The remaining 13125 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0008  0.8580  0.9500  0.8970  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          6.97e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Sample size for the above analysis:  14339"
## [1] ""
## [1] ""
## [1] "####################################"
## [1] "Analysis of AJSC 1107"
## [1] "####################################"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 1211 1396 1330 1344 1239 1484 1274 1506 1337 1345 1567 1550 1725 1532 1470
## 2011 2012
## 1559 1674
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  717  731  764  771  671  656  791  879  822  863 1016 1031 1066  960  861
## 2011 2012
```

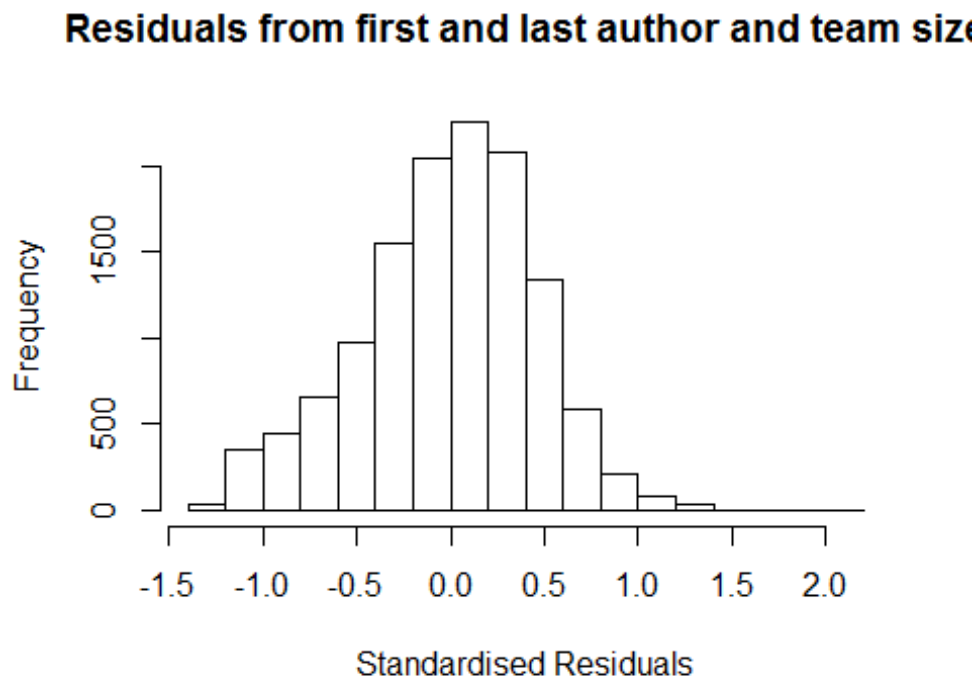
```
##  940  996
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  635  675  672  691  585  562  686  760  725  742  878  876  907  844  743
## 2011 2012
##  798  843
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
##  Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 30, df = 16, p-value = 0.02
```




```
##
##  Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 2.6, df = 1, p-value = 0.1
```

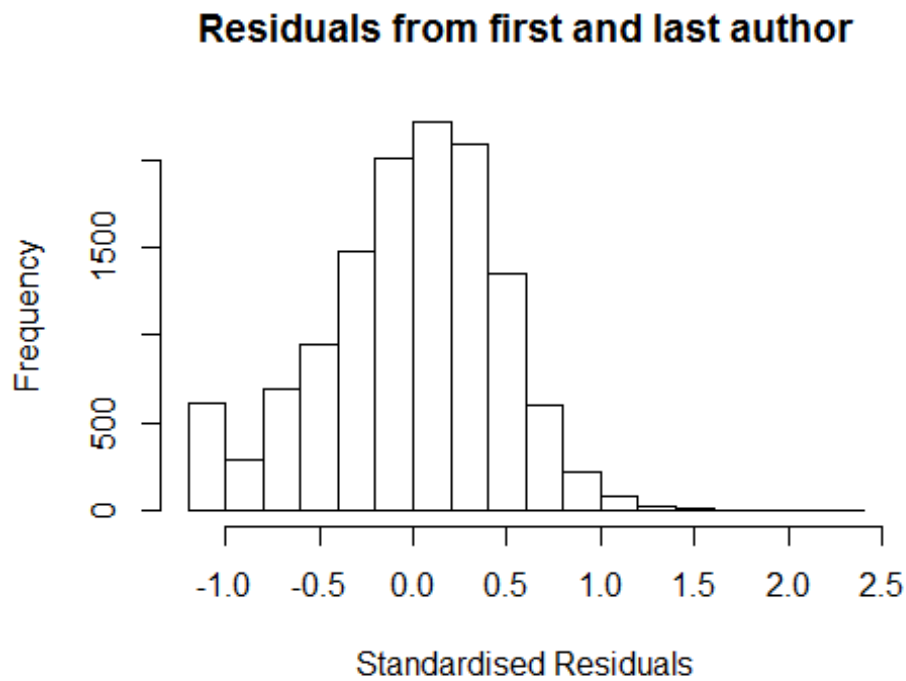



```
## [1] "Female first author team size 2018 geometric mean: 3.64087334060268"
## [1] "Male first author team size 2018 geometric mean: 3.41811031760486"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 70000, p-value = 0.6
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 3.62850296091416"
## [1] "Male last author team size 2018 geometric mean: 3.44795819413544"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 59000, p-value = 0.3
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.033  1           1.016
## LastAuthorFemale  1.020  1           1.010
## UniqueAuthors     1.080  4           1.010
## Year              1.098 16           1.003
```



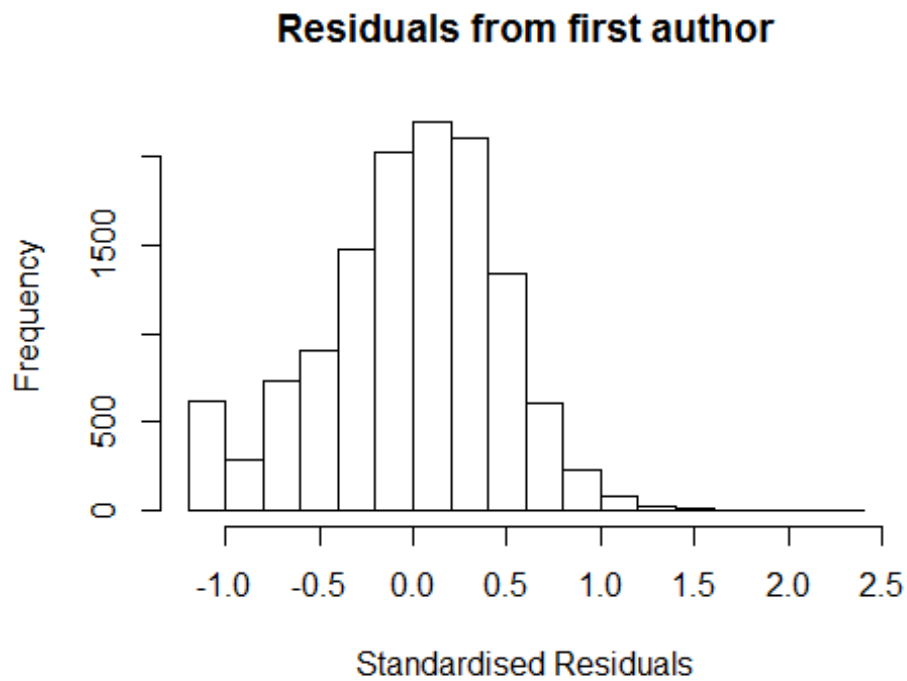

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
UniqueAuthors +
##     Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##     k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.3142 -0.3003  0.0253  0.3023  2.1118
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         0.98972    0.02129   46.50  < 2e-16 ***
## FirstAuthorFemale1 -0.01925    0.01054   -1.83  0.06787 .
## LastAuthorFemale1  -0.04074    0.01192   -3.42  0.00063 ***
## UniqueAuthors2      0.14474    0.01461    9.90  < 2e-16 ***
## UniqueAuthors3      0.17748    0.01519   11.69  < 2e-16 ***
## UniqueAuthors4      0.24130    0.01603   15.05  < 2e-16 ***
## UniqueAuthors5      0.31906    0.01564   20.40  < 2e-16 ***
## Year1997           -0.00613    0.02662   -0.23  0.81787
## Year1998           -0.01957    0.02676   -0.73  0.46445
## Year1999           -0.03702    0.02626   -1.41  0.15867
```

```
## Year2000           -0.04125    0.02689   -1.53  0.12508
## Year2001            0.00541    0.02675    0.20  0.83978
## Year2002           -0.10984    0.02773   -3.96  7.5e-05 ***
## Year2003           -0.08636    0.02496   -3.46  0.00054 ***
## Year2004           -0.09464    0.02560   -3.70  0.00022 ***
## Year2005           -0.09920    0.02599   -3.82  0.00014 ***
## Year2006           -0.10313    0.02444   -4.22  2.5e-05 ***
## Year2007           -0.09282    0.02456   -3.78  0.00016 ***
## Year2008           -0.06589    0.02379   -2.77  0.00561 **
## Year2009           -0.07180    0.02521   -2.85  0.00440 **
## Year2010           -0.08412    0.02693   -3.12  0.00179 **
## Year2011           -0.07071    0.02517   -2.81  0.00497 **
## Year2012           -0.08485    0.02555   -3.32  0.00090 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.442
## Multiple R-squared:  0.0466, Adjusted R-squared:  0.0449
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
##  observation 8193 is an outlier with |weight| = 0 ( < 7.9e-06);
##  1009 weights are ~= 1. The remaining 11612 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0518  0.8650  0.9500  0.9000  0.9840  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          7.92e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                    psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.028  1           1.014
## LastAuthorFemale  1.017  1           1.009
## Year              1.025 16           1.001
```



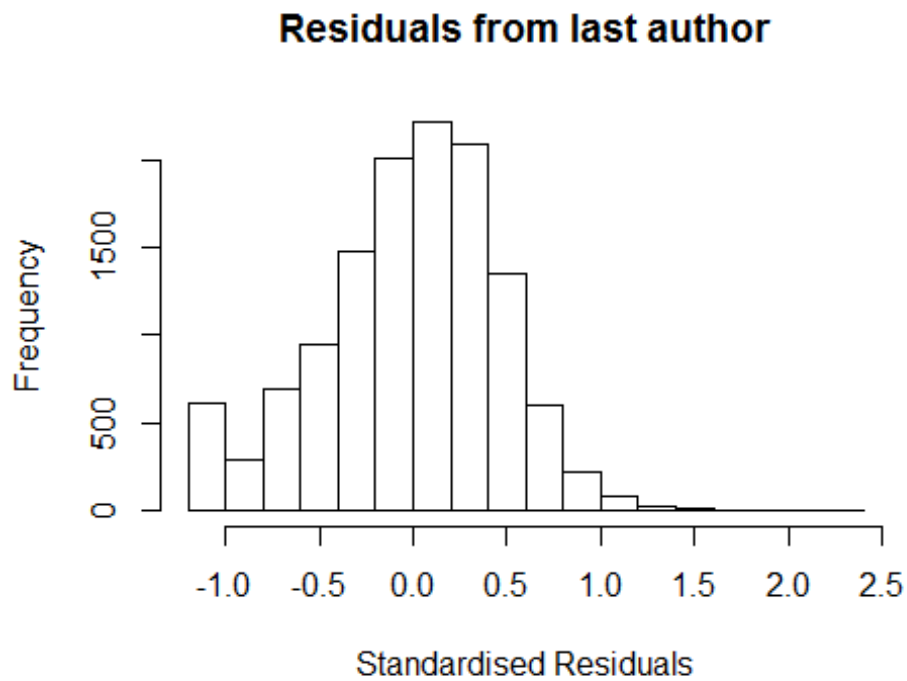

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.1609 -0.3033  0.0259  0.3090  2.2477
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)          1.12529    0.01905   59.08  < 2e-16 ***
## FirstAuthorFemale1 -0.01309    0.01082   -1.21  0.22667
## LastAuthorFemale1  -0.04383    0.01231   -3.56  0.00037 ***
## Year1997           -0.01168    0.02694   -0.43  0.66455
## Year1998           -0.00419    0.02706   -0.15  0.87695
## Year1999           -0.02068    0.02636   -0.78  0.43280
## Year2000           -0.02800    0.02694   -1.04  0.29869
## Year2001            0.03563    0.02689    1.33  0.18511
## Year2002           -0.07885    0.02784   -2.83  0.00462 **
## Year2003           -0.05004    0.02520   -1.99  0.04710 *
## Year2004           -0.04897    0.02581   -1.90  0.05781 .
## Year2005           -0.06100    0.02648   -2.30  0.02128 *
```

```
## Year2006           -0.05621    0.02470   -2.28  0.02291 *
## Year2007           -0.04219    0.02477   -1.70  0.08853 .
## Year2008           -0.01619    0.02399   -0.67  0.49988
## Year2009           -0.01717    0.02532   -0.68  0.49763
## Year2010           -0.03413    0.02724   -1.25  0.21035
## Year2011           -0.01158    0.02543   -0.46  0.64876
## Year2012           -0.02372    0.02581   -0.92  0.35813
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.448
## Multiple R-squared:  0.00445,    Adjusted R-squared:  0.00303
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  observation 8193 is an outlier with |weight| = 0 ( < 7.9e-06);
##  1046 weights are ~= 1. The remaining 11575 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.108   0.863   0.949   0.898   0.984   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          7.92e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.019  1           1.009
## Year              1.019 16           1.001
```




```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.1553 -0.3040  0.0257  0.3098  2.2104
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.12066    0.01901   58.94   <2e-16 ***
## FirstAuthorFemale1 -0.01994    0.01086   -1.84   0.0665 .
## Year1997           -0.01165    0.02694   -0.43   0.6654
## Year1998           -0.00402    0.02705   -0.15   0.8819
## Year1999           -0.02124    0.02631   -0.81   0.4195
## Year2000           -0.02821    0.02698   -1.05   0.2958
## Year2001            0.03468    0.02687    1.29   0.1969
## Year2002           -0.07954    0.02787   -2.85   0.0043 **
## Year2003           -0.05152    0.02523   -2.04   0.0412 *
## Year2004           -0.04951    0.02583   -1.92   0.0553 .
## Year2005           -0.06149    0.02650   -2.32   0.0203 *
## Year2006           -0.05664    0.02471   -2.29   0.0219 *
```

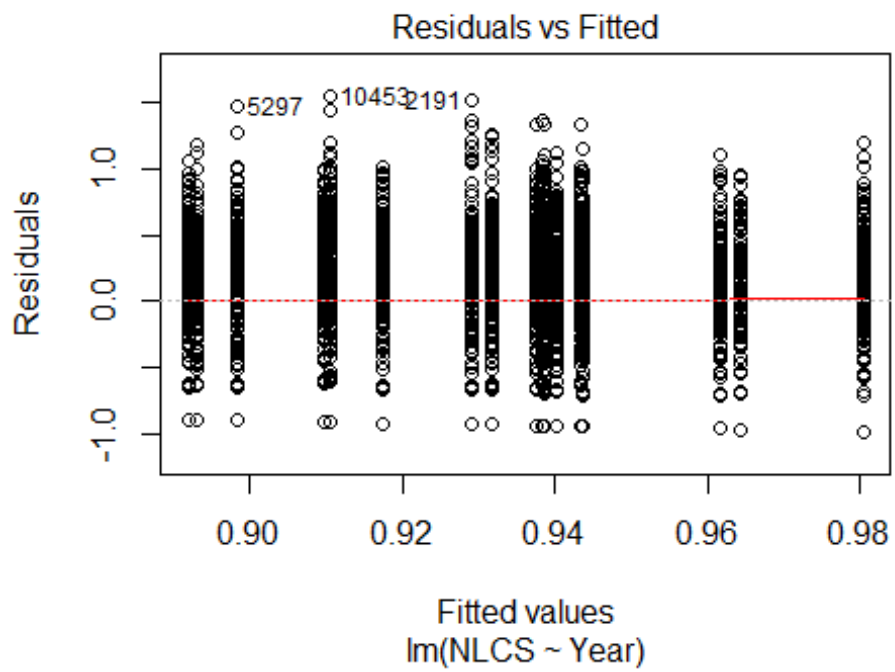
```
## Year2007           -0.04407    0.02477   -1.78   0.0752 .
## Year2008           -0.01717    0.02399   -0.72   0.4741
## Year2009           -0.01940    0.02534   -0.77   0.4439
## Year2010           -0.03550    0.02724   -1.30   0.1924
## Year2011           -0.01267    0.02542   -0.50   0.6182
## Year2012           -0.02645    0.02578   -1.03   0.3051
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.448
## Multiple R-squared:  0.00326,    Adjusted R-squared:  0.00191
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  observation 8193 is an outlier with |weight| = 0 ( < 7.9e-06);
##  1063 weights are ~= 1. The remaining 11558 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.101   0.862   0.949   0.898   0.984   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          7.92e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##                   GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.008  1           1.004
## Year             1.008 16           1.000
```



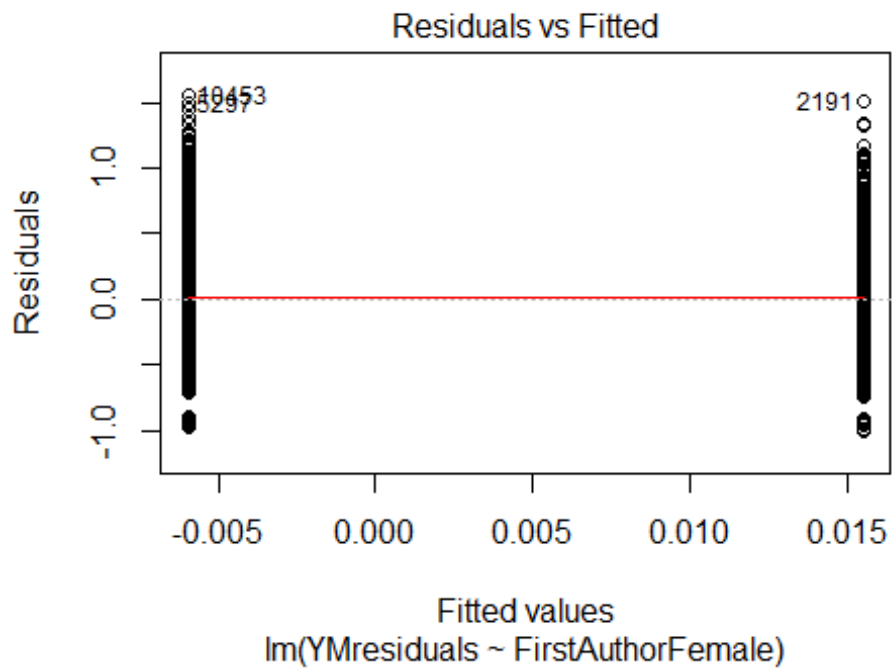

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##    Min     1Q Median     3Q    Max
## -1.159 -0.302  0.026  0.308  2.253
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.12385    0.01900   59.14  < 2e-16 ***
## LastAuthorFemale1 -0.04637    0.01231   -3.77  0.00017 ***
## Year1997          -0.01209    0.02693   -0.45  0.65342
## Year1998          -0.00438    0.02706   -0.16  0.87140
## Year1999          -0.02119    0.02637   -0.80  0.42151
## Year2000          -0.02878    0.02694   -1.07  0.28539
## Year2001           0.03524    0.02691    1.31  0.19028
## Year2002          -0.07968    0.02784   -2.86  0.00421 **
## Year2003          -0.05057    0.02521   -2.01  0.04483 *
## Year2004          -0.04992    0.02579   -1.94  0.05298 .
## Year2005          -0.06198    0.02646   -2.34  0.01916 *
## Year2006          -0.05745    0.02470   -2.33  0.02004 *
```

```
## Year2007          -0.04326    0.02475   -1.75  0.08058 .
## Year2008          -0.01793    0.02394   -0.75  0.45410
## Year2009          -0.01837    0.02530   -0.73  0.46789
## Year2010          -0.03525    0.02724   -1.29  0.19572
## Year2011          -0.01340    0.02537   -0.53  0.59746
## Year2012          -0.02529    0.02577   -0.98  0.32651
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.448
## Multiple R-squared:  0.00431,    Adjusted R-squared:  0.00297
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  observation 8193 is an outlier with |weight| = 0 ( < 7.9e-06);
##  1078 weights are ~= 1. The remaining 11543 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.114   0.862   0.949   0.898   0.984   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          7.92e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                    psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Sample size for the above analysis:  12622"
## [1] ""
## [1] ""
## [1] "####################################"
## [1] "Analysis of AJSC 1108"
## [1] "####################################"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  634  582  593  557  539  590  577  465  604  634  551  529  640  568  566
## 2011 2012
##  508  560
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  342  307  327  304  235  134  317  265  340  410  350  340  409  384  396
```

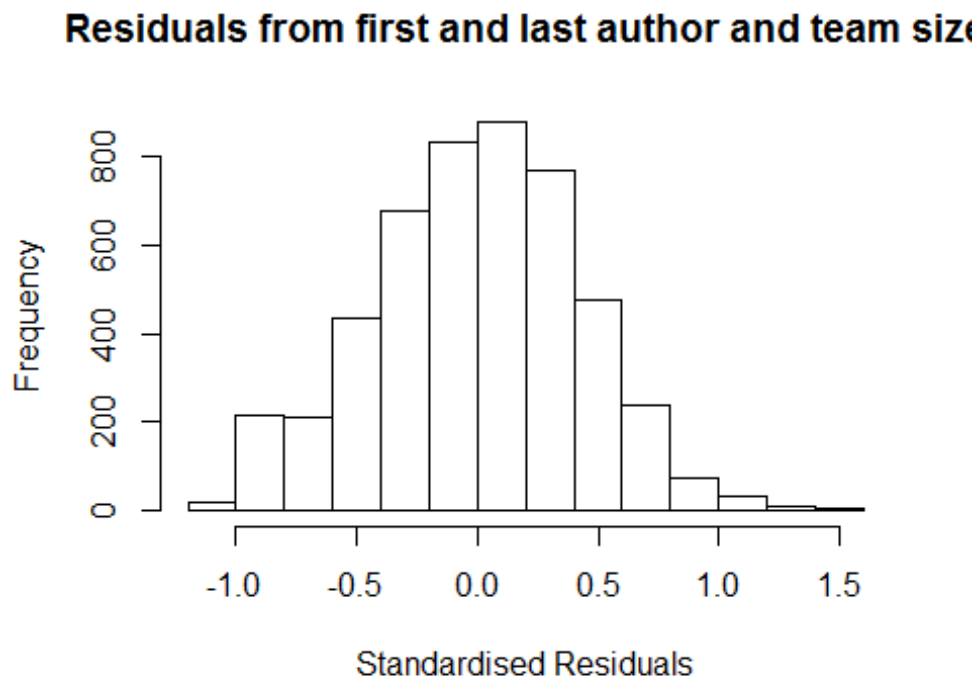
```
## 2011 2012
##  345  393
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  295  268  287  274  202  115  277  238  303  360  313  296  333  330  334
## 2011 2012
##  302  336
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
##  Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 39, df = 16, p-value = 0.001
```




```
##
##  Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 1.9, df = 1, p-value = 0.2
```

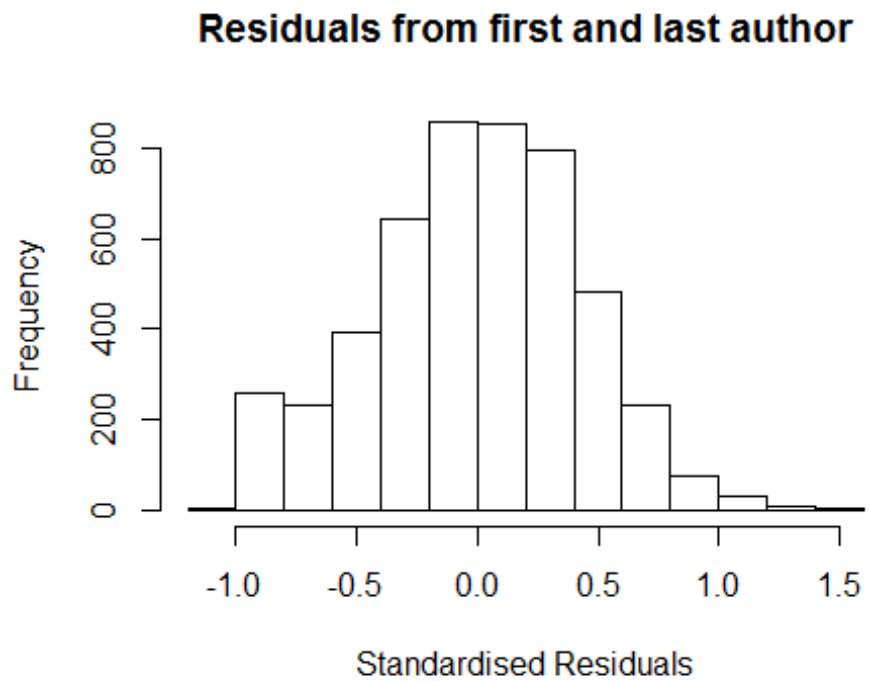



```
## [1] "Female first author team size 2018 geometric mean: 3.38499197155992"
## [1] "Male first author team size 2018 geometric mean: 3.17174159883503"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 8400, p-value = 0.4
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 3.10519095819335"
## [1] "Male last author team size 2018 geometric mean: 3.31245451158794"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 6600, p-value = 0.4
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##                     GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.033  1           1.016
## LastAuthorFemale  1.031  1           1.015
## UniqueAuthors     1.088  4           1.011
## Year              1.115 16           1.003
```



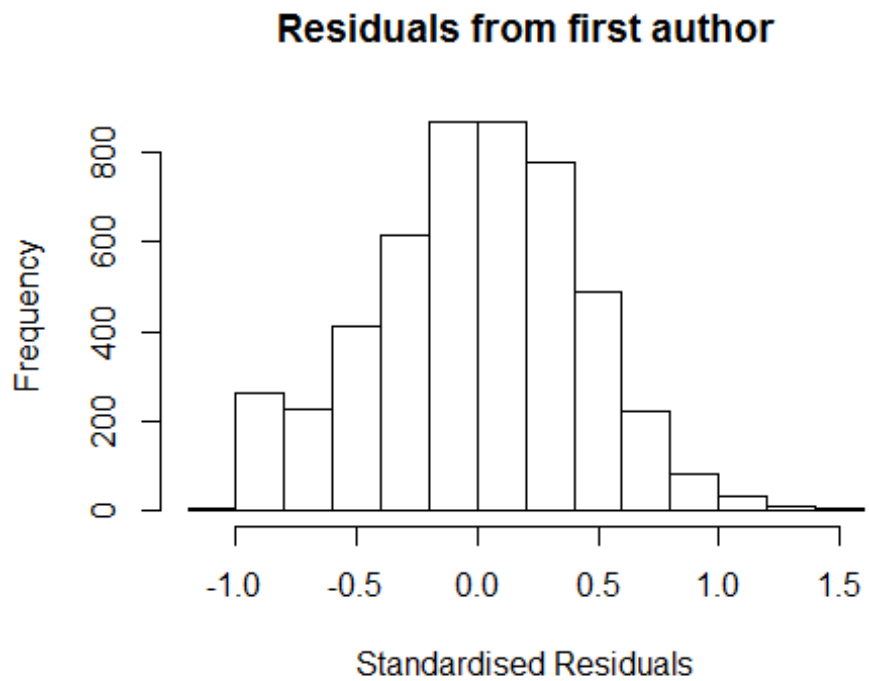

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
UniqueAuthors +
##     Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##     k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.0641 -0.2889  0.0111  0.2946  1.5637
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)          0.8831     0.0308   28.66  < 2e-16 ***
## FirstAuthorFemale1   0.0260     0.0141    1.85   0.0645 .
## LastAuthorFemale1    0.0281     0.0158    1.78   0.0759 .
## UniqueAuthors2       0.0740     0.0229    3.23   0.0012 **
## UniqueAuthors3       0.1076     0.0235    4.57  4.9e-06 ***
## UniqueAuthors4       0.1545     0.0248    6.22  5.3e-10 ***
## UniqueAuthors5       0.2065     0.0257    8.03  1.3e-15 ***
## Year1997            -0.0140     0.0362   -0.39   0.6987
## Year1998            -0.0392     0.0353   -1.11   0.2671
## Year1999            -0.0479     0.0371   -1.29   0.1973
```

```
## Year2000            -0.0486     0.0402   -1.21   0.2268
## Year2001            -0.0515     0.0554   -0.93   0.3526
## Year2002            -0.1027     0.0361   -2.84   0.0045 **
## Year2003            -0.0712     0.0377   -1.89   0.0587 .
## Year2004            -0.1136     0.0356   -3.19   0.0014 **
## Year2005            -0.1039     0.0341   -3.05   0.0023 **
## Year2006            -0.0785     0.0349   -2.25   0.0244 *
## Year2007            -0.0990     0.0358   -2.77   0.0057 **
## Year2008            -0.0814     0.0354   -2.30   0.0214 *
## Year2009            -0.0896     0.0350   -2.56   0.0105 *
## Year2010            -0.0787     0.0352   -2.23   0.0255 *
## Year2011            -0.0667     0.0375   -1.78   0.0753 .
## Year2012            -0.1169     0.0399   -2.93   0.0034 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.431
## Multiple R-squared:  0.0236, Adjusted R-squared:  0.0192
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  390 weights are ~= 1. The remaining 4473 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.161   0.873   0.951   0.908   0.985   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          2.06e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.031  1           1.016
## LastAuthorFemale  1.030  1           1.015
## Year              1.034 16           1.001
```



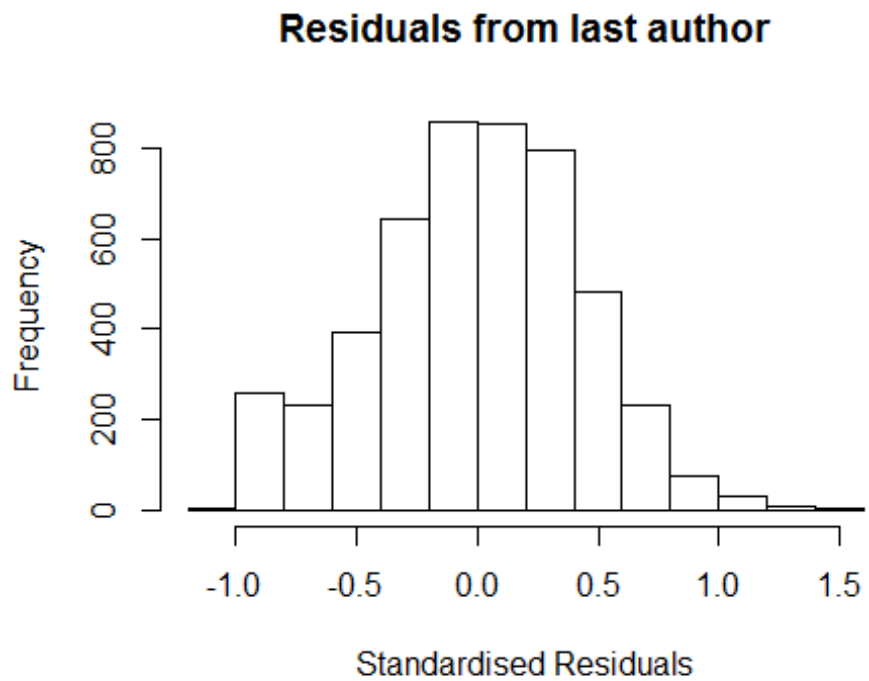

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.0046 -0.2862  0.0102  0.2962  1.5710
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)           0.9737     0.0253   38.55   <2e-16 ***
## FirstAuthorFemale1    0.0309     0.0142    2.17   0.0298 *
## LastAuthorFemale1     0.0243     0.0160    1.52   0.1287
## Year1997             -0.0180     0.0358   -0.50   0.6155
## Year1998             -0.0412     0.0355   -1.16   0.2458
## Year1999             -0.0454     0.0369   -1.23   0.2187
## Year2000             -0.0504     0.0401   -1.26   0.2088
## Year2001             -0.0410     0.0567   -0.72   0.4703
## Year2002             -0.0960     0.0363   -2.64   0.0083 **
## Year2003             -0.0596     0.0379   -1.57   0.1161
## Year2004             -0.1009     0.0354   -2.85   0.0044 **
## Year2005             -0.0904     0.0339   -2.67   0.0076 **
```

```
## Year2006            -0.0549     0.0345   -1.59   0.1114
## Year2007            -0.0744     0.0354   -2.10   0.0356 *
## Year2008            -0.0580     0.0352   -1.65   0.0989 .
## Year2009            -0.0680     0.0348   -1.95   0.0507 .
## Year2010            -0.0553     0.0350   -1.58   0.1143
## Year2011            -0.0465     0.0371   -1.25   0.2098
## Year2012            -0.0877     0.0401   -2.19   0.0289 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.433
## Multiple R-squared:  0.00515,    Adjusted R-squared:  0.00145
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  389 weights are ~= 1. The remaining 4474 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.161   0.870   0.951   0.907   0.985   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          2.06e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.018  1           1.009
## Year              1.018 16           1.001
```




```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.0118 -0.2898  0.0122  0.2962  1.5652
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)           0.9771     0.0251   38.87   <2e-16 ***
## FirstAuthorFemale1    0.0347     0.0142    2.45   0.0144 *
## Year1997             -0.0182     0.0358   -0.51   0.6112
## Year1998             -0.0417     0.0355   -1.17   0.2403
## Year1999             -0.0451     0.0369   -1.22   0.2216
## Year2000             -0.0493     0.0401   -1.23   0.2181
## Year2001             -0.0402     0.0567   -0.71   0.4783
## Year2002             -0.0961     0.0364   -2.64   0.0083 **
## Year2003             -0.0593     0.0379   -1.56   0.1183
## Year2004             -0.1005     0.0354   -2.84   0.0046 **
## Year2005             -0.0893     0.0338   -2.64   0.0083 **
## Year2006             -0.0533     0.0344   -1.55   0.1213
```

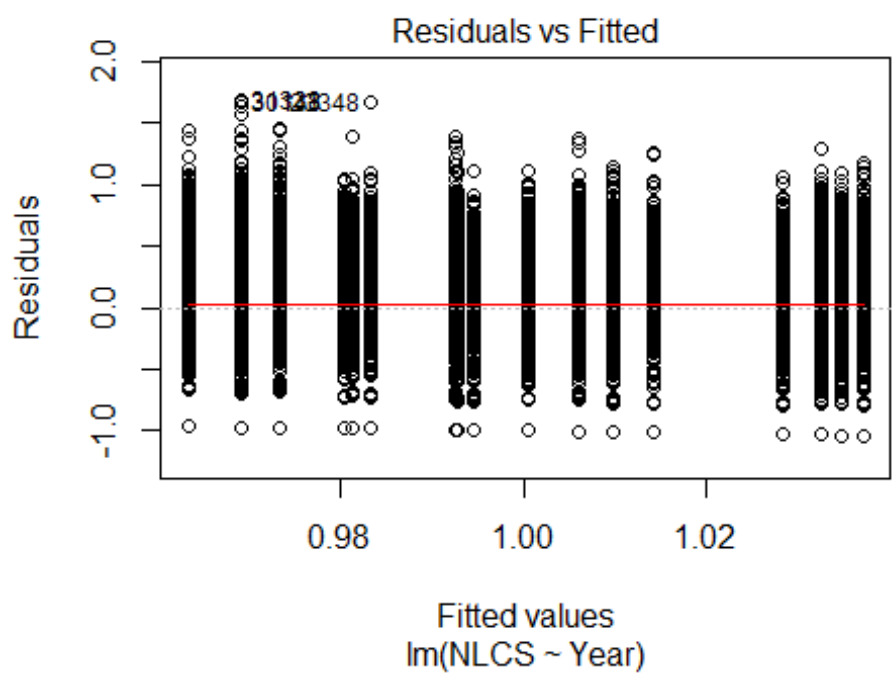
```
## Year2007            -0.0737     0.0354   -2.08   0.0377 *
## Year2008            -0.0575     0.0352   -1.63   0.1022
## Year2009            -0.0672     0.0348   -1.93   0.0538 .
## Year2010            -0.0548     0.0350   -1.57   0.1174
## Year2011            -0.0454     0.0371   -1.22   0.2216
## Year2012            -0.0853     0.0402   -2.12   0.0338 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.433
## Multiple R-squared:  0.00463,    Adjusted R-squared:  0.00114
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  396 weights are ~= 1. The remaining 4467 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.164   0.871   0.951   0.907   0.985   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          2.06e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##                   GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.017  1           1.008
## Year             1.017 16           1.001
```



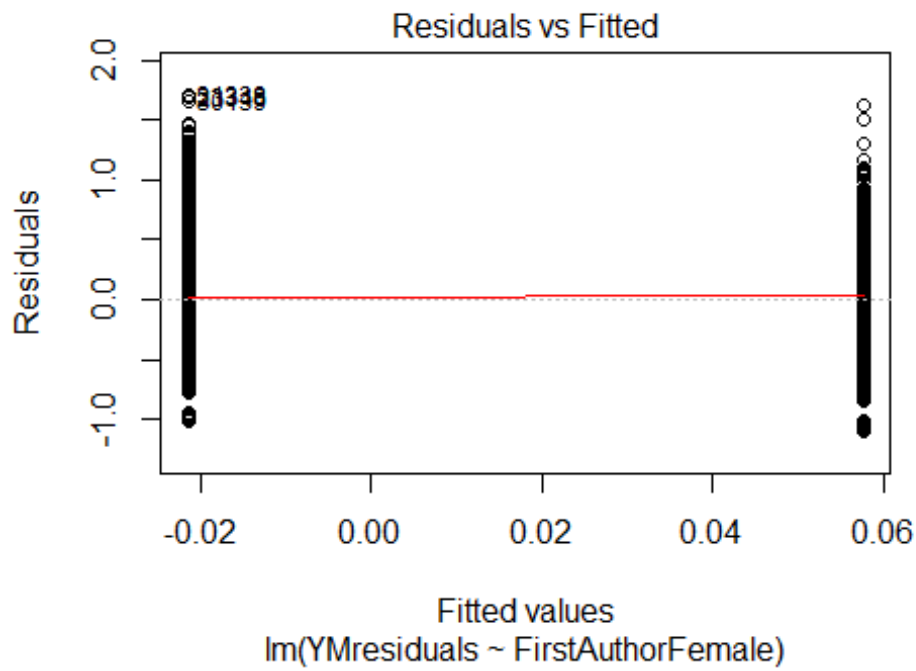

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -0.9929 -0.2912  0.0117  0.2967  1.5612
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         0.9783     0.0252   38.85   <2e-16 ***
## LastAuthorFemale1   0.0301     0.0159    1.89   0.0584 .
## Year1997           -0.0155     0.0358   -0.43   0.6654
## Year1998           -0.0388     0.0355   -1.09   0.2743
## Year1999           -0.0441     0.0370   -1.19   0.2331
## Year2000           -0.0488     0.0402   -1.21   0.2247
## Year2001           -0.0395     0.0566   -0.70   0.4848
## Year2002           -0.0930     0.0362   -2.57   0.0103 *
## Year2003           -0.0557     0.0379   -1.47   0.1417
## Year2004           -0.0981     0.0354   -2.77   0.0056 **
## Year2005           -0.0879     0.0339   -2.59   0.0095 **
## Year2006           -0.0511     0.0344   -1.48   0.1378
```

```
## Year2007           -0.0718     0.0354   -2.03   0.0425 *
## Year2008           -0.0553     0.0351   -1.57   0.1156
## Year2009           -0.0638     0.0347   -1.84   0.0663 .
## Year2010           -0.0505     0.0349   -1.45   0.1482
## Year2011           -0.0436     0.0370   -1.18   0.2382
## Year2012           -0.0825     0.0400   -2.06   0.0392 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.434
## Multiple R-squared:  0.00414,    Adjusted R-squared:  0.000642
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  396 weights are ~= 1. The remaining 4467 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.167   0.872   0.951   0.907   0.985   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          2.06e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Sample size for the above analysis:  4863"
## [1] ""
## [1] ""
## [1] "####################################"
## [1] "Analysis of AJSC 1109"
## [1] "####################################"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 1798 1664 1664 1578 1795 1770 1756 1794 1820 1842 1750 1769 1780 1799 1775
## 2011 2012
## 1790 1723
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 1116 1015 1045  992  867  682 1135 1173 1209 1200 1149 1173 1251 1202 1209
## 2011 2012
```

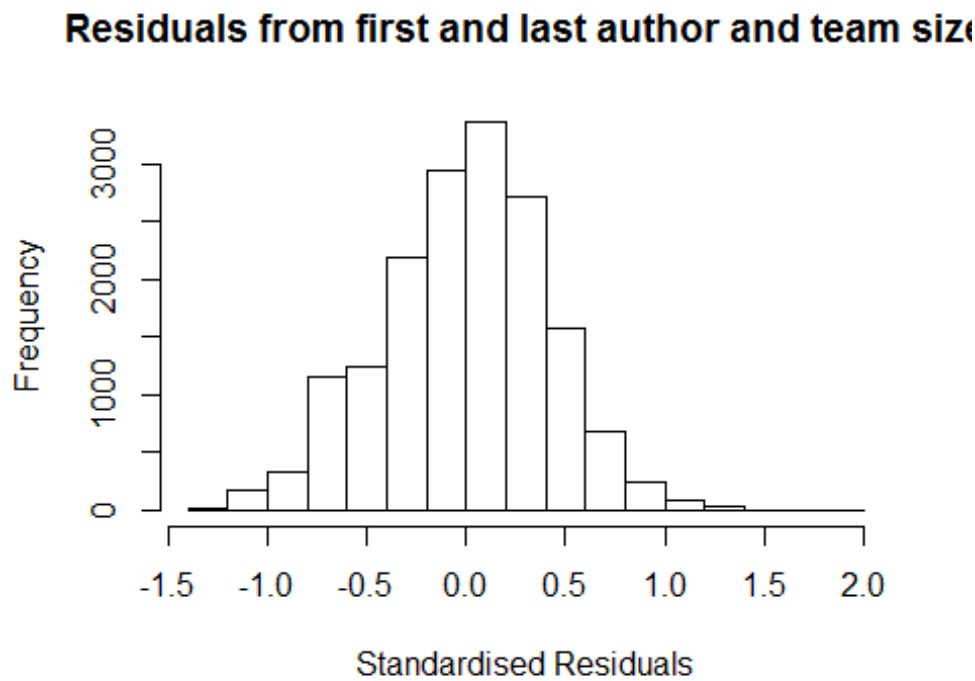
```
## 1250 1192
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 1009  913  935  894  766  606 1011 1038 1081 1046 1036 1039 1118 1048 1046
## 2011 2012
## 1097 1075
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
##  Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 170, df = 16, p-value <2e-16
```




```
##
##  Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 42, df = 1, p-value = 1e-10
```

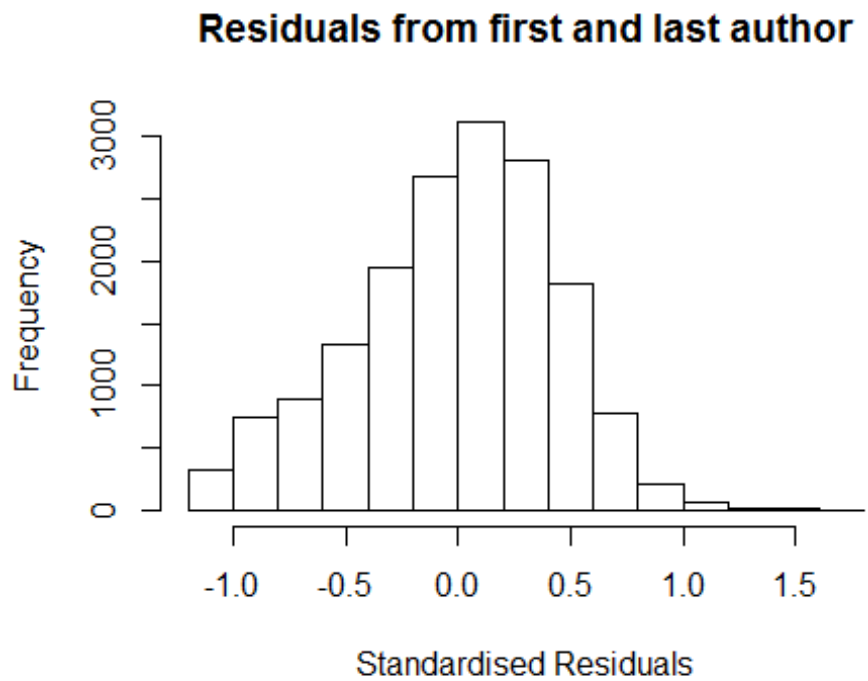



```
## [1] "Female first author team size 2018 geometric mean: 3.71445681807482"
## [1] "Male first author team size 2018 geometric mean: 3.13803164839961"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 120000, p-value = 9e-04
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 3.71477681478097"
## [1] "Male last author team size 2018 geometric mean: 3.24095845930526"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 95000, p-value = 0.003
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##                     GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.039  1           1.019
## LastAuthorFemale  1.016  1           1.008
## UniqueAuthors     1.072  4           1.009
## Year              1.055 16           1.002
```



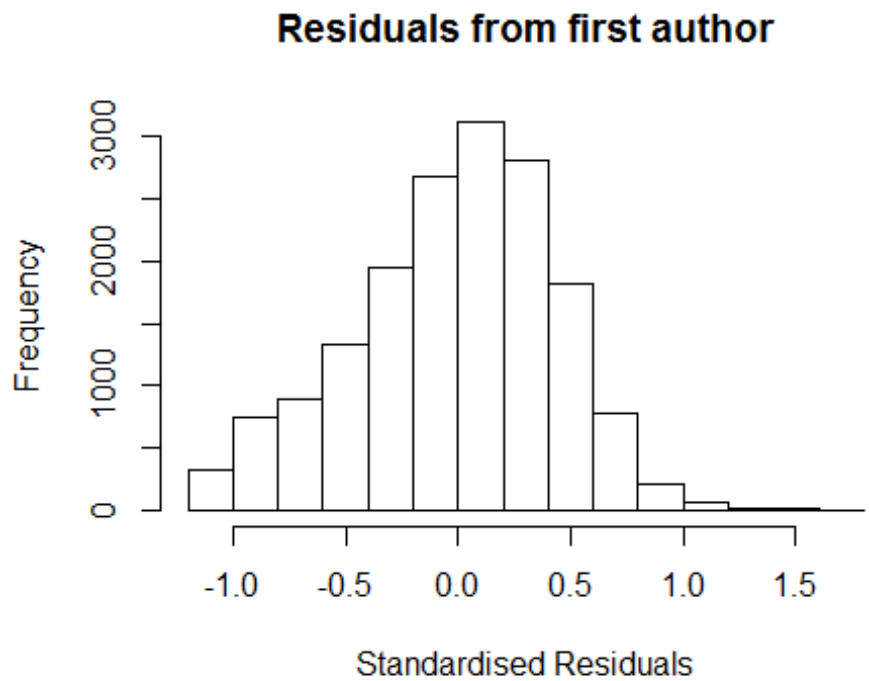

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
UniqueAuthors +
##     Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##     k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.3307 -0.2743  0.0188  0.2734  1.9948
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         0.72626    0.01603   45.31  < 2e-16 ***
## FirstAuthorFemale1  0.04032    0.00730    5.52  3.4e-08 ***
## LastAuthorFemale1  -0.01053    0.00849   -1.24  0.21500
## UniqueAuthors2      0.29160    0.01217   23.97  < 2e-16 ***
## UniqueAuthors3      0.36892    0.01223   30.16  < 2e-16 ***
## UniqueAuthors4      0.41182    0.01299   31.71  < 2e-16 ***
## UniqueAuthors5      0.56106    0.01192   47.06  < 2e-16 ***
## Year1997           -0.03597    0.01976   -1.82  0.06869 .
## Year1998            0.00302    0.01959    0.15  0.87736
## Year1999           -0.00523    0.01923   -0.27  0.78555
```

```
## Year2000           -0.01242    0.01948   -0.64  0.52390
## Year2001           -0.00809    0.02142   -0.38  0.70565
## Year2002           -0.05489    0.01797   -3.05  0.00226 **
## Year2003           -0.06043    0.01797   -3.36  0.00077 ***
## Year2004           -0.07988    0.01757   -4.55  5.5e-06 ***
## Year2005           -0.07066    0.01822   -3.88  0.00011 ***
## Year2006           -0.06911    0.01879   -3.68  0.00024 ***
## Year2007           -0.07880    0.01841   -4.28  1.9e-05 ***
## Year2008           -0.07630    0.01925   -3.96  7.4e-05 ***
## Year2009           -0.08211    0.01948   -4.22  2.5e-05 ***
## Year2010           -0.11081    0.01963   -5.64  1.7e-08 ***
## Year2011           -0.11811    0.01961   -6.02  1.7e-09 ***
## Year2012           -0.11813    0.01988   -5.94  2.9e-09 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.406
## Multiple R-squared:  0.163,  Adjusted R-squared:  0.162
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
##  observation 14778 is an outlier with |weight| = 0 ( < 6e-06);
##  1409 weights are ~= 1. The remaining 15348 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.007   0.867   0.950   0.903   0.985   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          5.97e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.017  1           1.009
## LastAuthorFemale  1.005  1           1.002
## Year              1.021 16           1.001
```



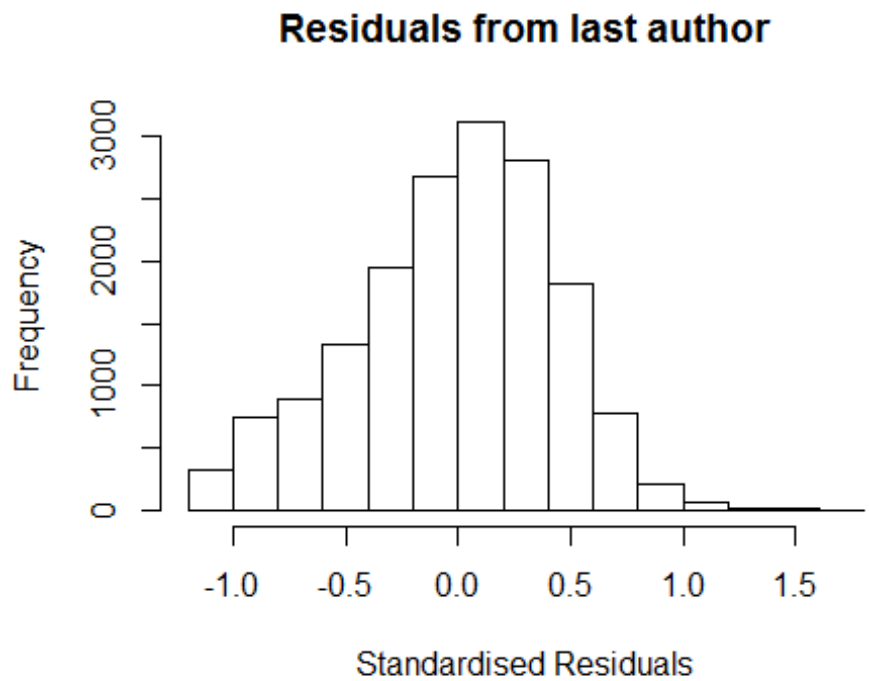

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##    Min     1Q Median     3Q    Max
## -1.118 -0.298  0.032  0.301  1.709
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)          1.01204    0.01450   69.78   <2e-16 ***
## FirstAuthorFemale1   0.09179    0.00767   11.97   <2e-16 ***
## LastAuthorFemale1   -0.00360    0.00897   -0.40   0.6884
## Year1997            -0.02822    0.02119   -1.33   0.1829
## Year1998             0.01388    0.02103    0.66   0.5094
## Year1999             0.01324    0.02068    0.64   0.5222
## Year2000             0.00602    0.02088    0.29   0.7730
## Year2001            -0.00635    0.02249   -0.28   0.7778
## Year2002            -0.02422    0.01943   -1.25   0.2127
## Year2003            -0.03284    0.01959   -1.68   0.0936 .
## Year2004            -0.04427    0.01900   -2.33   0.0198 *
## Year2005            -0.02515    0.01954   -1.29   0.1982
```

```
## Year2006           -0.02216    0.02019   -1.10   0.2724
## Year2007           -0.03655    0.02015   -1.81   0.0697 .
## Year2008           -0.02962    0.02062   -1.44   0.1508
## Year2009           -0.02240    0.02068   -1.08   0.2787
## Year2010           -0.06154    0.02165   -2.84   0.0045 **
## Year2011           -0.06634    0.02147   -3.09   0.0020 **
## Year2012           -0.05925    0.02177   -2.72   0.0065 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.433
## Multiple R-squared:  0.0104, Adjusted R-squared:  0.00931
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
##  1314 weights are ~= 1. The remaining 15444 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0851  0.8660  0.9490  0.9010  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          5.97e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.017  1           1.009
## Year              1.017 16           1.001
```




```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.1168 -0.2986  0.0317  0.3010  1.7100
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)          1.01147    0.01441   70.17   <2e-16 ***
## FirstAuthorFemale1   0.09136    0.00773   11.83   <2e-16 ***
## Year1997            -0.02819    0.02119   -1.33   0.1834
## Year1998             0.01394    0.02103    0.66   0.5075
## Year1999             0.01328    0.02068    0.64   0.5207
## Year2000             0.00617    0.02088    0.30   0.7674
## Year2001            -0.00627    0.02249   -0.28   0.7804
## Year2002            -0.02424    0.01943   -1.25   0.2124
## Year2003            -0.03279    0.01959   -1.67   0.0941 .
## Year2004            -0.04424    0.01900   -2.33   0.0199 *
## Year2005            -0.02507    0.01954   -1.28   0.1997
## Year2006            -0.02218    0.02019   -1.10   0.2720
```

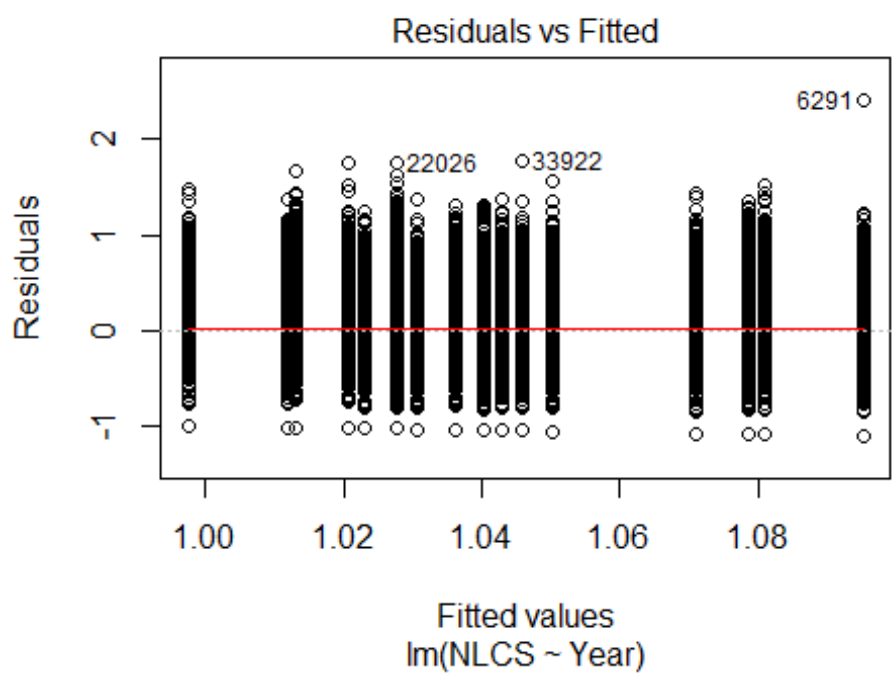
```
## Year2007           -0.03661    0.02014   -1.82   0.0691 .
## Year2008           -0.02963    0.02061   -1.44   0.1507
## Year2009           -0.02240    0.02068   -1.08   0.2787
## Year2010           -0.06155    0.02165   -2.84   0.0045 **
## Year2011           -0.06643    0.02146   -3.10   0.0020 **
## Year2012           -0.05928    0.02177   -2.72   0.0065 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.433
## Multiple R-squared:  0.0104, Adjusted R-squared:  0.00937
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
##  1325 weights are ~= 1. The remaining 15433 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0847  0.8650  0.9490  0.9010  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          5.97e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##                   GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.005  1           1.003
## Year             1.005 16           1.000
```



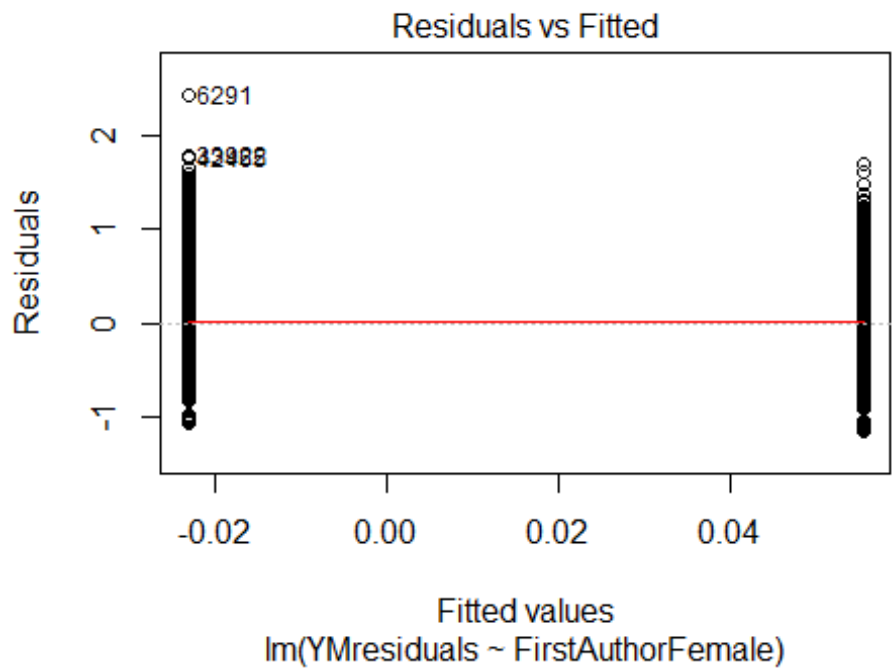

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.0561 -0.3035  0.0326  0.3028  1.6843
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.03187    0.01445   71.43   <2e-16 ***
## LastAuthorFemale1   0.01083    0.00897    1.21   0.2272
## Year1997           -0.02831    0.02123   -1.33   0.1824
## Year1998            0.01342    0.02111    0.64   0.5250
## Year1999            0.01272    0.02079    0.61   0.5408
## Year2000            0.00718    0.02096    0.34   0.7318
## Year2001           -0.01004    0.02251   -0.45   0.6556
## Year2002           -0.02286    0.01953   -1.17   0.2419
## Year2003           -0.03145    0.01965   -1.60   0.1095
## Year2004           -0.04175    0.01910   -2.19   0.0289 *
## Year2005           -0.02039    0.01963   -1.04   0.2990
## Year2006           -0.01941    0.02029   -0.96   0.3388
```

```
## Year2007          -0.03116    0.02026   -1.54   0.1241
## Year2008          -0.02760    0.02075   -1.33   0.1835
## Year2009          -0.01575    0.02082   -0.76   0.4494
## Year2010          -0.05288    0.02176   -2.43   0.0151 *
## Year2011          -0.06113    0.02158   -2.83   0.0046 **
## Year2012          -0.04902    0.02192   -2.24   0.0253 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.435
## Multiple R-squared:  0.00228,    Adjusted R-squared:  0.00127
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
##  1349 weights are ~= 1. The remaining 15409 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.101   0.867   0.948   0.901   0.985   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          5.97e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Sample size for the above analysis:  16758"
## [1] ""
## [1] ""
## [1] "####################################"
## [1] "Analysis of AJSC 1110"
## [1] "####################################"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 2877 2737 2653 2505 2378 2524 2583 2299 2511 2404 2429 2493 2461 2371 2314
## 2011 2012
## 2241 2375
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 1611 1540 1520 1501 1080  815 1614 1428 1582 1569 1609 1713 1653 1654 1612
## 2011 2012
```

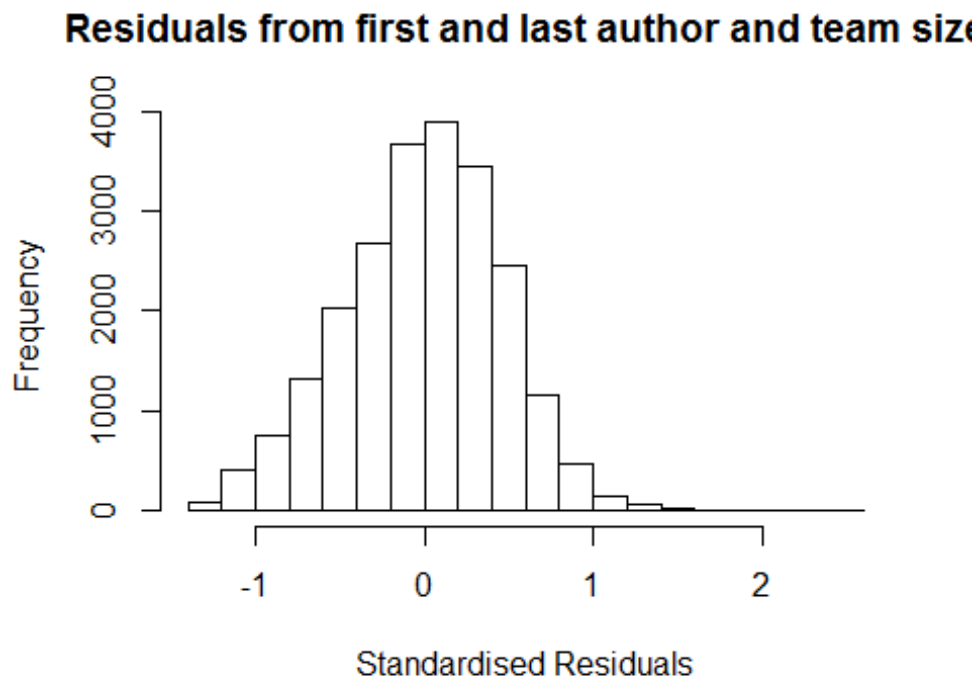
```
## 1540 1662
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 1416 1388 1353 1337  972  702 1382 1261 1390 1410 1426 1496 1447 1442 1417
## 2011 2012
## 1325 1444
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
##  Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 95, df = 16, p-value = 3e-13
```




```
##
##  Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 26, df = 1, p-value = 3e-07
```

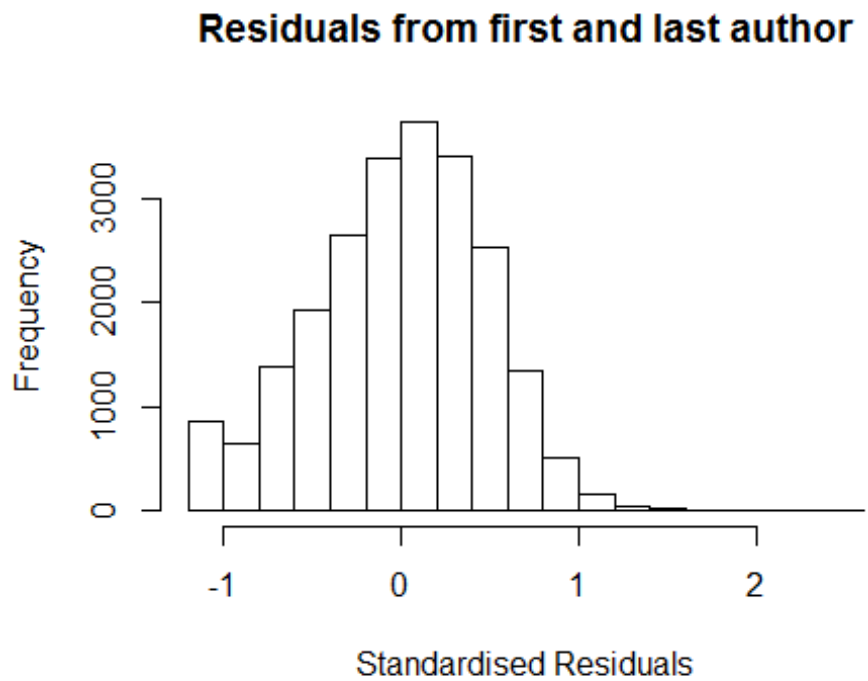



```
## [1] "Female first author team size 2018 geometric mean: 3.55309722332158"
## [1] "Male first author team size 2018 geometric mean: 3.2764675578008"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 2e+05, p-value = 0.05
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 3.35415454055799"
## [1] "Male last author team size 2018 geometric mean: 3.38319508664327"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 170000, p-value = 0.7
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.045  1           1.022
## LastAuthorFemale  1.037  1           1.018
## UniqueAuthors     1.057  4           1.007
## Year              1.063 16           1.002
```



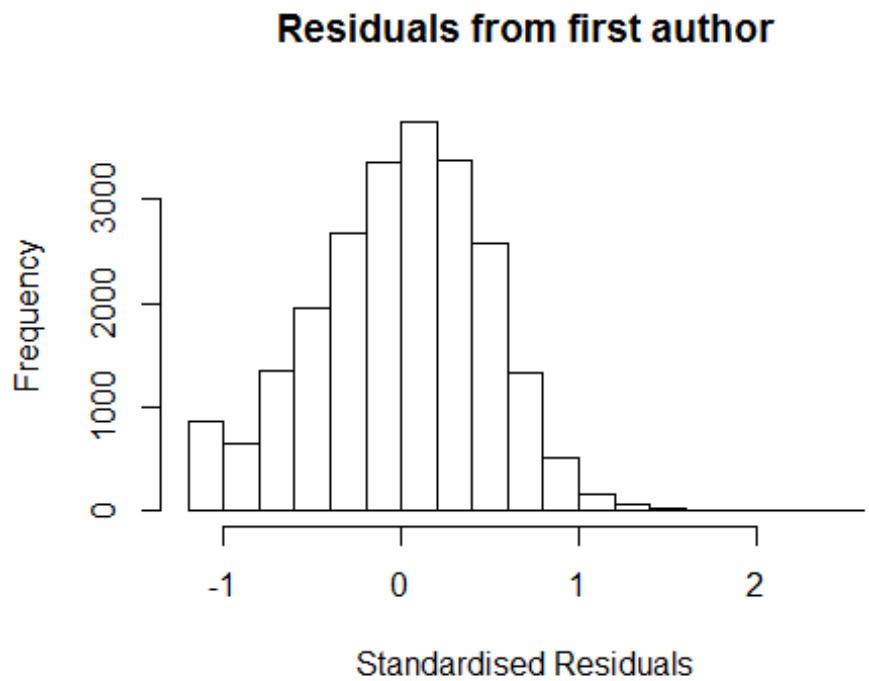

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
UniqueAuthors +
##     Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##     k.max = 1000))
##  \--> method = "MM"
## Residuals:
##    Min     1Q Median     3Q    Max
## -1.389 -0.314  0.018  0.314  2.445
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)          0.84198    0.01528   55.10  < 2e-16 ***
## FirstAuthorFemale1   0.05958    0.00696    8.56  < 2e-16 ***
## LastAuthorFemale1    0.04470    0.00765    5.84  5.3e-09 ***
## UniqueAuthors2       0.20943    0.01034   20.25  < 2e-16 ***
## UniqueAuthors3       0.27448    0.01068   25.71  < 2e-16 ***
## UniqueAuthors4       0.36311    0.01180   30.77  < 2e-16 ***
## UniqueAuthors5       0.48695    0.01156   42.11  < 2e-16 ***
## Year1997             0.00216    0.01871    0.12  0.90805
## Year1998             0.00905    0.01857    0.49  0.62621
## Year1999            -0.04084    0.01877   -2.18  0.02958 *
```

```
## Year2000           -0.03178    0.01994   -1.59  0.11096
## Year2001           -0.06997    0.02215   -3.16  0.00158 **
## Year2002           -0.07113    0.01846   -3.85  0.00012 ***
## Year2003           -0.07880    0.01804   -4.37  1.3e-05 ***
## Year2004           -0.10330    0.01824   -5.66  1.5e-08 ***
## Year2005           -0.12760    0.01849   -6.90  5.4e-12 ***
## Year2006           -0.10004    0.01816   -5.51  3.7e-08 ***
## Year2007           -0.09360    0.01813   -5.16  2.4e-07 ***
## Year2008           -0.07587    0.01811   -4.19  2.8e-05 ***
## Year2009           -0.10269    0.01867   -5.50  3.8e-08 ***
## Year2010           -0.11931    0.01923   -6.21  5.5e-10 ***
## Year2011           -0.14132    0.01962   -7.20  6.1e-13 ***
## Year2012           -0.15046    0.01918   -7.84  4.5e-15 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.459
## Multiple R-squared:  0.102,  Adjusted R-squared:  0.101
## Convergence in 10 IRWLS iterations
##
## Robustness weights:
##  observation 2827 is an outlier with |weight| = 0 ( < 4.4e-06);
##  1936 weights are ~= 1. The remaining 20671 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0634  0.8700  0.9490  0.9060  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          4.42e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.034  1           1.017
## LastAuthorFemale  1.030  1           1.015
## Year              1.020 16           1.001
```



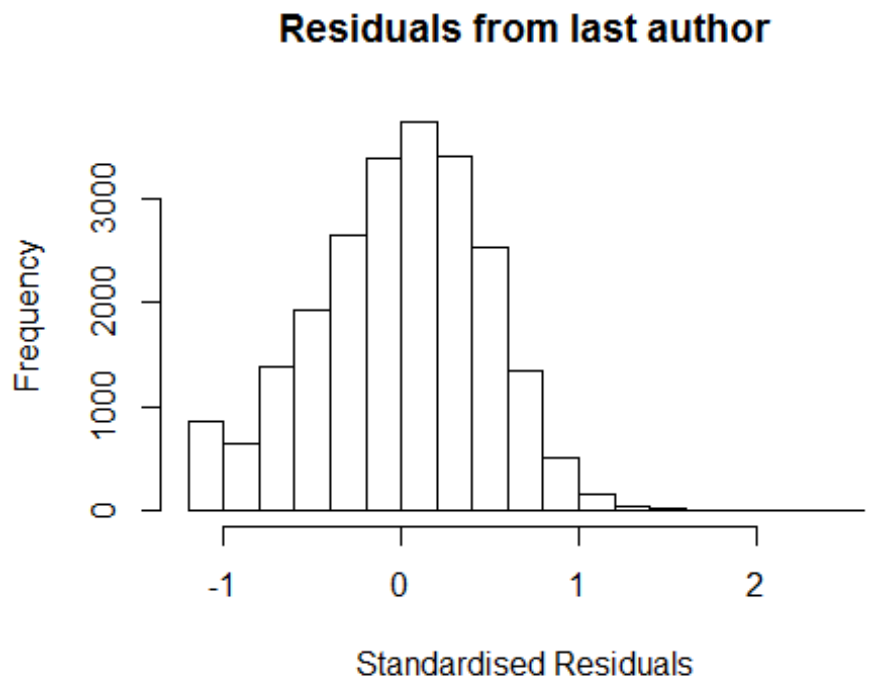

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.1803 -0.3279  0.0264  0.3340  2.4426
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)          1.05219    0.01427   73.74  < 2e-16 ***
## FirstAuthorFemale1   0.08331    0.00724   11.51  < 2e-16 ***
## LastAuthorFemale1    0.03450    0.00801    4.31  1.7e-05 ***
## Year1997             0.00722    0.01946    0.37  0.71067
## Year1998             0.01025    0.01958    0.52  0.60055
## Year1999            -0.03192    0.01953   -1.63  0.10230
## Year2000            -0.01594    0.02052   -0.78  0.43727
## Year2001            -0.04455    0.02274   -1.96  0.05011 .
## Year2002            -0.04568    0.01942   -2.35  0.01866 *
## Year2003            -0.05130    0.01917   -2.68  0.00747 **
## Year2004            -0.07070    0.01932   -3.66  0.00025 ***
## Year2005            -0.08557    0.01961   -4.36  1.3e-05 ***
```

```
## Year2006           -0.05460    0.01902   -2.87  0.00411 **
## Year2007           -0.04088    0.01892   -2.16  0.03071 *
## Year2008           -0.02121    0.01908   -1.11  0.26642
## Year2009           -0.04280    0.01965   -2.18  0.02940 *
## Year2010           -0.06467    0.02055   -3.15  0.00165 **
## Year2011           -0.07260    0.02066   -3.51  0.00044 ***
## Year2012           -0.07693    0.02029   -3.79  0.00015 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.483
## Multiple R-squared:  0.0107, Adjusted R-squared:  0.00995
## Convergence in 10 IRWLS iterations
##
## Robustness weights:
##  observation 2827 is an outlier with |weight| = 0 ( < 4.4e-06);
##  1887 weights are ~= 1. The remaining 20720 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.131   0.871   0.948   0.907   0.985   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          4.42e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                  psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.013  1           1.006
## Year              1.013 16           1.000
```




```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##    Min     1Q Median     3Q    Max
## -1.157 -0.328  0.025  0.335  2.438
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)          1.05761    0.01422   74.38  < 2e-16 ***
## FirstAuthorFemale1   0.08980    0.00717   12.52  < 2e-16 ***
## Year1997             0.00715    0.01948    0.37  0.71371
## Year1998             0.00979    0.01960    0.50  0.61752
## Year1999            -0.03209    0.01956   -1.64  0.10086
## Year2000            -0.01541    0.02051   -0.75  0.45264
## Year2001            -0.04441    0.02276   -1.95  0.05104 .
## Year2002            -0.04523    0.01941   -2.33  0.01980 *
## Year2003            -0.05102    0.01916   -2.66  0.00777 **
## Year2004            -0.07099    0.01933   -3.67  0.00024 ***
## Year2005            -0.08634    0.01962   -4.40  1.1e-05 ***
## Year2006            -0.05435    0.01903   -2.86  0.00429 **
```

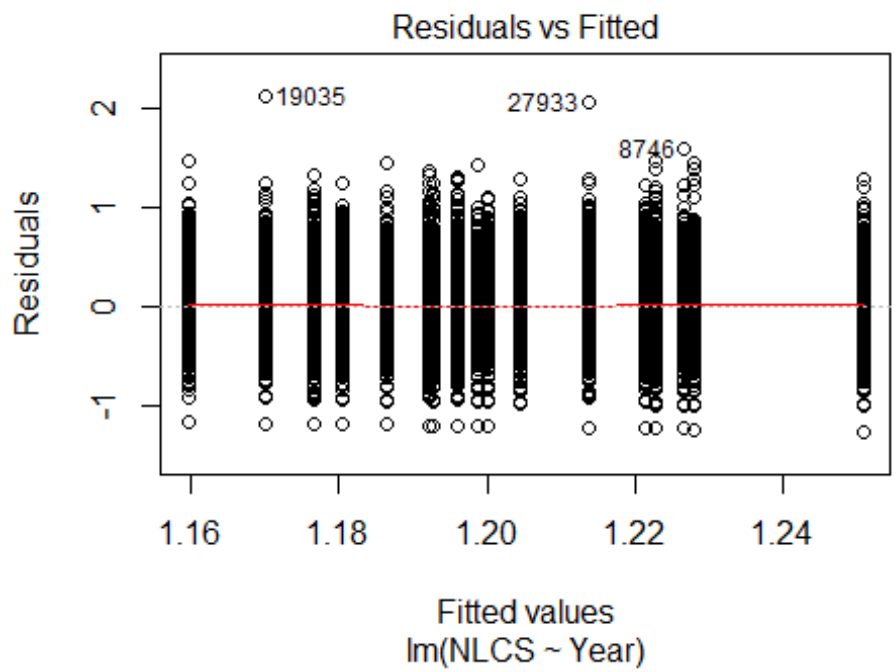
```
## Year2007           -0.04053    0.01892   -2.14  0.03222 *
## Year2008           -0.02075    0.01909   -1.09  0.27706
## Year2009           -0.04166    0.01965   -2.12  0.03405 *
## Year2010           -0.06303    0.02055   -3.07  0.00216 **
## Year2011           -0.07109    0.02068   -3.44  0.00059 ***
## Year2012           -0.07540    0.02028   -3.72  0.00020 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.484
## Multiple R-squared:  0.00984,    Adjusted R-squared:  0.00909
## Convergence in 10 IRWLS iterations
##
## Robustness weights:
##  observation 2827 is an outlier with |weight| = 0 ( < 4.4e-06);
##  1832 weights are ~= 1. The remaining 20775 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.136   0.871   0.949   0.907   0.985   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          4.42e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##                   GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.009  1           1.004
## Year             1.009 16           1.000
```



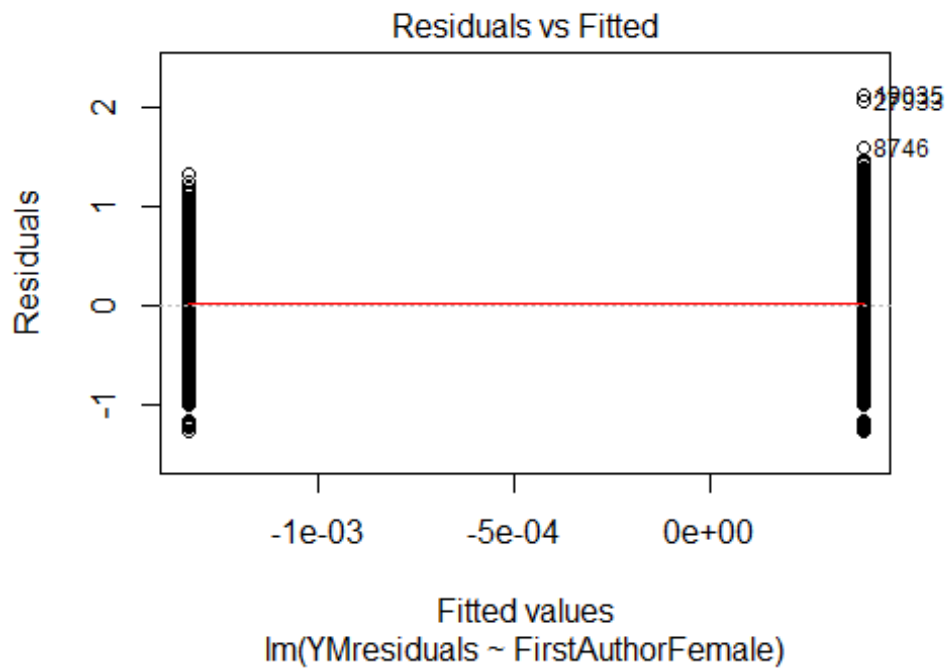

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.1343 -0.3314  0.0254  0.3354  2.4239
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.06802    0.01428   74.77  < 2e-16 ***
## LastAuthorFemale1   0.05329    0.00791    6.73  1.7e-11 ***
## Year1997            0.00755    0.01954    0.39  0.69935
## Year1998            0.01303    0.01962    0.66  0.50652
## Year1999           -0.03066    0.01961   -1.56  0.11796
## Year2000           -0.01302    0.02060   -0.63  0.52754
## Year2001           -0.04139    0.02276   -1.82  0.06899 .
## Year2002           -0.04041    0.01946   -2.08  0.03789 *
## Year2003           -0.04718    0.01928   -2.45  0.01441 *
## Year2004           -0.06698    0.01943   -3.45  0.00057 ***
## Year2005           -0.08013    0.01975   -4.06  5.0e-05 ***
## Year2006           -0.04930    0.01912   -2.58  0.00992 **
```

```
## Year2007          -0.03367    0.01899   -1.77  0.07631 .
## Year2008          -0.01439    0.01917   -0.75  0.45273
## Year2009          -0.03770    0.01969   -1.91  0.05561 .
## Year2010          -0.05681    0.02058   -2.76  0.00577 **
## Year2011          -0.06216    0.02072   -3.00  0.00270 **
## Year2012          -0.06800    0.02034   -3.34  0.00083 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.485
## Multiple R-squared:  0.00494,    Adjusted R-squared:  0.0042
## Convergence in 10 IRWLS iterations
##
## Robustness weights:
##  observation 2827 is an outlier with |weight| = 0 ( < 4.4e-06);
##  1852 weights are ~= 1. The remaining 20755 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.147   0.872   0.949   0.907   0.985   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          4.42e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                    psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Sample size for the above analysis:  22608"
## [1] ""
## [1] ""
## [1] "####################################"
## [1] "Analysis of AJSC 1111"
## [1] "####################################"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 1235 1455 1343 1314 1186 1448 1285 1508 1247 1267 1523 1566 1584 1400 1405
## 2011 2012
## 1422 1501
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  553  626  570  577  484  557  650  685  583  616  807  897  823  718  670
```

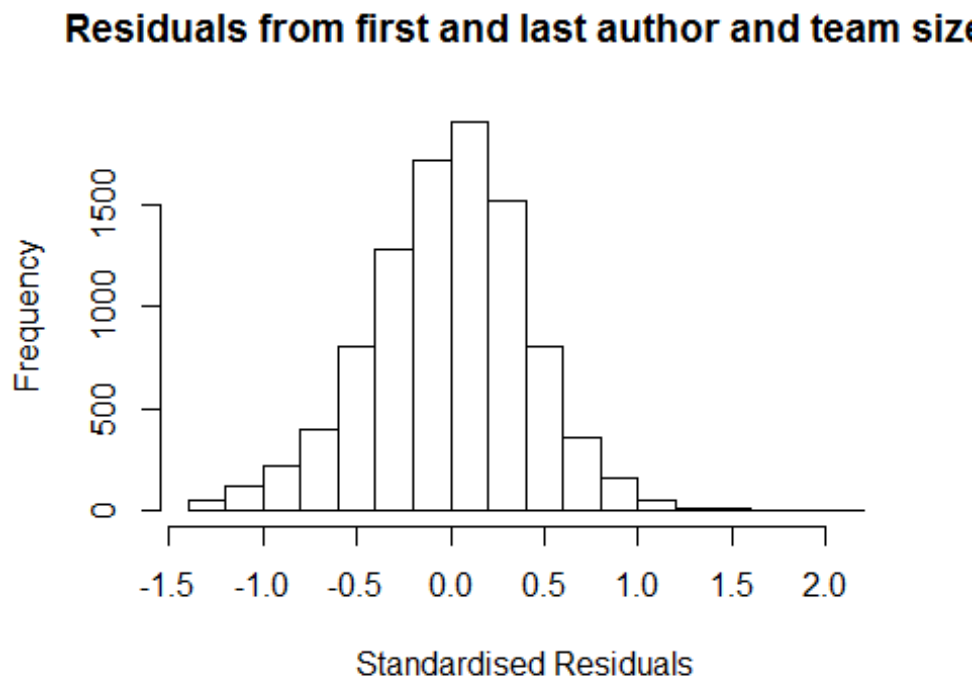
```
## 2011 2012
##  702  728
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  484  558  481  494  422  466  544  573  480  511  660  729  659  608  551
## 2011 2012
##  595  588
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
##  Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 61, df = 16, p-value = 3e-07
```




```
##
##  Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 14, df = 1, p-value = 2e-04
```

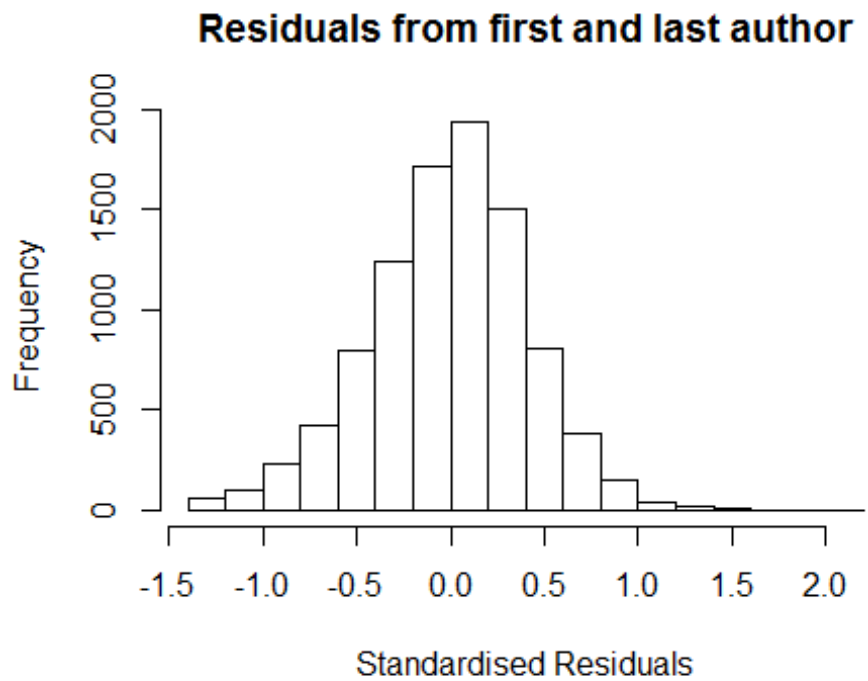



```
## [1] "Female first author team size 2018 geometric mean: 4.05957545854093"
## [1] "Male first author team size 2018 geometric mean: 3.39731424352347"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 54000, p-value = 0.003
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 3.67593038378301"
## [1] "Male last author team size 2018 geometric mean: 3.5644591274955"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 39000, p-value = 0.8
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.057  1           1.028
## LastAuthorFemale  1.033  1           1.016
## UniqueAuthors     1.095  4           1.011
## Year              1.107 16           1.003
```



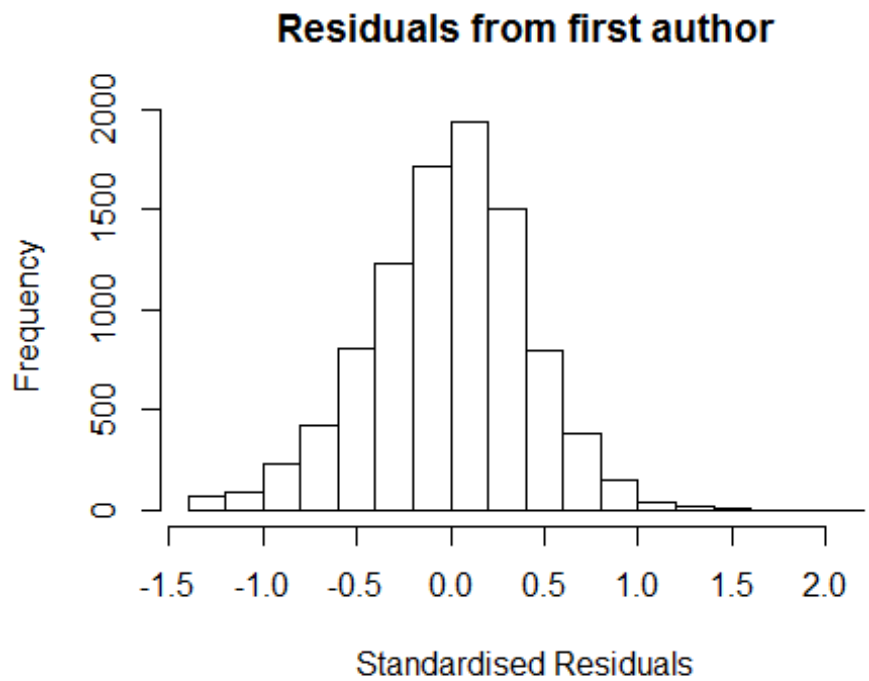

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
UniqueAuthors +
##     Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##     k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.3596 -0.2700  0.0117  0.2639  2.1456
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.19441    0.02265   52.73  < 2e-16 ***
## FirstAuthorFemale1 -0.00728    0.01002   -0.73  0.46709
## LastAuthorFemale1   0.00650    0.01199    0.54  0.58753
## UniqueAuthors2      0.07790    0.01540    5.06  4.3e-07 ***
## UniqueAuthors3      0.09301    0.01568    5.93  3.1e-09 ***
## UniqueAuthors4      0.12125    0.01662    7.29  3.3e-13 ***
## UniqueAuthors5      0.19230    0.01611   11.93  < 2e-16 ***
## Year1997           -0.02712    0.02666   -1.02  0.30923
## Year1998           -0.03168    0.02758   -1.15  0.25068
## Year1999           -0.07905    0.02721   -2.91  0.00367 **
```

```
## Year2000           -0.07227    0.02810   -2.57  0.01012 *
## Year2001           -0.05037    0.02840   -1.77  0.07616 .
## Year2002           -0.09693    0.02845   -3.41  0.00066 ***
## Year2003           -0.09494    0.02583   -3.68  0.00024 ***
## Year2004           -0.12226    0.02702   -4.53  6.1e-06 ***
## Year2005           -0.08104    0.02611   -3.10  0.00192 **
## Year2006           -0.12805    0.02508   -5.11  3.4e-07 ***
## Year2007           -0.13083    0.02485   -5.26  1.4e-07 ***
## Year2008           -0.10933    0.02597   -4.21  2.6e-05 ***
## Year2009           -0.08001    0.02568   -3.12  0.00184 **
## Year2010           -0.10080    0.02680   -3.76  0.00017 ***
## Year2011           -0.10711    0.02729   -3.93  8.7e-05 ***
## Year2012           -0.06999    0.02678   -2.61  0.00897 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.394
## Multiple R-squared:  0.0231, Adjusted R-squared:  0.0208
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  2 observations c(6161,9096) are outliers with |weight| = 0 ( < 1.1e-05);
##  793 weights are ~= 1. The remaining 8608 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.033   0.868   0.951   0.897   0.985   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.06e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.045  1           1.022
## LastAuthorFemale  1.028  1           1.014
## Year              1.029 16           1.001
```



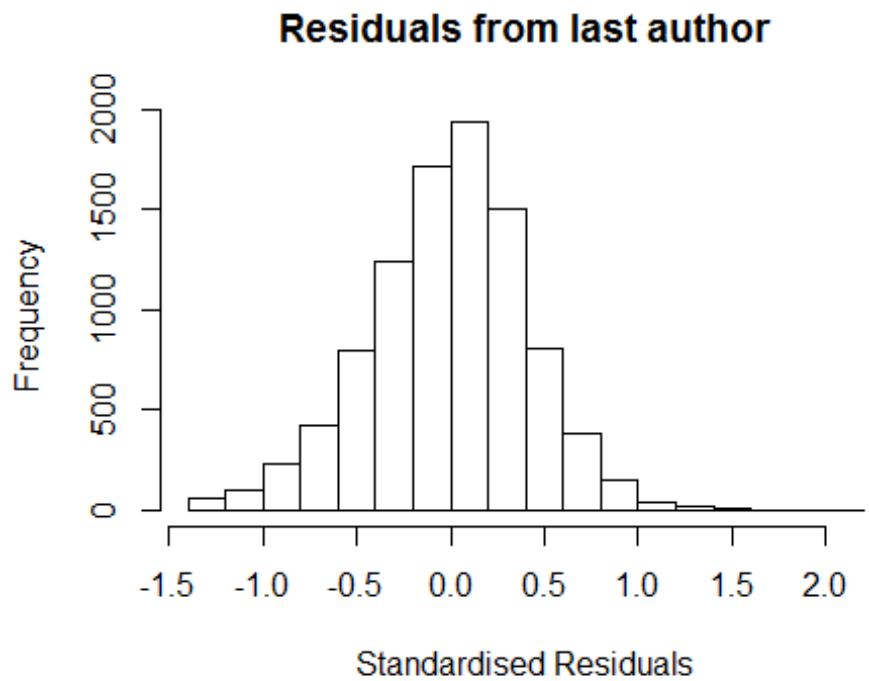

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.2837 -0.2697  0.0142  0.2691  2.1031
##
## Coefficients:
##                      Estimate Std. Error t value Pr(>|t|)
## (Intercept)          1.271416   0.019697   64.55  < 2e-16 ***
## FirstAuthorFemale1   0.000421   0.010076    0.04  0.96668
## LastAuthorFemale1    0.011856   0.012083    0.98  0.32650
## Year1997            -0.032428   0.026746   -1.21  0.22538
## Year1998            -0.022858   0.027288   -0.84  0.40225
## Year1999            -0.071999   0.027121   -2.65  0.00795 **
## Year2000            -0.067944   0.028101   -2.42  0.01563 *
## Year2001            -0.029859   0.028401   -1.05  0.29312
## Year2002            -0.078995   0.028402   -2.78  0.00543 **
## Year2003            -0.067859   0.025753   -2.63  0.00843 **
## Year2004            -0.093869   0.026876   -3.49  0.00048 ***
## Year2005            -0.058663   0.026035   -2.25  0.02427 *
```

```
## Year2006           -0.098422   0.024955   -3.94  8.1e-05 ***
## Year2007           -0.099342   0.024664   -4.03  5.7e-05 ***
## Year2008           -0.079519   0.025894   -3.07  0.00214 **
## Year2009           -0.043744   0.025493   -1.72  0.08621 .
## Year2010           -0.072632   0.026578   -2.73  0.00629 **
## Year2011           -0.073206   0.027295   -2.68  0.00733 **
## Year2012           -0.034238   0.026755   -1.28  0.20070
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.397
## Multiple R-squared:  0.00465,    Adjusted R-squared:  0.00274
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  2 observations c(6161,9096) are outliers with |weight| = 0 ( < 1.1e-05);
##  772 weights are ~= 1. The remaining 8629 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0809  0.8660  0.9500  0.8970  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.06e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.023  1           1.011
## Year              1.023 16           1.001
```




```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.2747 -0.2701  0.0137  0.2692  2.1136
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)          1.27281    0.01966   64.74  < 2e-16 ***
## FirstAuthorFemale1  0.00190    0.00997    0.19  0.84900
## Year1997           -0.03256    0.02675   -1.22  0.22361
## Year1998           -0.02288    0.02729   -0.84  0.40180
## Year1999           -0.07192    0.02714   -2.65  0.00805 **
## Year2000           -0.06788    0.02808   -2.42  0.01565 *
## Year2001           -0.02980    0.02841   -1.05  0.29415
## Year2002           -0.07899    0.02840   -2.78  0.00543 **
## Year2003           -0.06760    0.02575   -2.63  0.00867 **
## Year2004           -0.09374    0.02688   -3.49  0.00049 ***
## Year2005           -0.05880    0.02604   -2.26  0.02398 *
## Year2006           -0.09849    0.02496   -3.95  8.0e-05 ***
```

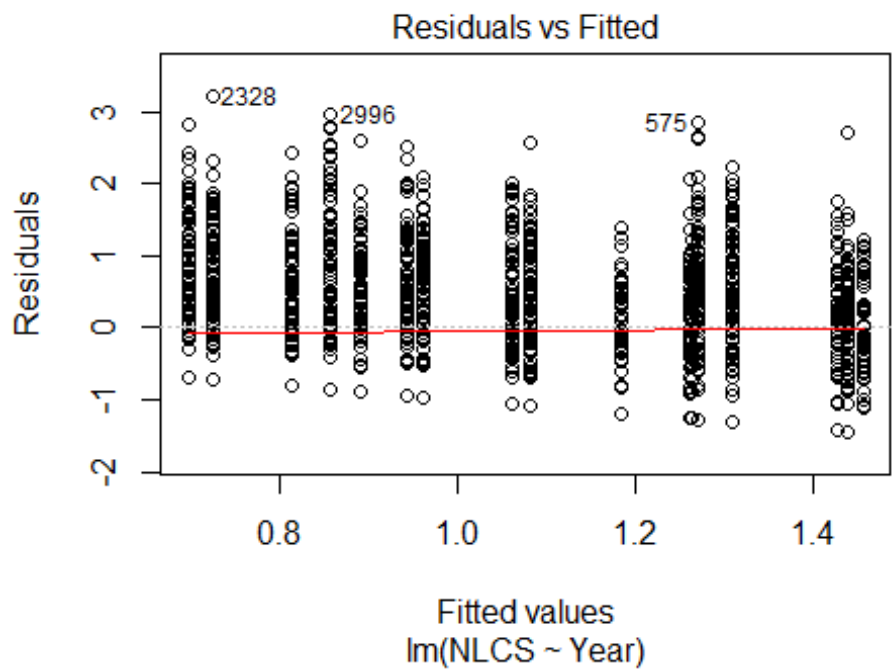
```
## Year2007           -0.09940    0.02468   -4.03  5.7e-05 ***
## Year2008           -0.07938    0.02590   -3.06  0.00219 **
## Year2009           -0.04346    0.02549   -1.71  0.08815 .
## Year2010           -0.07230    0.02658   -2.72  0.00653 **
## Year2011           -0.07317    0.02731   -2.68  0.00740 **
## Year2012           -0.03359    0.02676   -1.26  0.20948
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.397
## Multiple R-squared:  0.00455,    Adjusted R-squared:  0.00274
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  2 observations c(6161,9096) are outliers with |weight| = 0 ( < 1.1e-05);
##  763 weights are ~= 1. The remaining 8638 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0818  0.8660  0.9500  0.8970  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.06e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##                   GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.006  1           1.003
## Year             1.006 16           1.000
```



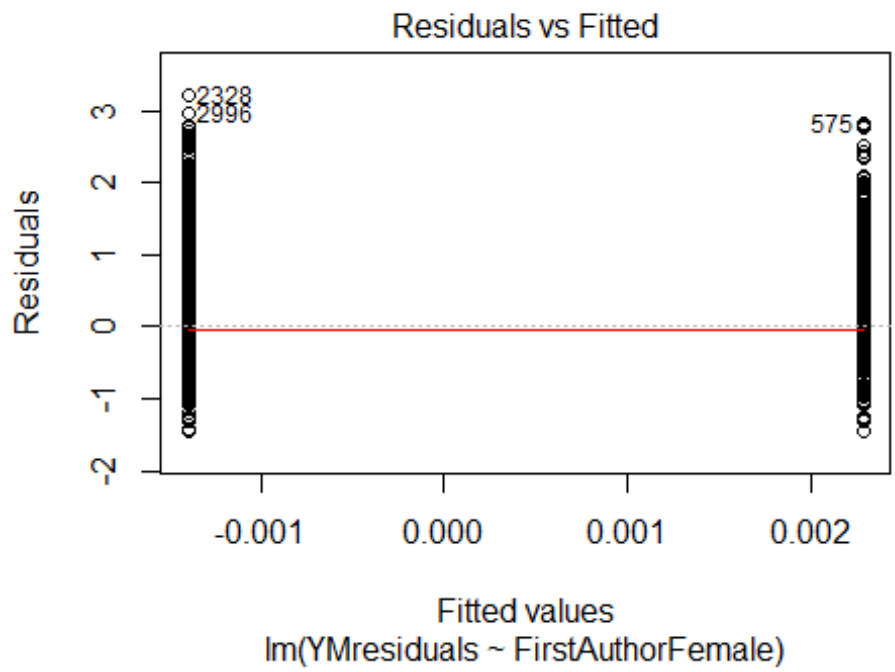

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.2834 -0.2698  0.0141  0.2693  2.1029
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)          1.2715     0.0197   64.70  < 2e-16 ***
## LastAuthorFemale1    0.0119     0.0120    1.00  0.31845
## Year1997            -0.0324     0.0267   -1.21  0.22550
## Year1998            -0.0229     0.0273   -0.84  0.40226
## Year1999            -0.0720     0.0271   -2.66  0.00792 **
## Year2000            -0.0679     0.0281   -2.42  0.01560 *
## Year2001            -0.0298     0.0284   -1.05  0.29334
## Year2002            -0.0790     0.0284   -2.78  0.00542 **
## Year2003            -0.0678     0.0257   -2.63  0.00843 **
## Year2004            -0.0938     0.0268   -3.50  0.00047 ***
## Year2005            -0.0586     0.0260   -2.25  0.02423 *
## Year2006            -0.0984     0.0249   -3.95  8.0e-05 ***
```

```
## Year2007           -0.0993     0.0246   -4.03  5.6e-05 ***
## Year2008           -0.0795     0.0259   -3.07  0.00213 **
## Year2009           -0.0437     0.0255   -1.72  0.08618 .
## Year2010           -0.0726     0.0265   -2.73  0.00626 **
## Year2011           -0.0731     0.0272   -2.69  0.00715 **
## Year2012           -0.0342     0.0267   -1.28  0.20051
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.397
## Multiple R-squared:  0.00465,    Adjusted R-squared:  0.00285
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  2 observations c(6161,9096) are outliers with |weight| = 0 ( < 1.1e-05);
##  771 weights are ~= 1. The remaining 8630 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0808  0.8660  0.9500  0.8970  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.06e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                    psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Sample size for the above analysis:  9403"
## [1] ""
## [1] ""
## [1] "####################################"
## [1] "Analysis of AJSC 1200"
## [1] "####################################"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##   88   70   88   63   72   81  175  165  215  198  199  224  208  189  285
## 2011 2012
##  339  381
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##   64   60   75   53   52   52  152  137  178  170  166  184  168  153  225
```

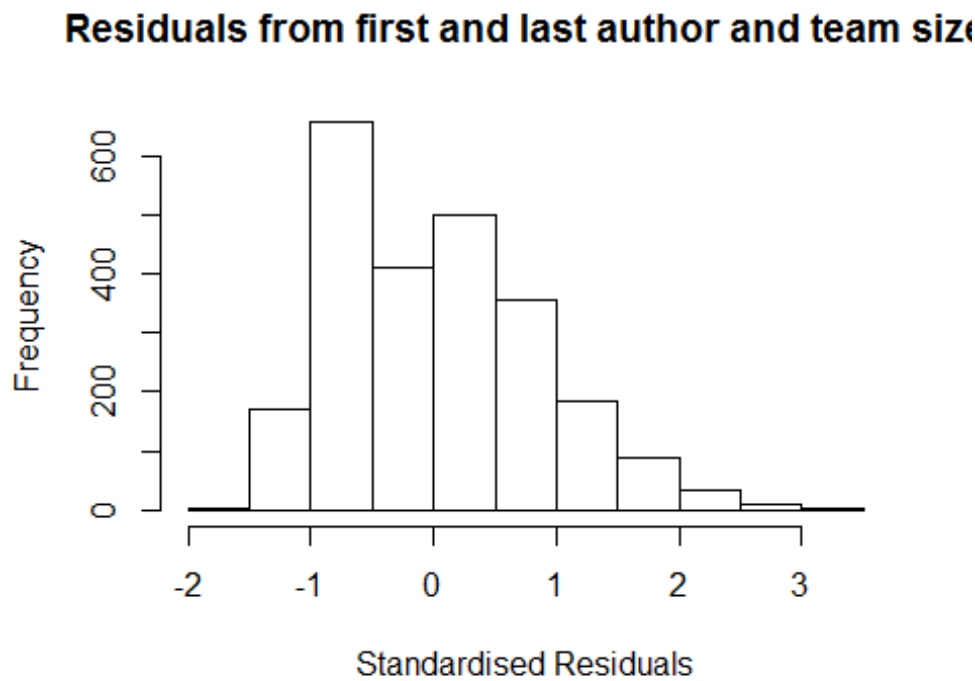
```
## 2011 2012
##  274  308
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##   60   60   73   51   52   52  149  130  177  166  166  179  162  152  221
## 2011 2012
##  269  296
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
##  Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 46, df = 16, p-value = 1e-04
```




```
##
##  Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 2.6, df = 1, p-value = 0.1
```



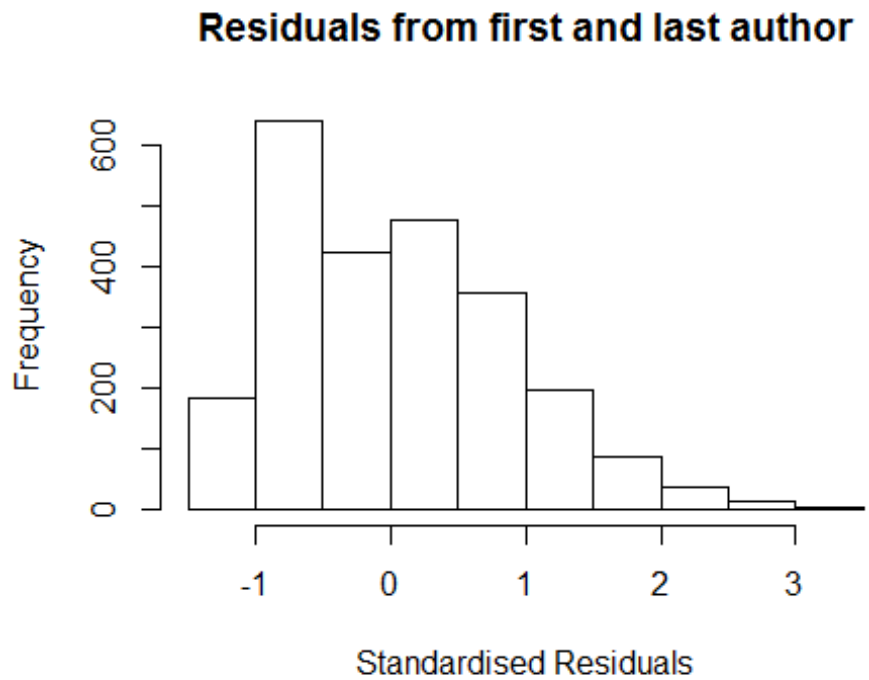

```
## [1] "Female first author team size 2018 geometric mean: 1.1965309174289"
## [1] "Male first author team size 2018 geometric mean: 1.25641672787251"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 6500, p-value = 0.8
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 1.27199591808464"
## [1] "Male last author team size 2018 geometric mean: 1.17859002407924"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 7000, p-value = 0.2
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 5.570  1           2.360
## LastAuthorFemale  5.552  1           2.356
## UniqueAuthors     1.276  4           1.031
## Year              1.274 16           1.008
```




```
## [1] "List of  13 outliers with residuals above  2.5"
##         ScopusId  NLCS Year OneField Fields residuals
## 168   0001180094 4.152 1998     1200      2     2.759
## 525   0038702075 3.925 2002     1200      2     2.501
## 575   0036622707 4.107 2002     1200      2     2.653
## 629  19844382943 3.903 2002     1200      2     2.730
## 1673 33845535808 3.450 2007     1200      2     2.566
## 2328 77951809379 3.936 2010     1200      2     3.336
## 2569 84856757333 3.507 2011     1200      2     2.903
## 2996 84868540587 3.821 2012     1200      2     3.129
## 2998 84868571646 3.628 2012     1200      2     2.906
## 2999 84868575831 3.219 2012     1200      2     2.527
## 3005 84872107429 3.611 2012     1200      2     2.919
## 3013 84863323052 3.244 2012     1200      2     2.552
## 3116 84858681871 3.360 2012     1200      3     2.668
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
UniqueAuthors +
##     Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##     k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.6747 -0.6296 -0.0294  0.6060  3.3362
##
```

```
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.42338    0.10279   13.85  < 2e-16 ***
## FirstAuthorFemale1  0.03433    0.08361    0.41  0.68144
## LastAuthorFemale1  -0.00447    0.08339   -0.05  0.95725
## UniqueAuthors2      0.25129    0.08109    3.10  0.00196 **
## UniqueAuthors3      0.42422    0.11815    3.59  0.00034 ***
## UniqueAuthors4      0.51028    0.22644    2.25  0.02432 *
## UniqueAuthors5      0.54949    0.26236    2.09  0.03633 *
## Year1997           -0.21947    0.14040   -1.56  0.11814
## Year1998           -0.03037    0.13624   -0.22  0.82362
## Year1999           -0.24198    0.15493   -1.56  0.11845
## Year2000           -0.08883    0.14569   -0.61  0.54213
## Year2001           -0.31416    0.13574   -2.31  0.02073 *
## Year2002           -0.25068    0.13538   -1.85  0.06420 .
## Year2003           -0.22556    0.14284   -1.58  0.11444
## Year2004           -0.43218    0.12449   -3.47  0.00053 ***
## Year2005           -0.60269    0.11926   -5.05  4.7e-07 ***
## Year2006           -0.39410    0.12260   -3.21  0.00132 **
## Year2007           -0.56962    0.12159   -4.68  3.0e-06 ***
## Year2008           -0.55871    0.12337   -4.53  6.2e-06 ***
## Year2009           -0.71459    0.12317   -5.80  7.4e-09 ***
## Year2010           -0.82361    0.11669   -7.06  2.2e-12 ***
## Year2011           -0.84901    0.11493   -7.39  2.1e-13 ***
## Year2012           -0.73137    0.11551   -6.33  2.9e-10 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.842
## Multiple R-squared:  0.095,  Adjusted R-squared:  0.0867
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  198 weights are ~= 1. The remaining 2217 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0815  0.8840  0.9400  0.9100  0.9800  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          4.14e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
```

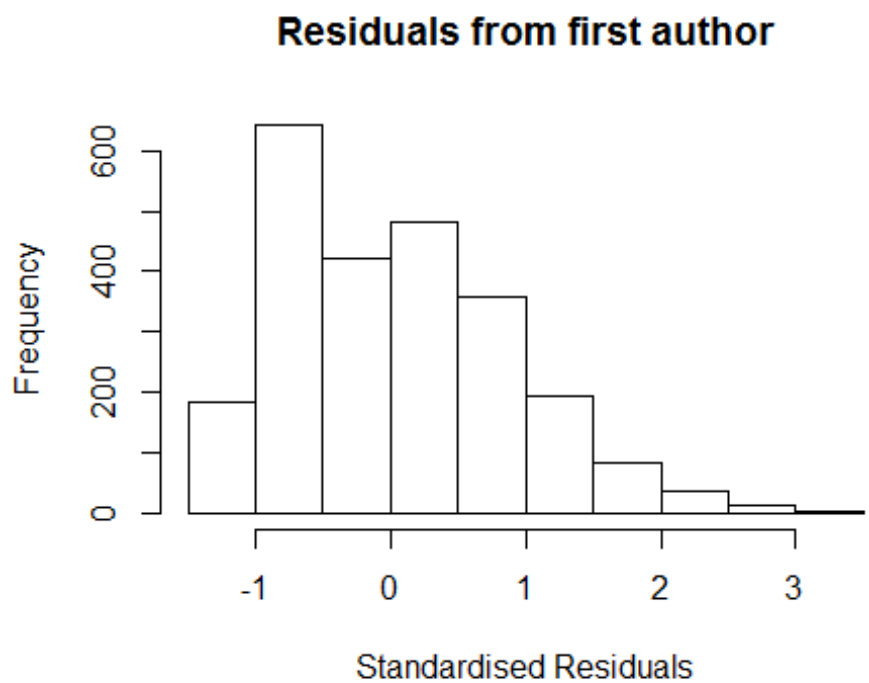
```
##                    "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##                     GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 5.861  1           2.421
## LastAuthorFemale  5.828  1           2.414
## Year              1.089 16           1.003
```




```
## [1] "List of  15 outliers with residuals above  2.5"
##         ScopusId  NLCS Year OneField Fields residuals
## 168   0001180094 4.152 1998     1200      2     2.726
## 525   0038702075 3.925 2002     1200      2     2.735
## 575   0036622707 4.107 2002     1200      2     2.885
## 629  19844382943 3.903 2002     1200      2     2.713
## 1168 33748623359 3.480 2005     1200      2     2.638
## 1409 33746888657 3.658 2006     1200      3     2.616
## 1673 33845535808 3.450 2007     1200      2     2.553
## 2328 77951809379 3.936 2010     1200      2     3.296
## 2569 84856757333 3.507 2011     1200      2     2.873
## 2791 79951671066 3.119 2011     1200      3     2.516
## 2996 84868540587 3.821 2012     1200      2     3.075
## 2998 84868571646 3.628 2012     1200      2     2.851
## 3005 84872107429 3.611 2012     1200      2     2.865
## 3116 84858681871 3.360 2012     1200      3     2.614
## 3150 84861657853 3.650 2012     1200      2     2.904
##
```

```
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.4573 -0.6716 -0.0401  0.6108  3.2956
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)          1.4477     0.1041   13.91  < 2e-16 ***
## FirstAuthorFemale1  -0.0530     0.0866   -0.61  0.54070
## LastAuthorFemale1    0.0841     0.0861    0.98  0.32867
## Year1997            -0.1748     0.1402   -1.25  0.21261
## Year1998            -0.0215     0.1360   -0.16  0.87414
## Year1999            -0.2202     0.1585   -1.39  0.16499
## Year2000            -0.0157     0.1452   -0.11  0.91387
## Year2001            -0.2816     0.1344   -2.09  0.03631 *
## Year2002            -0.2572     0.1364   -1.89  0.05944 .
## Year2003            -0.2312     0.1443   -1.60  0.10923
## Year2004            -0.4335     0.1256   -3.45  0.00056 ***
## Year2005            -0.6057     0.1214   -4.99  6.6e-07 ***
## Year2006            -0.4053     0.1242   -3.26  0.00111 **
## Year2007            -0.5816     0.1229   -4.73  2.4e-06 ***
## Year2008            -0.5495     0.1254   -4.38  1.2e-05 ***
## Year2009            -0.7105     0.1240   -5.73  1.1e-08 ***
## Year2010            -0.8072     0.1181   -6.84  1.0e-11 ***
## Year2011            -0.8448     0.1160   -7.28  4.5e-13 ***
## Year2012            -0.7016     0.1168   -6.00  2.2e-09 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.844
## Multiple R-squared:  0.0816, Adjusted R-squared:  0.0747
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  174 weights are ~= 1. The remaining 2241 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0935  0.8930  0.9430  0.9100  0.9810  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          4.14e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
```

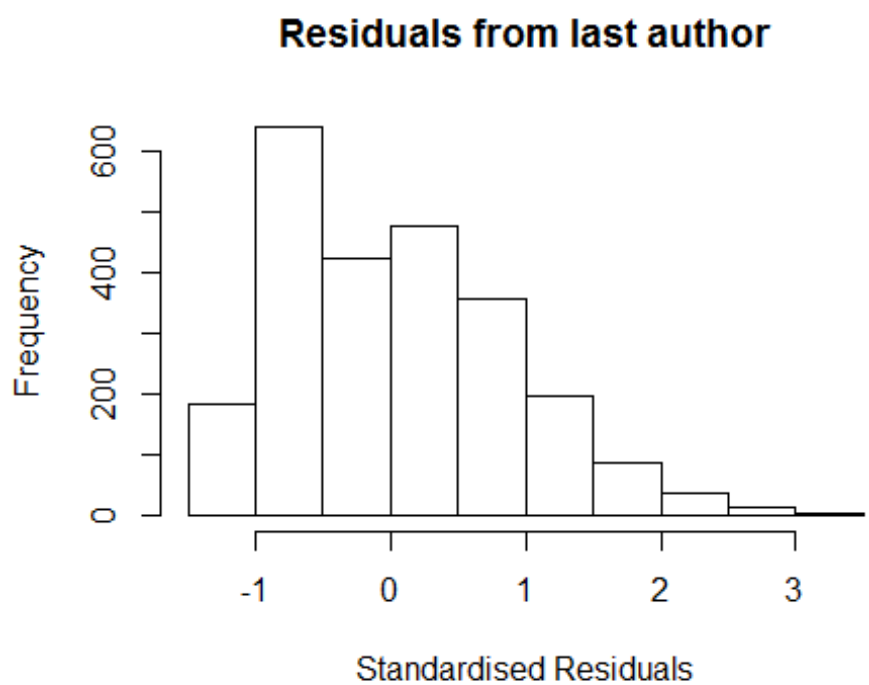
```
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                    cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.065  1           1.032
## Year              1.065 16           1.002
```




```
## [1] "List of  15 outliers with residuals above  2.5"
##         ScopusId  NLCS Year OneField Fields residuals
## 168   0001180094 4.152 1998     1200      2     2.726
## 525   0038702075 3.925 2002     1200      2     2.735
## 575   0036622707 4.107 2002     1200      2     2.885
## 629  19844382943 3.903 2002     1200      2     2.713
## 1168 33748623359 3.480 2005     1200      2     2.638
## 1409 33746888657 3.658 2006     1200      3     2.616
## 1673 33845535808 3.450 2007     1200      2     2.553
## 2328 77951809379 3.936 2010     1200      2     3.296
## 2569 84856757333 3.507 2011     1200      2     2.873
## 2791 79951671066 3.119 2011     1200      3     2.516
## 2996 84868540587 3.821 2012     1200      2     3.075
## 2998 84868571646 3.628 2012     1200      2     2.851
## 3005 84872107429 3.611 2012     1200      2     2.865
```

```
## 3116 84858681871 3.360 2012     1200      3     2.614
## 3150 84861657853 3.650 2012     1200      2     2.904
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.4550 -0.6459 -0.0429  0.6130  3.2907
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.44805    0.10414   13.90  < 2e-16 ***
## FirstAuthorFemale1  0.02349    0.03687    0.64  0.52418
## Year1997           -0.17396    0.14000   -1.24  0.21417
## Year1998           -0.01651    0.13594   -0.12  0.90334
## Year1999           -0.21017    0.15839   -1.33  0.18467
## Year2000           -0.00595    0.14459   -0.04  0.96715
## Year2001           -0.27458    0.13443   -2.04  0.04120 *
## Year2002           -0.25549    0.13637   -1.87  0.06111 .
## Year2003           -0.22879    0.14434   -1.59  0.11306
## Year2004           -0.43005    0.12560   -3.42  0.00063 ***
## Year2005           -0.60244    0.12155   -4.96  7.7e-07 ***
## Year2006           -0.40328    0.12431   -3.24  0.00119 **
## Year2007           -0.58037    0.12305   -4.72  2.5e-06 ***
## Year2008           -0.54779    0.12546   -4.37  1.3e-05 ***
## Year2009           -0.70632    0.12413   -5.69  1.4e-08 ***
## Year2010           -0.80277    0.11794   -6.81  1.3e-11 ***
## Year2011           -0.84118    0.11604   -7.25  5.6e-13 ***
## Year2012           -0.69855    0.11677   -5.98  2.5e-09 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.844
## Multiple R-squared:  0.0813, Adjusted R-squared:  0.0748
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  173 weights are ~= 1. The remaining 2242 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0943  0.8930  0.9440  0.9100  0.9810  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          4.14e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
```

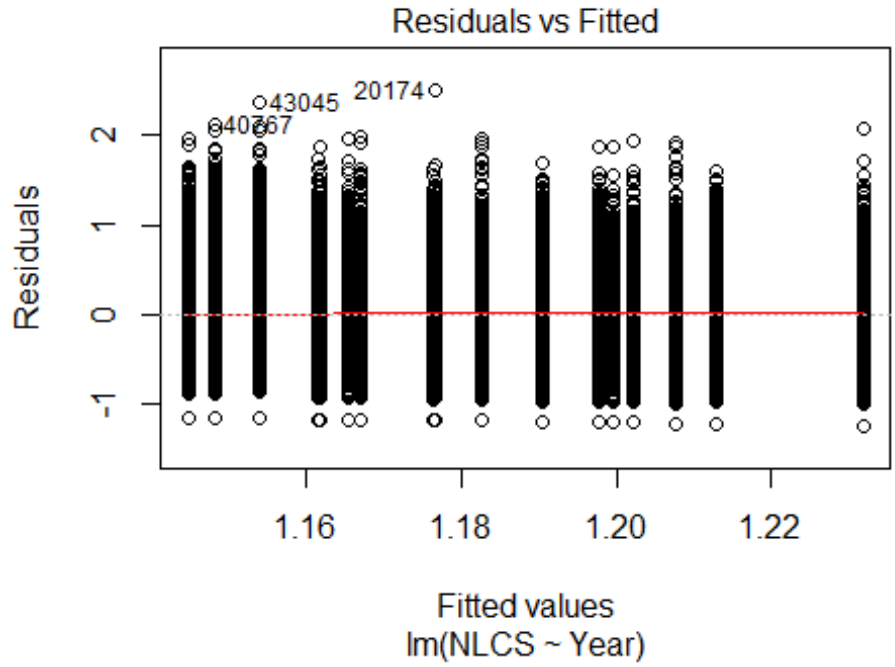
```
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##                   GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.058  1           1.029
## Year             1.058 16           1.002
```



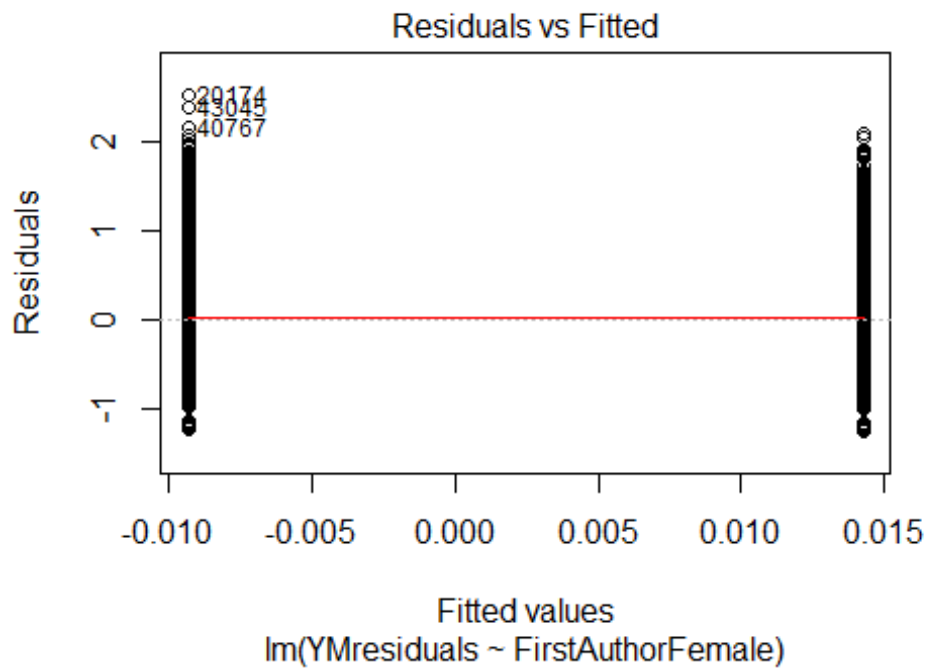

```
## [1] "List of  15 outliers with residuals above  2.5"
##         ScopusId  NLCS Year OneField Fields residuals
## 168   0001180094 4.152 1998     1200      2     2.726
## 525   0038702075 3.925 2002     1200      2     2.735
## 575   0036622707 4.107 2002     1200      2     2.885
## 629  19844382943 3.903 2002     1200      2     2.713
## 1168 33748623359 3.480 2005     1200      2     2.638
## 1409 33746888657 3.658 2006     1200      3     2.616
## 1673 33845535808 3.450 2007     1200      2     2.553
## 2328 77951809379 3.936 2010     1200      2     3.296
## 2569 84856757333 3.507 2011     1200      2     2.873
## 2791 79951671066 3.119 2011     1200      3     2.516
## 2996 84868540587 3.821 2012     1200      2     3.075
```

```
## 2998 84868571646 3.628 2012     1200      2     2.851
## 3005 84872107429 3.611 2012     1200      2     2.865
## 3116 84858681871 3.360 2012     1200      3     2.614
## 3150 84861657853 3.650 2012     1200      2     2.904
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.4624 -0.6632 -0.0413  0.6123  3.2975
##
## Coefficients:
##                   Estimate Std. Error t value Pr(>|t|)
## (Intercept)        1.44513    0.10379   13.92  < 2e-16 ***
## LastAuthorFemale1  0.03664    0.03670    1.00  0.31822
## Year1997          -0.17435    0.13992   -1.25  0.21287
## Year1998          -0.01934    0.13574   -0.14  0.88670
## Year1999          -0.21540    0.15813   -1.36  0.17328
## Year2000          -0.00909    0.14469   -0.06  0.94993
## Year2001          -0.27685    0.13419   -2.06  0.03920 *
## Year2002          -0.25718    0.13617   -1.89  0.05906 .
## Year2003          -0.23068    0.14420   -1.60  0.10980
## Year2004          -0.43193    0.12538   -3.45  0.00058 ***
## Year2005          -0.60465    0.12136   -4.98  6.7e-07 ***
## Year2006          -0.40588    0.12405   -3.27  0.00108 **
## Year2007          -0.58137    0.12286   -4.73  2.4e-06 ***
## Year2008          -0.54977    0.12527   -4.39  1.2e-05 ***
## Year2009          -0.70869    0.12390   -5.72  1.2e-08 ***
## Year2010          -0.80663    0.11795   -6.84  1.0e-11 ***
## Year2011          -0.84416    0.11591   -7.28  4.4e-13 ***
## Year2012          -0.70190    0.11670   -6.01  2.1e-09 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.844
## Multiple R-squared:  0.0815, Adjusted R-squared:  0.075
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  173 weights are ~= 1. The remaining 2242 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0928  0.8920  0.9430  0.9100  0.9810  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          4.14e-05          1.82e-12
```

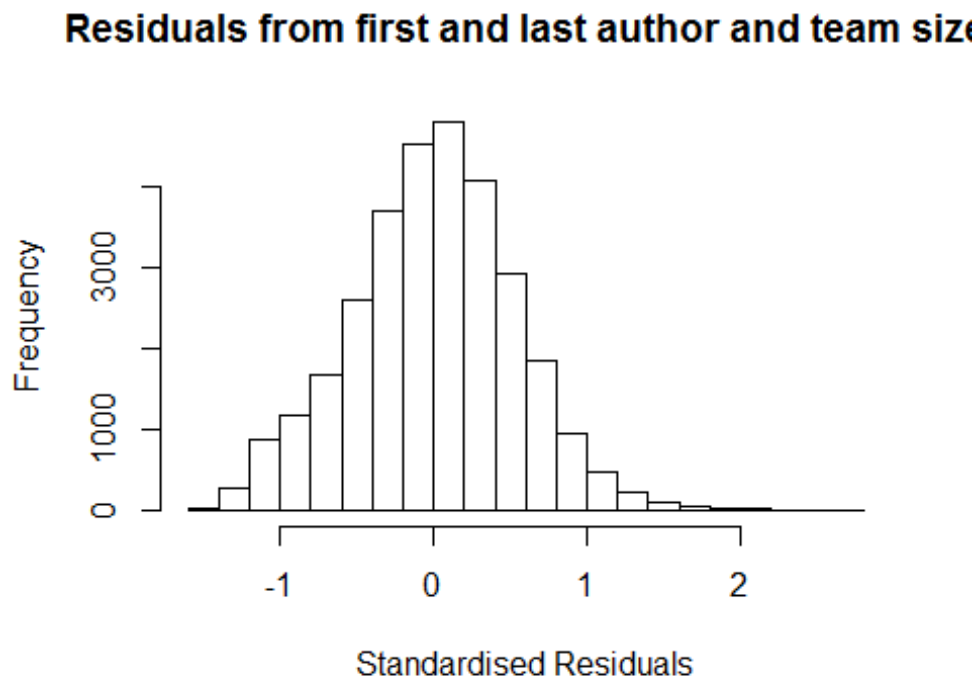
```
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Sample size for the above analysis:  2415"
## [1] ""
## [1] ""
## [1] "####################################"
## [1] "Analysis of AJSC 1201"
## [1] "####################################"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 2222 2061 1829 1718 2054 2222 2132 1825 1968 1980 2457 2667 2640 2872 2821
## 2011 2012
## 3244 3393
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 1832 1670 1490 1387 1535 1444 1795 1565 1664 1628 2083 2246 2237 2382 2363
## 2011 2012
## 2688 2790
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 1714 1550 1390 1279 1421 1356 1653 1456 1525 1468 1904 2042 2063 2212 2194
## 2011 2012
## 2475 2550
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
##  Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 250, df = 16, p-value <2e-16
```




```
##
##  Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 11, df = 1, p-value = 7e-04
```

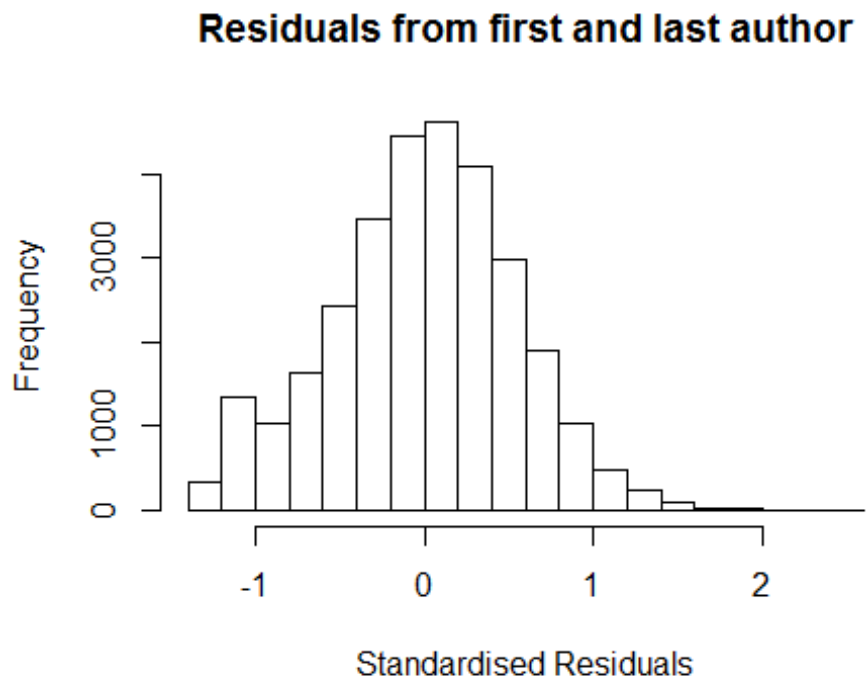



```
## [1] "Female first author team size 2018 geometric mean: 2.22378291965835"
## [1] "Male first author team size 2018 geometric mean: 2.04381649663967"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 640000, p-value = 0.004
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 2.08118885308021"
## [1] "Male last author team size 2018 geometric mean: 2.17540904570402"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 570000, p-value = 0.1
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.171  1           1.082
## LastAuthorFemale  1.172  1           1.083
## UniqueAuthors     1.023  4           1.003
## Year              1.029 16           1.001
```



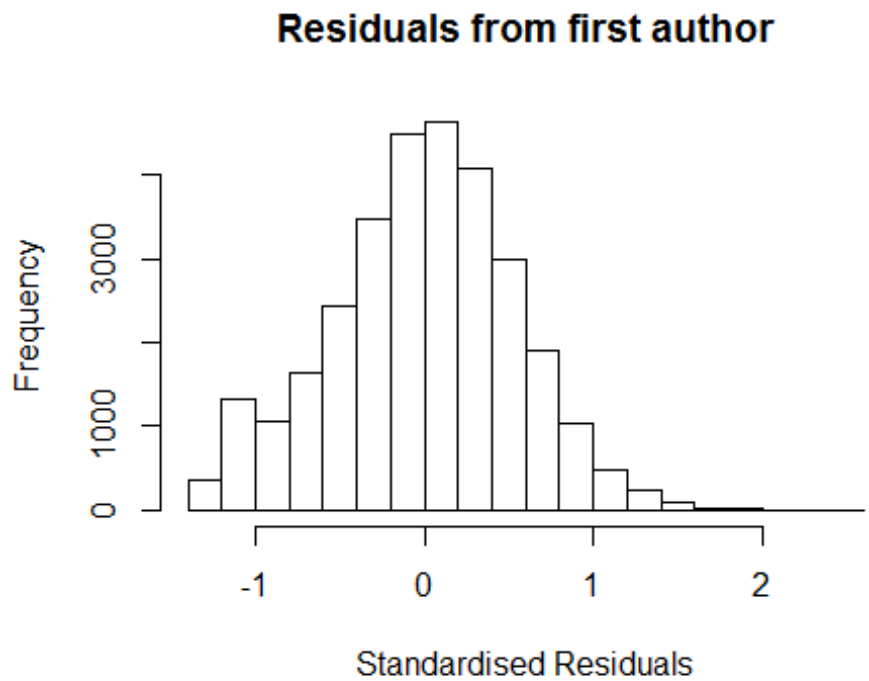

```
## [1] "List of  2 outliers with residuals above  2.5"
##          ScopusId  NLCS Year OneField Fields residuals
## 20174 28244490466 3.676 2004     1201      4     2.672
## 43045 84861899557 3.516 2012     1201      3     2.572
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
UniqueAuthors +
##     Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##     k.max = 1000))
##  \--> method = "MM"
## Residuals:
##    Min     1Q Median     3Q    Max
## -1.532 -0.342  0.013  0.346  2.672
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.05e+00   1.61e-02   65.38  < 2e-16 ***
## FirstAuthorFemale1  2.27e-02   6.81e-03    3.34  0.00083 ***
## LastAuthorFemale1   8.21e-03   6.98e-03    1.18  0.23975
## UniqueAuthors2      2.20e-01   8.18e-03   26.94  < 2e-16 ***
## UniqueAuthors3      2.55e-01   8.98e-03   28.36  < 2e-16 ***
## UniqueAuthors4      2.85e-01   1.10e-02   25.86  < 2e-16 ***
## UniqueAuthors5      4.47e-01   1.00e-02   44.71  < 2e-16 ***
## Year1997           -1.98e-02   2.15e-02   -0.92  0.35724
## Year1998           -1.72e-05   2.08e-02    0.00  0.99934
```

```
## Year1999           -1.04e-02   2.07e-02   -0.50  0.61476
## Year2000           -6.14e-02   2.08e-02   -2.95  0.00318 **
## Year2001           -1.13e-02   2.08e-02   -0.54  0.58840
## Year2002           -3.93e-02   2.00e-02   -1.96  0.04978 *
## Year2003           -6.19e-02   2.06e-02   -3.01  0.00266 **
## Year2004           -5.02e-02   2.00e-02   -2.51  0.01201 *
## Year2005           -3.57e-02   2.01e-02   -1.78  0.07534 .
## Year2006           -6.13e-02   1.88e-02   -3.27  0.00109 **
## Year2007           -9.26e-02   1.89e-02   -4.89  1.0e-06 ***
## Year2008           -7.84e-02   1.91e-02   -4.11  3.9e-05 ***
## Year2009           -8.29e-02   1.89e-02   -4.39  1.2e-05 ***
## Year2010           -1.05e-01   1.93e-02   -5.43  5.5e-08 ***
## Year2011           -1.09e-01   1.91e-02   -5.72  1.1e-08 ***
## Year2012           -1.10e-01   1.96e-02   -5.64  1.7e-08 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.511
## Multiple R-squared:  0.0808, Adjusted R-squared:  0.0802
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
##  2 observations c(13139,28911) are outliers with |weight| = 0 ( < 3.3e-
06);
##  2527 weights are ~= 1. The remaining 27723 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0095  0.8650  0.9500  0.9010  0.9860  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          3.31e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.177  1           1.085
## LastAuthorFemale  1.179  1           1.086
## Year              1.009 16           1.000
```



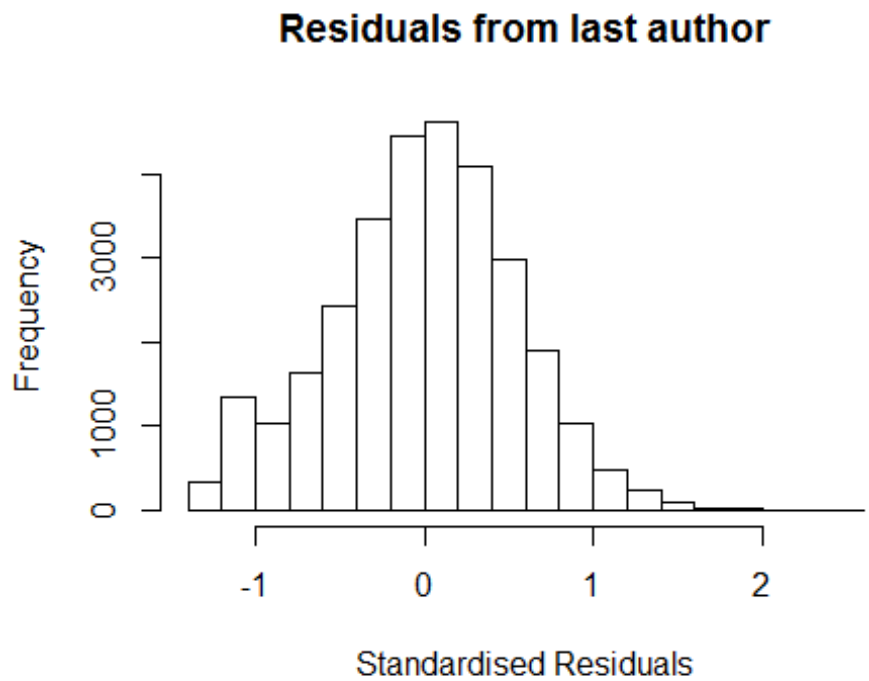

```
## [1] "List of  1 outliers with residuals above  2.5"
##          ScopusId  NLCS Year OneField Fields residuals
## 20174 28244490466 3.676 2004     1201      4     2.502
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.2608 -0.3497  0.0169  0.3581  2.5019
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)          1.21892    0.01596   76.39  < 2e-16 ***
## FirstAuthorFemale1  0.03403    0.00711    4.79  1.7e-06 ***
## LastAuthorFemale1  -0.01162    0.00728   -1.60  0.11026
## Year1997           -0.02141    0.02235   -0.96  0.33800
## Year1998            0.00783    0.02172    0.36  0.71846
## Year1999           -0.02077    0.02145   -0.97  0.33285
## Year2000           -0.06498    0.02158   -3.01  0.00260 **
## Year2001           -0.01248    0.02149   -0.58  0.56147
## Year2002           -0.03560    0.02070   -1.72  0.08550 .
## Year2003           -0.05212    0.02133   -2.44  0.01454 *
## Year2004           -0.04482    0.02055   -2.18  0.02921 *
## Year2005           -0.01942    0.02073   -0.94  0.34892
```

```
## Year2006           -0.05377    0.01930   -2.79  0.00533 **
## Year2007           -0.07087    0.01952   -3.63  0.00028 ***
## Year2008           -0.06207    0.01969   -3.15  0.00162 **
## Year2009           -0.06897    0.01951   -3.53  0.00041 ***
## Year2010           -0.08762    0.01999   -4.38  1.2e-05 ***
## Year2011           -0.08397    0.01967   -4.27  2.0e-05 ***
## Year2012           -0.08227    0.02024   -4.06  4.8e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.528
## Multiple R-squared:  0.00343,    Adjusted R-squared:  0.00284
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
##  observation 13139 is an outlier with |weight| = 0 ( < 3.3e-06);
##  2581 weights are ~= 1. The remaining 27670 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0054  0.8620  0.9500  0.9000  0.9860  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          3.31e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.005  1           1.003
## Year              1.005 16           1.000
```




```
## [1] "List of  1 outliers with residuals above  2.5"
##          ScopusId  NLCS Year OneField Fields residuals
## 20174 28244490466 3.676 2004     1201      4     2.502
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.2537 -0.3506  0.0173  0.3579  2.5040
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)          1.21679    0.01586   76.72  < 2e-16 ***
## FirstAuthorFemale1   0.02865    0.00660    4.34  1.4e-05 ***
## Year1997            -0.02128    0.02234   -0.95  0.34086
## Year1998             0.00822    0.02170    0.38  0.70498
## Year1999            -0.02067    0.02145   -0.96  0.33511
## Year2000            -0.06461    0.02158   -2.99  0.00275 **
## Year2001            -0.01275    0.02148   -0.59  0.55286
## Year2002            -0.03565    0.02070   -1.72  0.08502 .
## Year2003            -0.05213    0.02132   -2.45  0.01449 *
## Year2004            -0.04478    0.02055   -2.18  0.02932 *
## Year2005            -0.01965    0.02073   -0.95  0.34309
## Year2006            -0.05405    0.01929   -2.80  0.00509 **
```

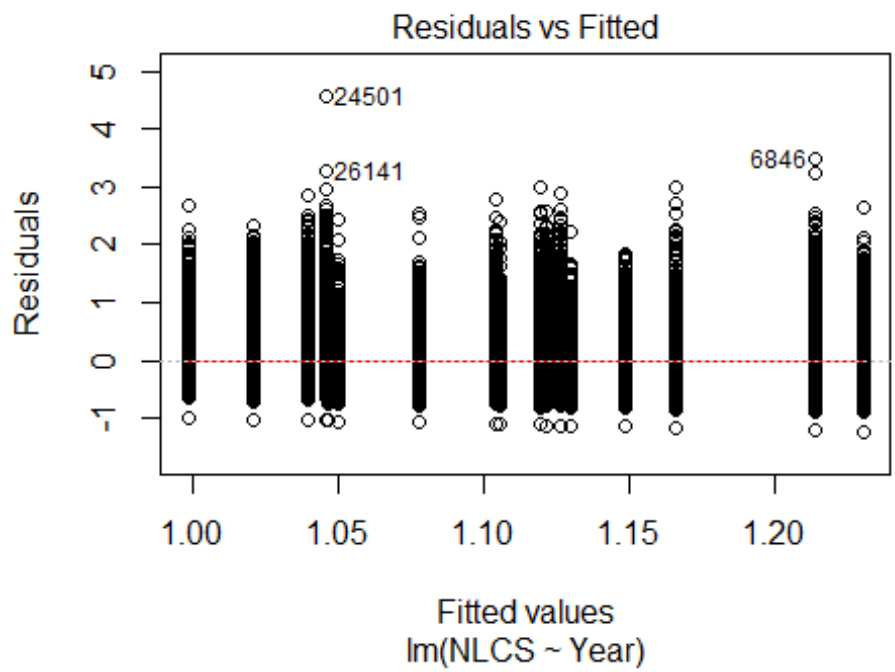
```
## Year2007           -0.07089    0.01951   -3.63  0.00028 ***
## Year2008           -0.06204    0.01968   -3.15  0.00163 **
## Year2009           -0.06880    0.01951   -3.53  0.00042 ***
## Year2010           -0.08775    0.01998   -4.39  1.1e-05 ***
## Year2011           -0.08422    0.01967   -4.28  1.9e-05 ***
## Year2012           -0.08214    0.02024   -4.06  4.9e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.528
## Multiple R-squared:  0.00336,    Adjusted R-squared:  0.0028
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
##  observation 13139 is an outlier with |weight| = 0 ( < 3.3e-06);
##  2593 weights are ~= 1. The remaining 27658 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0052  0.8620  0.9500  0.8990  0.9860  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          3.31e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##                   GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.006  1           1.003
## Year             1.006 16           1.000
```



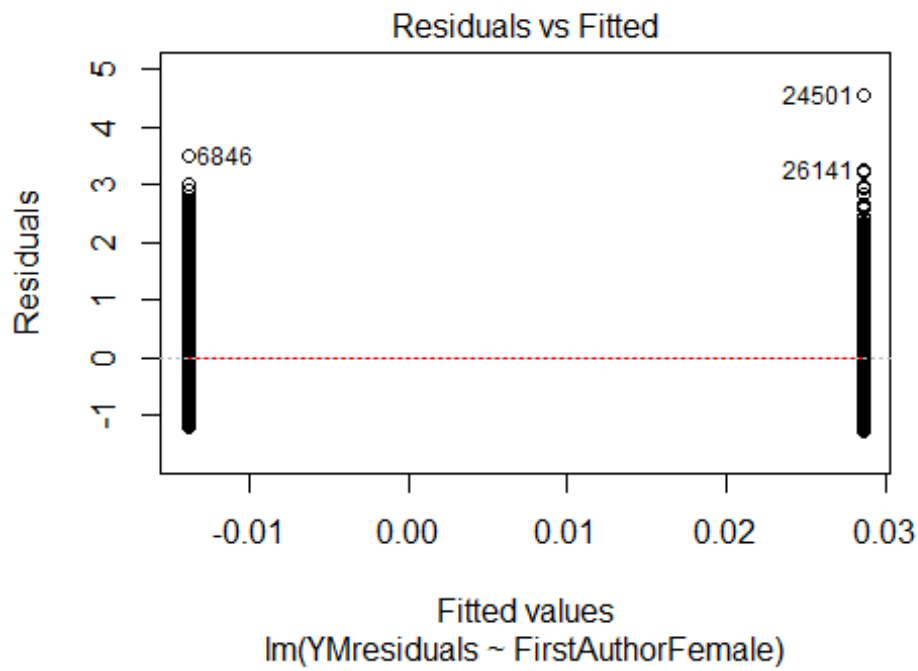

```
## [1] "List of  1 outliers with residuals above  2.5"
##          ScopusId  NLCS Year OneField Fields residuals
## 20174 28244490466 3.676 2004     1201      4     2.502
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.2385 -0.3522  0.0183  0.3582  2.4944
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.22520    0.01588   77.16  < 2e-16 ***
## LastAuthorFemale1   0.00463    0.00674    0.69  0.49250
## Year1997           -0.02131    0.02237   -0.95  0.34094
## Year1998            0.00870    0.02173    0.40  0.68875
## Year1999           -0.02010    0.02148   -0.94  0.34943
## Year2000           -0.06383    0.02161   -2.95  0.00314 **
## Year2001           -0.01112    0.02150   -0.52  0.60506
## Year2002           -0.03498    0.02072   -1.69  0.09145 .
## Year2003           -0.05189    0.02133   -2.43  0.01498 *
## Year2004           -0.04364    0.02058   -2.12  0.03395 *
## Year2005           -0.01844    0.02076   -0.89  0.37440
## Year2006           -0.05252    0.01931   -2.72  0.00654 **
```

```
## Year2007          -0.06958    0.01955   -3.56  0.00037 ***
## Year2008          -0.06036    0.01971   -3.06  0.00220 **
## Year2009          -0.06644    0.01952   -3.40  0.00067 ***
## Year2010          -0.08521    0.02000   -4.26  2.0e-05 ***
## Year2011          -0.08217    0.01969   -4.17  3.0e-05 ***
## Year2012          -0.07955    0.02027   -3.92  8.7e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.528
## Multiple R-squared:  0.00273,    Adjusted R-squared:  0.00217
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
##  observation 13139 is an outlier with |weight| = 0 ( < 3.3e-06);
##  2576 weights are ~= 1. The remaining 27675 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0067  0.8620  0.9500  0.9000  0.9860  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          3.31e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Sample size for the above analysis:  30252"
## [1] ""
## [1] ""
## [1] "####################################"
## [1] "Analysis of AJSC 1202"
## [1] "####################################"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  860  771  827  941 1089 1123 1365 1434 1561 1541 1609 1911 1992 1789 2047
## 2011 2012
## 2352 2368
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  737  630  700  792  890  895 1153 1240 1356 1337 1419 1655 1725 1544 1777
```

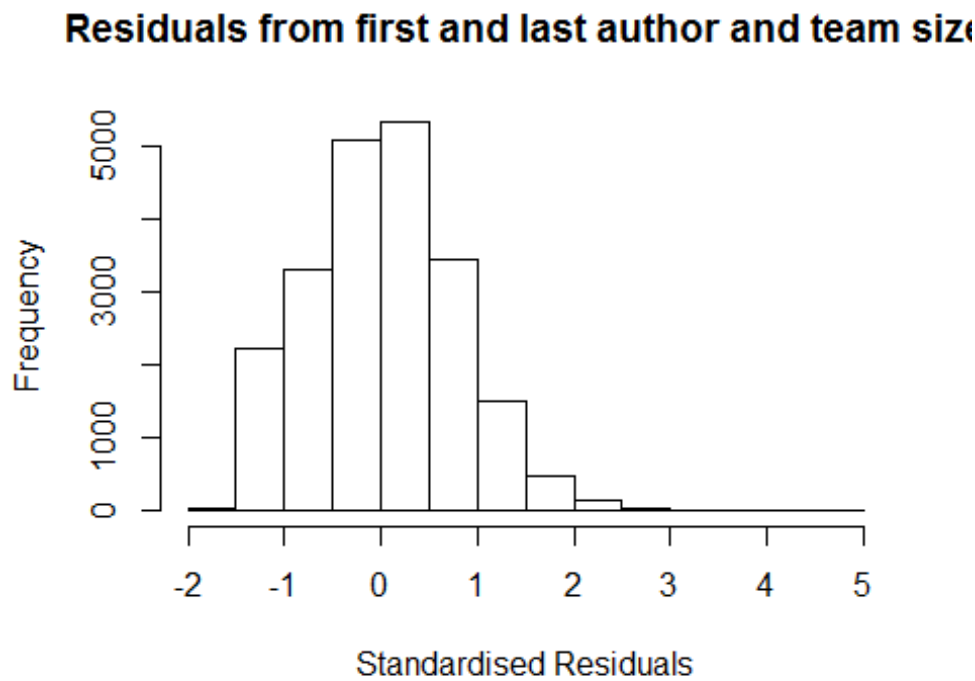
```
## 2011 2012
## 2027 2033
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  723  627  692  781  874  880 1137 1225 1350 1324 1400 1625 1688 1515 1750
## 2011 2012
## 1994 2005
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
##  Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 310, df = 16, p-value <2e-16
```




```
##
##  Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 1.6, df = 1, p-value = 0.2
```



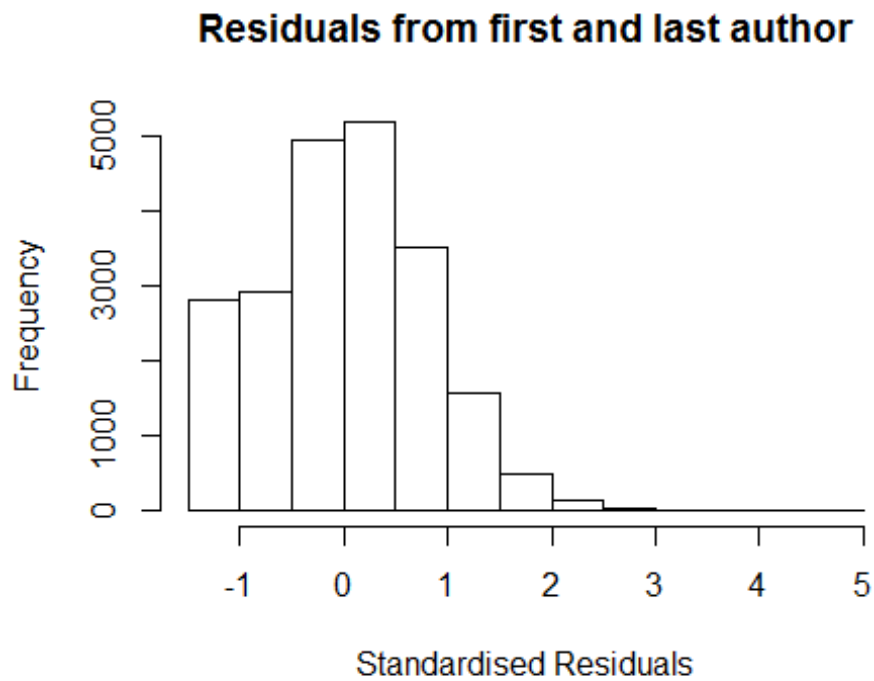

```
## [1] "Female first author team size 2018 geometric mean: 1.13215723113873"
## [1] "Male first author team size 2018 geometric mean: 1.10019995796152"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 5e+05, p-value = 0.1
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 1.1217644609329"
## [1] "Male last author team size 2018 geometric mean: 1.10710937594789"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 490000, p-value = 0.9
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 5.399  1           2.324
## LastAuthorFemale  5.402  1           2.324
## UniqueAuthors     1.020  4           1.002
## Year              1.024 16           1.001
```




```
## [1] "List of  34 outliers with residuals above  2.5"
##          ScopusId  NLCS Year OneField Fields residuals
## 4069   0034345564 3.547 2000     1202      2     2.540
## 5178   0035101649 3.482 2001     1202      3     2.506
## 6846  61149169393 4.690 2002     1202      2     3.566
## 6974   0036596824 4.474 2002     1202      3     3.304
## 7211  85012444559 3.746 2002     1202      3     2.622
## 7261   1442335800 3.694 2002     1202      6     2.570
## 8085   0038067828 3.866 2003     1202      1     2.707
## 8430   0042885451 3.882 2003     1202      3     2.677
## 9496   4043158357 4.118 2004     1202      1     3.083
## 10051 28244490466 3.676 2004     1201      4     2.641
## 10072 33749860006 3.693 2004     1202      3     2.658
## 10477 33746417589 3.577 2005     1202      2     2.538
## 10637 33845250551 4.010 2005     1202      4     2.971
## 10693 60950110696 3.592 2005     1202      3     2.507
## 10773 62749191854 3.744 2005     1202      3     2.659
## 12886 33645730942 3.886 2006     1202      1     2.801
## 13187 41449090508 4.162 2006     1202      2     3.031
## 15126 43249133166 3.704 2007     1202      2     2.676
## 16120 65749261392 3.585 2008     1202      3     2.536
## 16767 46849103881 3.875 2008     1202      3     2.872
## 20813 79956372974 3.668 2010     1202      4     2.762
## 22299 80155194058 3.906 2011     1202      5     2.924
## 24442 84868280635 3.541 2012     1202      3     2.628
## 24501 84871394290 5.615 2012     1202      3     4.656
## 24608 84874341423 3.664 2012     1202      3     2.751
```

```
## 24912 84867202611 3.525 2012     1202      2     2.612
## 25152 84865713540 3.451 2012     1202      3     2.538
## 25321 84864019885 3.507 2012     1202      3     2.548
## 26055 84860734606 3.999 2012     1202      3     3.040
## 26141 84856468758 4.331 2012     1202      3     3.372
## 26219 84857525840 3.491 2012     1202      3     2.532
## 26222 84857597635 3.715 2012     1202      3     2.756
## 26290 81855220212 3.475 2012     1202      2     2.562
## 26293 84855242044 3.474 2012     1202      2     2.561
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
UniqueAuthors +
##     Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##     k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.6860 -0.5162  0.0129  0.5216  4.6557
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)          1.08293    0.02663   40.67  < 2e-16 ***
## FirstAuthorFemale1  0.05980    0.02631    2.27   0.0230 *
## LastAuthorFemale1  -0.01365    0.02635   -0.52   0.6044
## UniqueAuthors2      0.37084    0.01939   19.12  < 2e-16 ***
## UniqueAuthors3      0.60993    0.03174   19.22  < 2e-16 ***
## UniqueAuthors4      0.65218    0.05312   12.28  < 2e-16 ***
## UniqueAuthors5      0.62146    0.06758    9.20  < 2e-16 ***
## Year1997           -0.01842    0.03856   -0.48   0.6329
## Year1998           -0.03504    0.03726   -0.94   0.3470
## Year1999           -0.10855    0.03518   -3.09   0.0020 **
## Year2000           -0.07639    0.03480   -2.20   0.0282 *
## Year2001           -0.10697    0.03472   -3.08   0.0021 **
## Year2002            0.04106    0.03626    1.13   0.2574
## Year2003            0.07562    0.03514    2.15   0.0314 *
## Year2004           -0.04777    0.03559   -1.34   0.1796
## Year2005           -0.04360    0.03539   -1.23   0.2180
## Year2006            0.00236    0.03372    0.07   0.9442
## Year2007           -0.05494    0.03275   -1.68   0.0934 .
## Year2008           -0.08008    0.03213   -2.49   0.0127 *
## Year2009           -0.15478    0.03269   -4.73  2.2e-06 ***
## Year2010           -0.17722    0.03144   -5.64  1.8e-08 ***
## Year2011           -0.14692    0.03223   -4.56  5.2e-06 ***
## Year2012           -0.16976    0.03319   -5.11  3.2e-07 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.787
```

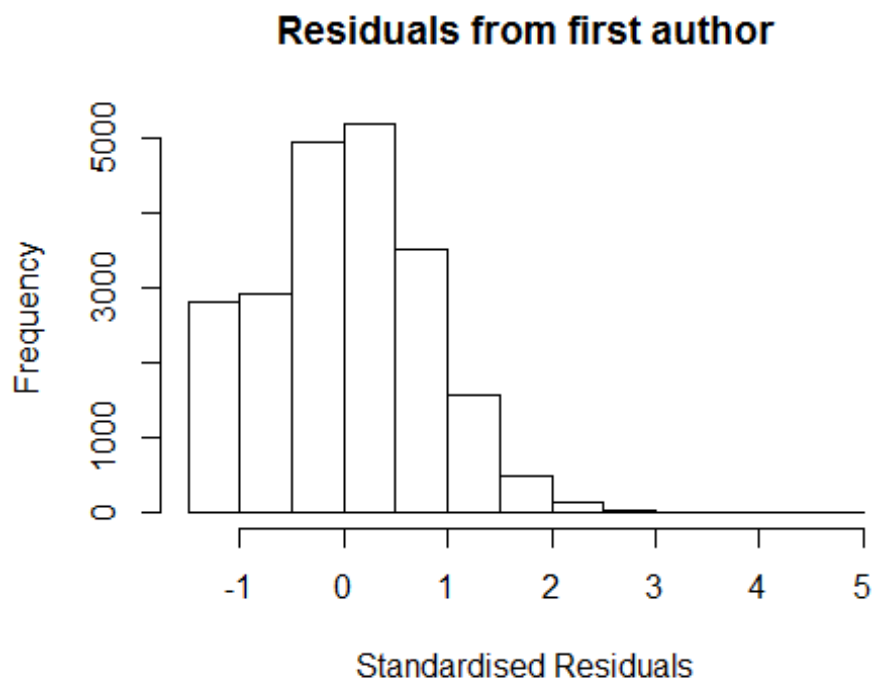
```
## Multiple R-squared:  0.0436, Adjusted R-squared:  0.0426
## Convergence in 10 IRWLS iterations
##
## Robustness weights:
##  observation 19867 is an outlier with |weight| = 0 ( < 4.6e-06);
##  1815 weights are ~= 1. The remaining 19774 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0041  0.8670  0.9520  0.9160  0.9860  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          4.63e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 6.800  1           2.608
## LastAuthorFemale  6.794  1           2.607
## Year              1.009 16           1.000
```




```
## [1] "List of  28 outliers with residuals above  2.5"
##          ScopusId  NLCS Year OneField Fields residuals
## 4067   0034344702 3.626 2000     1202      2     2.574
## 6846  61149169393 4.690 2002     1202      2     3.520
## 6974   0036596824 4.474 2002     1202      3     3.261
## 7211  85012444559 3.746 2002     1202      3     2.576
## 7261   1442335800 3.694 2002     1202      6     2.524
## 8085   0038067828 3.866 2003     1202      1     2.664
## 8430   0042885451 3.882 2003     1202      3     2.637
## 9496   4043158357 4.118 2004     1202      1     3.038
## 10051 28244490466 3.676 2004     1201      4     2.596
## 10072 33749860006 3.693 2004     1202      3     2.613
## 10637 33845250551 4.010 2005     1202      4     2.929
## 10773 62749191854 3.744 2005     1202      3     2.620
## 12886 33645730942 3.886 2006     1202      1     2.748
## 13186 40849109927 3.694 2006     1202      2     2.556
## 13187 41449090508 4.162 2006     1202      2     2.981
## 15126 43249133166 3.704 2007     1202      2     2.620
## 16767 46849103881 3.875 2008     1202      3     2.808
## 20813 79956372974 3.668 2010     1202      4     2.707
## 22299 80155194058 3.906 2011     1202      5     2.878
## 24442 84868280635 3.541 2012     1202      3     2.566
## 24501 84871394290 5.615 2012     1202      3     4.597
## 24608 84874341423 3.664 2012     1202      3     2.689
## 24912 84867202611 3.525 2012     1202      2     2.550
## 25219 84873204638 3.715 2012     1202      3     2.686
## 26055 84860734606 3.999 2012     1202      3     2.981
```

```
## 26141 84856468758 4.331 2012     1202      3     3.313
## 26222 84857597635 3.715 2012     1202      3     2.697
## 26290 81855220212 3.475 2012     1202      2     2.500
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.2446 -0.5364  0.0113  0.5354  4.5974
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.12948    0.02706   41.74  < 2e-16 ***
## FirstAuthorFemale1  0.05468    0.03014    1.81   0.0696 .
## LastAuthorFemale1  -0.01183    0.03016   -0.39   0.6948
## Year1997           -0.01579    0.03885   -0.41   0.6844
## Year1998           -0.04368    0.03770   -1.16   0.2467
## Year1999           -0.10514    0.03596   -2.92   0.0035 **
## Year2000           -0.07715    0.03529   -2.19   0.0288 *
## Year2001           -0.10282    0.03544   -2.90   0.0037 **
## Year2002            0.04053    0.03708    1.09   0.2744
## Year2003            0.07229    0.03568    2.03   0.0428 *
## Year2004           -0.04990    0.03609   -1.38   0.1667
## Year2005           -0.04880    0.03602   -1.35   0.1755
## Year2006            0.00875    0.03435    0.25   0.7988
## Year2007           -0.04591    0.03343   -1.37   0.1697
## Year2008           -0.06258    0.03289   -1.90   0.0571 .
## Year2009           -0.14609    0.03346   -4.37  1.3e-05 ***
## Year2010           -0.16886    0.03217   -5.25  1.5e-07 ***
## Year2011           -0.14425    0.03292   -4.38  1.2e-05 ***
## Year2012           -0.15477    0.03403   -4.55  5.4e-06 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.809
## Multiple R-squared:  0.00879,    Adjusted R-squared:  0.00796
## Convergence in 10 IRWLS iterations
##
## Robustness weights:
##  observation 19867 is an outlier with |weight| = 0 ( < 4.6e-06);
##  1835 weights are ~= 1. The remaining 19754 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0187  0.8650  0.9530  0.9170  0.9860  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
```

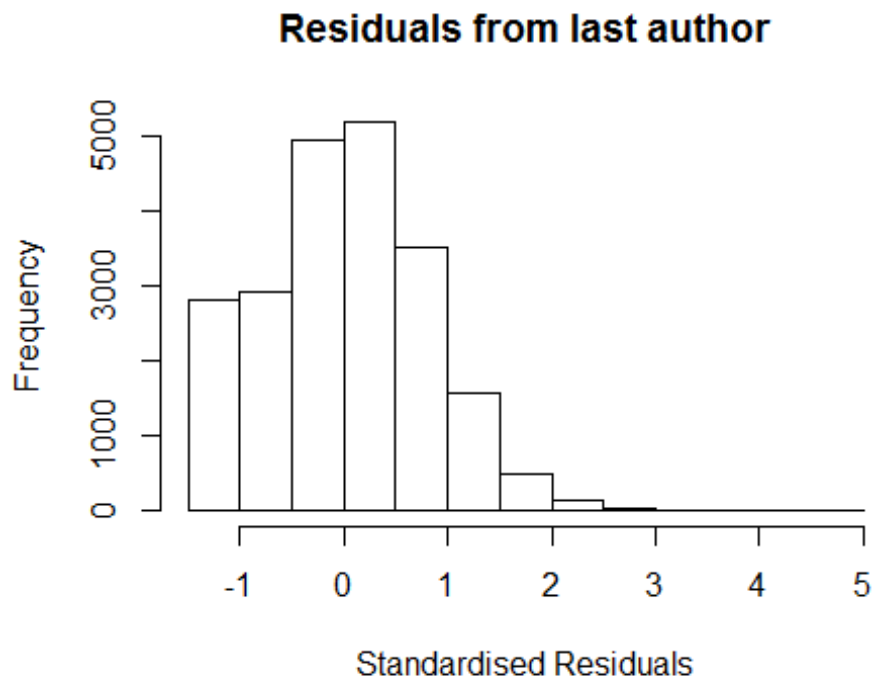
```
##          1.00e-07          1.00e-07          4.63e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.006  1           1.003
## Year              1.006 16           1.000
```




```
## [1] "List of  28 outliers with residuals above  2.5"
##          ScopusId  NLCS Year OneField Fields residuals
## 4067   0034344702 3.626 2000     1202      2     2.574
## 6846  61149169393 4.690 2002     1202      2     3.520
## 6974   0036596824 4.474 2002     1202      3     3.261
## 7211  85012444559 3.746 2002     1202      3     2.576
## 7261   1442335800 3.694 2002     1202      6     2.524
## 8085   0038067828 3.866 2003     1202      1     2.664
## 8430   0042885451 3.882 2003     1202      3     2.637
## 9496   4043158357 4.118 2004     1202      1     3.038
```

```
## 10051 28244490466 3.676 2004     1201      4     2.596
## 10072 33749860006 3.693 2004     1202      3     2.613
## 10637 33845250551 4.010 2005     1202      4     2.929
## 10773 62749191854 3.744 2005     1202      3     2.620
## 12886 33645730942 3.886 2006     1202      1     2.748
## 13186 40849109927 3.694 2006     1202      2     2.556
## 13187 41449090508 4.162 2006     1202      2     2.981
## 15126 43249133166 3.704 2007     1202      2     2.620
## 16767 46849103881 3.875 2008     1202      3     2.808
## 20813 79956372974 3.668 2010     1202      4     2.707
## 22299 80155194058 3.906 2011     1202      5     2.878
## 24442 84868280635 3.541 2012     1202      3     2.566
## 24501 84871394290 5.615 2012     1202      3     4.597
## 24608 84874341423 3.664 2012     1202      3     2.689
## 24912 84867202611 3.525 2012     1202      2     2.550
## 25219 84873204638 3.715 2012     1202      3     2.686
## 26055 84860734606 3.999 2012     1202      3     2.981
## 26141 84856468758 4.331 2012     1202      3     3.313
## 26222 84857597635 3.715 2012     1202      3     2.697
## 26290 81855220212 3.475 2012     1202      2     2.500
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.2453 -0.5359  0.0113  0.5357  4.5967
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.12912    0.02705   41.75  < 2e-16 ***
## FirstAuthorFemale1  0.04388    0.01160    3.78  0.00015 ***
## Year1997           -0.01570    0.03884   -0.40  0.68618
## Year1998           -0.04362    0.03770   -1.16  0.24718
## Year1999           -0.10510    0.03596   -2.92  0.00347 **
## Year2000           -0.07699    0.03528   -2.18  0.02911 *
## Year2001           -0.10286    0.03543   -2.90  0.00370 **
## Year2002            0.04060    0.03707    1.10  0.27349
## Year2003            0.07235    0.03568    2.03  0.04263 *
## Year2004           -0.04984    0.03609   -1.38  0.16724
## Year2005           -0.04874    0.03602   -1.35  0.17597
## Year2006            0.00883    0.03435    0.26  0.79717
## Year2007           -0.04589    0.03343   -1.37  0.16986
## Year2008           -0.06243    0.03288   -1.90  0.05765 .
## Year2009           -0.14610    0.03346   -4.37  1.3e-05 ***
## Year2010           -0.16892    0.03216   -5.25  1.5e-07 ***
## Year2011           -0.14420    0.03292   -4.38  1.2e-05 ***
## Year2012           -0.15466    0.03402   -4.55  5.5e-06 ***
```

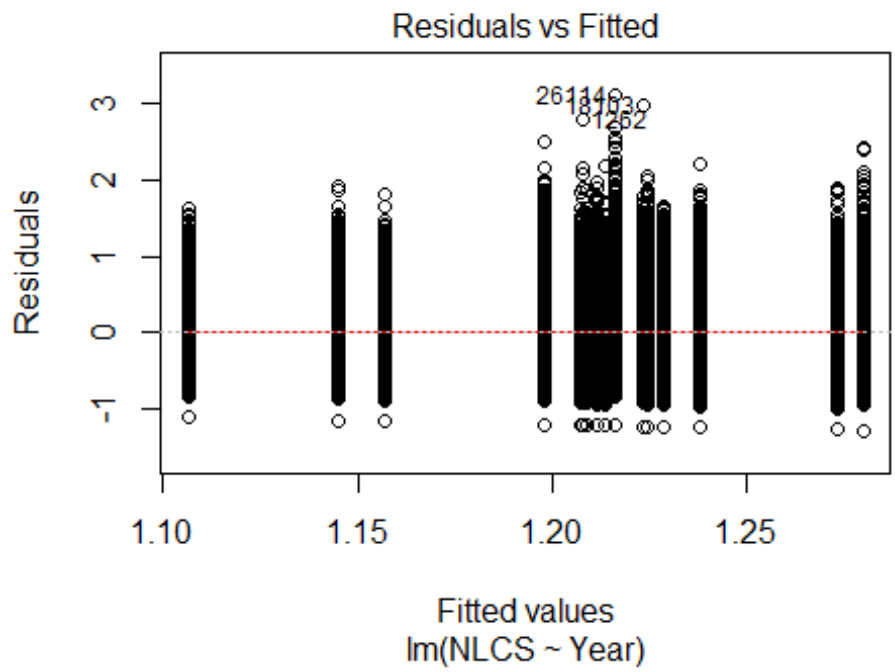
```
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.809
## Multiple R-squared:  0.00877,    Adjusted R-squared:  0.00799
## Convergence in 10 IRWLS iterations
##
## Robustness weights:
##  observation 19867 is an outlier with |weight| = 0 ( < 4.6e-06);
##  1836 weights are ~= 1. The remaining 19753 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0189  0.8650  0.9530  0.9170  0.9860  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          4.63e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                    psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##                   GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.005  1           1.003
## Year             1.005 16           1.000
```



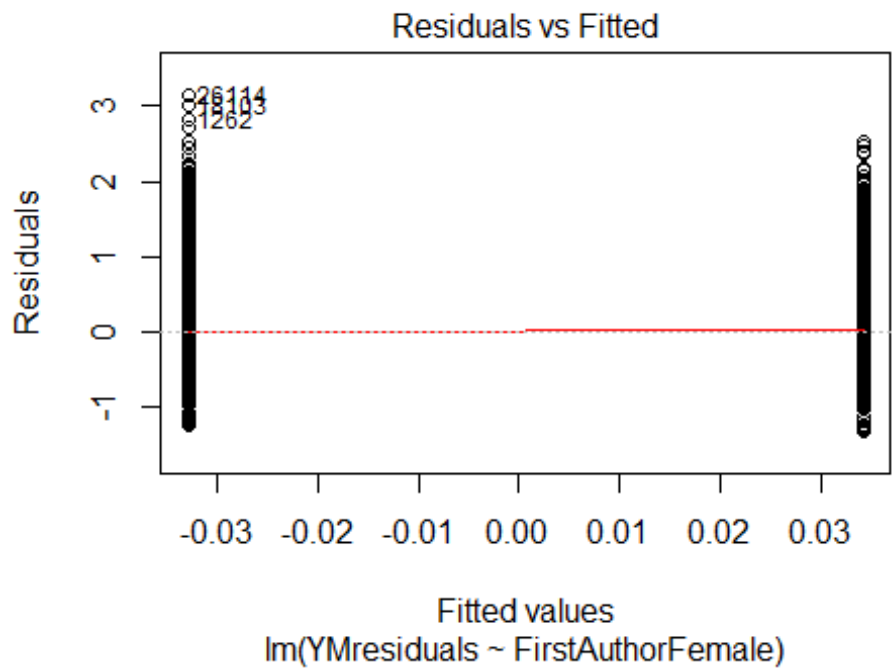

```
## [1] "List of  28 outliers with residuals above  2.5"
##          ScopusId  NLCS Year OneField Fields residuals
## 4067   0034344702 3.626 2000     1202      2     2.574
## 6846  61149169393 4.690 2002     1202      2     3.520
## 6974   0036596824 4.474 2002     1202      3     3.261
## 7211  85012444559 3.746 2002     1202      3     2.576
## 7261   1442335800 3.694 2002     1202      6     2.524
## 8085   0038067828 3.866 2003     1202      1     2.664
## 8430   0042885451 3.882 2003     1202      3     2.637
## 9496   4043158357 4.118 2004     1202      1     3.038
## 10051 28244490466 3.676 2004     1201      4     2.596
## 10072 33749860006 3.693 2004     1202      3     2.613
## 10637 33845250551 4.010 2005     1202      4     2.929
## 10773 62749191854 3.744 2005     1202      3     2.620
## 12886 33645730942 3.886 2006     1202      1     2.748
## 13186 40849109927 3.694 2006     1202      2     2.556
## 13187 41449090508 4.162 2006     1202      2     2.981
## 15126 43249133166 3.704 2007     1202      2     2.620
## 16767 46849103881 3.875 2008     1202      3     2.808
## 20813 79956372974 3.668 2010     1202      4     2.707
## 22299 80155194058 3.906 2011     1202      5     2.878
## 24442 84868280635 3.541 2012     1202      3     2.566
## 24501 84871394290 5.615 2012     1202      3     4.597
## 24608 84874341423 3.664 2012     1202      3     2.689
## 24912 84867202611 3.525 2012     1202      2     2.550
## 25219 84873204638 3.715 2012     1202      3     2.686
## 26055 84860734606 3.999 2012     1202      3     2.981
```

```
## 26141 84856468758 4.331 2012     1202      3     3.313
## 26222 84857597635 3.715 2012     1202      3     2.697
## 26290 81855220212 3.475 2012     1202      2     2.500
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.2413 -0.5360  0.0112  0.5367  4.6001
##
## Coefficients:
##                   Estimate Std. Error t value Pr(>|t|)
## (Intercept)        1.13017    0.02706   41.76  < 2e-16 ***
## LastAuthorFemale1  0.03804    0.01159    3.28   0.0010 **
## Year1997          -0.01503    0.03885   -0.39   0.6989
## Year1998          -0.04320    0.03769   -1.15   0.2516
## Year1999          -0.10463    0.03597   -2.91   0.0036 **
## Year2000          -0.07609    0.03528   -2.16   0.0310 *
## Year2001          -0.10278    0.03543   -2.90   0.0037 **
## Year2002           0.04127    0.03708    1.11   0.2657
## Year2003           0.07305    0.03568    2.05   0.0407 *
## Year2004          -0.04882    0.03608   -1.35   0.1761
## Year2005          -0.04787    0.03602   -1.33   0.1839
## Year2006           0.00988    0.03434    0.29   0.7735
## Year2007          -0.04500    0.03343   -1.35   0.1783
## Year2008          -0.06091    0.03287   -1.85   0.0639 .
## Year2009          -0.14520    0.03346   -4.34  1.4e-05 ***
## Year2010          -0.16829    0.03217   -5.23  1.7e-07 ***
## Year2011          -0.14320    0.03293   -4.35  1.4e-05 ***
## Year2012          -0.15334    0.03400   -4.51  6.5e-06 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.81
## Multiple R-squared:  0.0086, Adjusted R-squared:  0.00782
## Convergence in 10 IRWLS iterations
##
## Robustness weights:
##  observation 19867 is an outlier with |weight| = 0 ( < 4.6e-06);
##  1826 weights are ~= 1. The remaining 19763 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0194  0.8660  0.9530  0.9180  0.9860  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          4.63e-06          1.82e-12
```

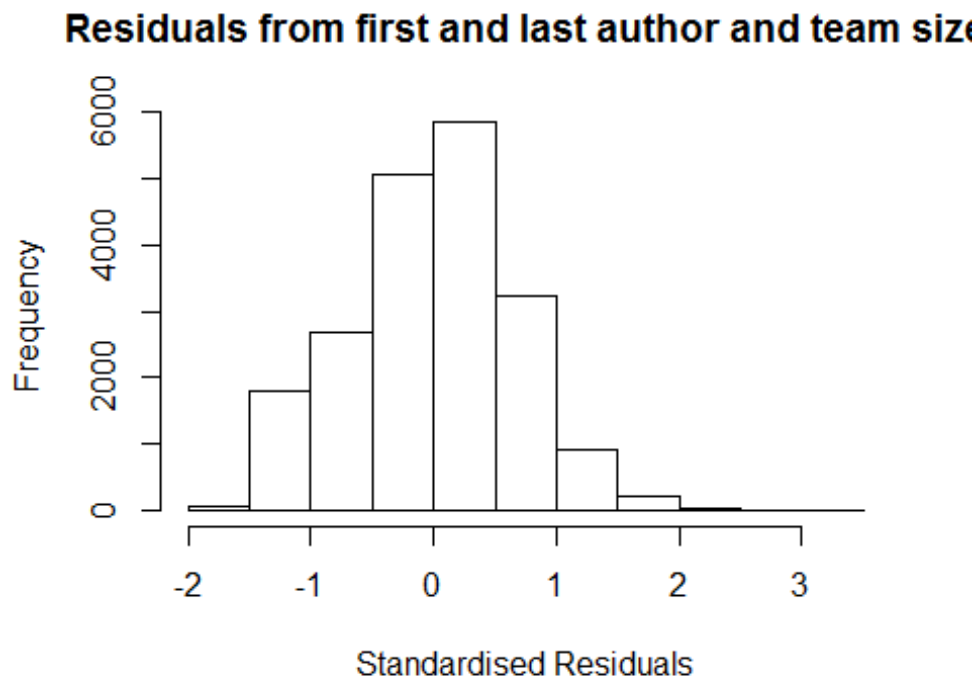
```
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Sample size for the above analysis:  21590"
## [1] ""
## [1] ""
## [1] "####################################"
## [1] "Analysis of AJSC 1203"
## [1] "####################################"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 1048 1117  992 1113 1184 1234 1357 1126 1308 1427 1408 1587 1654 1853 1886
## 2011 2012
## 2139 2317
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  901  912  859  930  998  963 1181  976 1115 1194 1186 1371 1401 1564 1614
## 2011 2012
## 1781 1963
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  865  890  824  896  949  917 1128  925 1053 1135 1139 1316 1329 1469 1509
## 2011 2012
## 1676 1866
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
##  Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 300, df = 16, p-value <2e-16
```




```
##
##  Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 50, df = 1, p-value = 2e-12
```



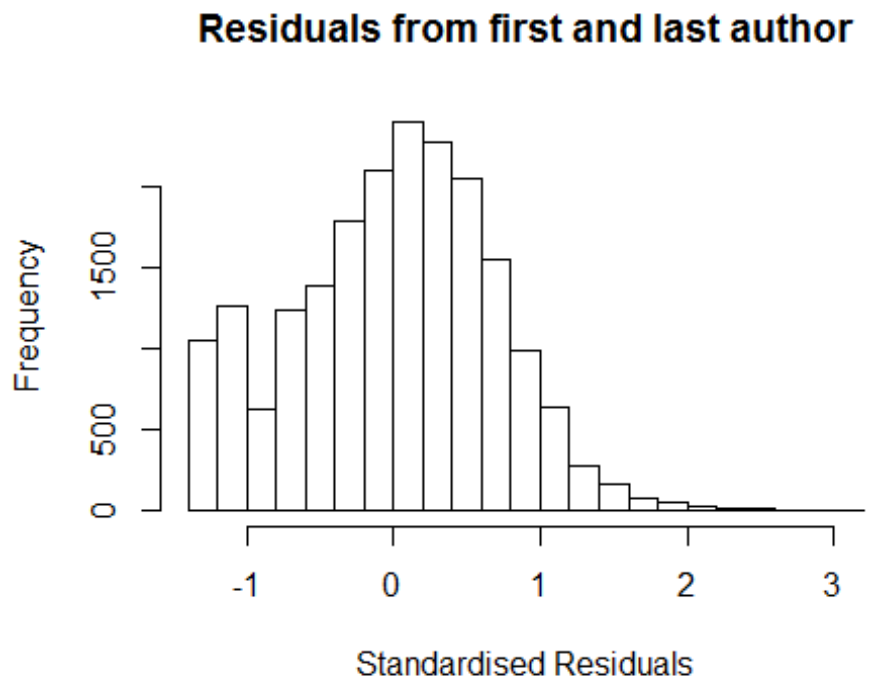

```
## [1] "Female first author team size 2018 geometric mean: 1.6406335973424"
## [1] "Male first author team size 2018 geometric mean: 1.51746959988562"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 370000, p-value = 0.005
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 1.59832107533102"
## [1] "Male last author team size 2018 geometric mean: 1.57428029493616"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 360000, p-value = 0.5
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.436  1           1.198
## LastAuthorFemale  1.433  1           1.197
## UniqueAuthors     1.026  4           1.003
## Year              1.027 16           1.001
```




```
## [1] "List of  6 outliers with residuals above  2.5"
##          ScopusId  NLCS Year OneField Fields residuals
## 1262   0031381525 3.993 1997     1203      3     2.542
## 18103 70949095633 4.199 2009     1203      2     3.203
## 22807 80054769198 3.695 2011     1203      2     2.655
## 23755 79955674091 3.677 2011     1203      6     2.581
## 26041 84872871724 3.721 2012     1203      3     2.700
## 26486 84856523526 3.906 2012     1203      3     2.512
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
UniqueAuthors +
##     Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##     k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.6109 -0.4503  0.0294  0.4412  3.2031
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.00254    0.02232   44.92   <2e-16 ***
## FirstAuthorFemale1  0.02439    0.01145    2.13   0.0333 *
## LastAuthorFemale1   0.03167    0.01143    2.77   0.0056 **
## UniqueAuthors2      0.42830    0.01135   37.72   <2e-16 ***
## UniqueAuthors3      0.45836    0.01451   31.58   <2e-16 ***
```

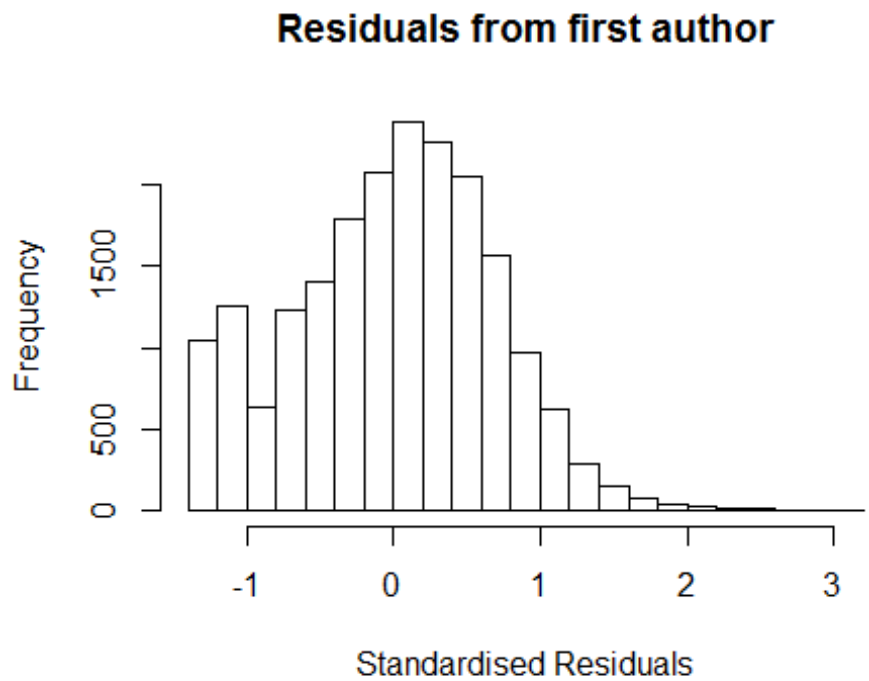
```
## UniqueAuthors4      0.47059    0.01969   23.90   <2e-16 ***
## UniqueAuthors5      0.51525    0.02309   22.31   <2e-16 ***
## Year1997            0.02043    0.03124    0.65   0.5131
## Year1998            0.01937    0.03075    0.63   0.5288
## Year1999           -0.04907    0.03014   -1.63   0.1035
## Year2000           -0.05402    0.02878   -1.88   0.0605 .
## Year2001           -0.09128    0.03185   -2.87   0.0042 **
## Year2002           -0.00135    0.03050   -0.04   0.9647
## Year2003            0.00542    0.03061    0.18   0.8595
## Year2004            0.01520    0.03170    0.48   0.6315
## Year2005            0.04788    0.02899    1.65   0.0986 .
## Year2006            0.00837    0.02853    0.29   0.7693
## Year2007            0.03598    0.02848    1.26   0.2064
## Year2008           -0.02943    0.02804   -1.05   0.2940
## Year2009           -0.00668    0.02705   -0.25   0.8049
## Year2010           -0.04494    0.02670   -1.68   0.0923 .
## Year2011            0.03702    0.02763    1.34   0.1802
## Year2012           -0.03722    0.02916   -1.28   0.2017
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.655
## Multiple R-squared:  0.106,  Adjusted R-squared:  0.105
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  observation 13403 is an outlier with |weight| = 0 ( < 5e-06);
##  1669 weights are ~= 1. The remaining 18216 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.051   0.855   0.949   0.908   0.986   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          5.03e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.561  1           1.249
```

```
## LastAuthorFemale  1.564  1           1.251
## Year              1.010 16           1.000
```



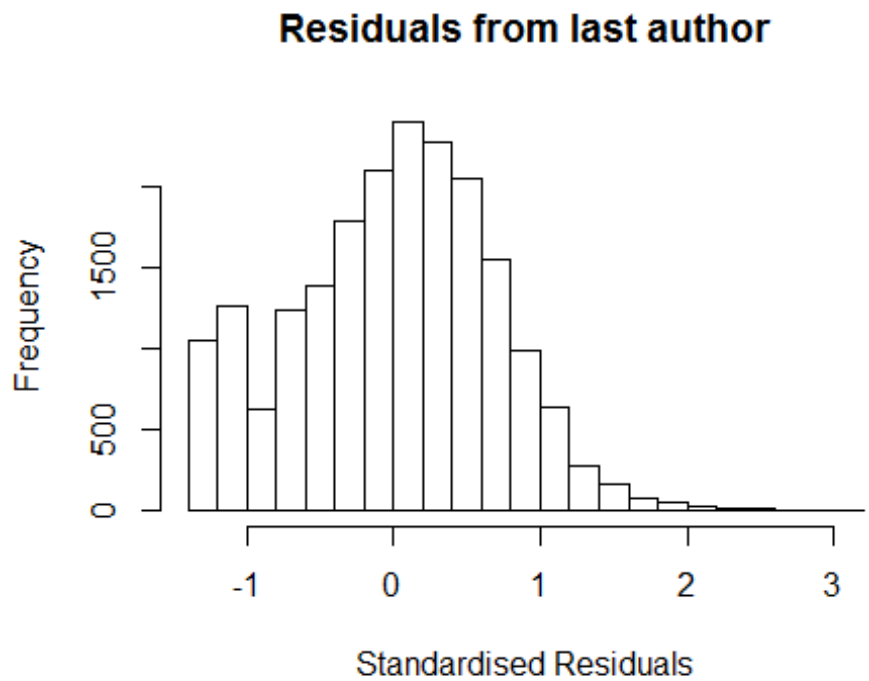

```
## [1] "List of  5 outliers with residuals above  2.5"
##          ScopusId  NLCS Year OneField Fields residuals
## 1262   0031381525 3.993 1997     1203      3     2.825
## 16265 60950608143 3.692 2008     1203      3     2.513
## 18103 70949095633 4.199 2009     1203      2     3.025
## 25949 84857205310 3.786 2012     1203      3     2.552
## 26486 84856523526 3.906 2012     1203      3     2.765
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.3169 -0.4703  0.0413  0.4785  3.0251
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)          1.16538    0.02328   50.06   <2e-16 ***
## FirstAuthorFemale1  0.05776    0.01275    4.53    6e-06 ***
## LastAuthorFemale1   0.03554    0.01275    2.79   0.0053 **
## Year1997            0.00253    0.03294    0.08   0.9389
## Year1998            0.00772    0.03241    0.24   0.8118
```

```
## Year1999           -0.06650    0.03197   -2.08   0.0375 *
## Year2000           -0.05189    0.03037   -1.71   0.0876 .
## Year2001           -0.10513    0.03290   -3.20   0.0014 **
## Year2002            0.00622    0.03250    0.19   0.8482
## Year2003           -0.00849    0.03292   -0.26   0.7966
## Year2004            0.00882    0.03380    0.26   0.7942
## Year2005            0.05822    0.03066    1.90   0.0576 .
## Year2006            0.01677    0.03034    0.55   0.5803
## Year2007            0.02788    0.03012    0.93   0.3546
## Year2008           -0.02175    0.02973   -0.73   0.4645
## Year2009            0.00855    0.02853    0.30   0.7644
## Year2010           -0.02424    0.02830   -0.86   0.3918
## Year2011            0.05788    0.02933    1.97   0.0485 *
## Year2012           -0.02453    0.03155   -0.78   0.4368
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.691
## Multiple R-squared:  0.00691,    Adjusted R-squared:  0.00601
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
##  1620 weights are ~= 1. The remaining 18266 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0161  0.8670  0.9480  0.9090  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          5.03e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.006  1           1.003
## Year              1.006 16           1.000
```




```
## [1] "List of  5 outliers with residuals above  2.5"
##          ScopusId  NLCS Year OneField Fields residuals
## 1262   0031381525 3.993 1997     1203      3     2.825
## 16265 60950608143 3.692 2008     1203      3     2.513
## 18103 70949095633 4.199 2009     1203      2     3.025
## 25949 84857205310 3.786 2012     1203      3     2.552
## 26486 84856523526 3.906 2012     1203      3     2.765
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.3107 -0.4723  0.0419  0.4777  3.0184
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.16928    0.02320   50.40  < 2e-16 ***
## FirstAuthorFemale1  0.08120    0.01027    7.91  2.8e-15 ***
## Year1997            0.00364    0.03294    0.11   0.9121
## Year1998            0.00930    0.03237    0.29   0.7739
## Year1999           -0.06531    0.03196   -2.04   0.0410 *
## Year2000           -0.05018    0.03036   -1.65   0.0984 .
## Year2001           -0.10336    0.03286   -3.15   0.0017 **
## Year2002            0.00855    0.03249    0.26   0.7925
```

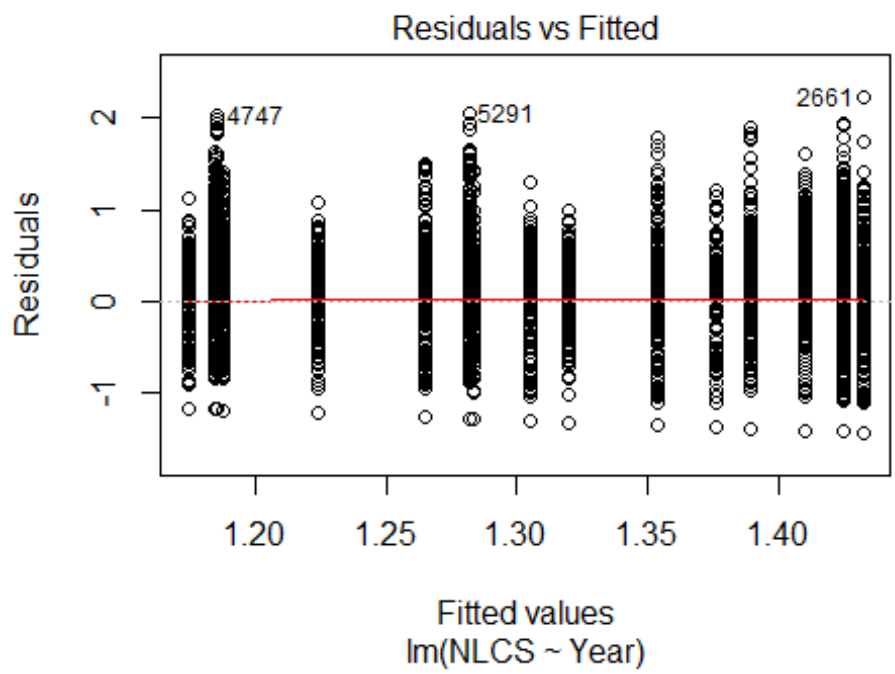
```
## Year2003           -0.00676    0.03290   -0.21   0.8371
## Year2004            0.01093    0.03377    0.32   0.7461
## Year2005            0.05953    0.03064    1.94   0.0520 .
## Year2006            0.01785    0.03032    0.59   0.5561
## Year2007            0.02979    0.03011    0.99   0.3225
## Year2008           -0.01983    0.02970   -0.67   0.5043
## Year2009            0.01131    0.02849    0.40   0.6913
## Year2010           -0.02223    0.02826   -0.79   0.4317
## Year2011            0.06020    0.02931    2.05   0.0400 *
## Year2012           -0.02296    0.03152   -0.73   0.4664
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.691
## Multiple R-squared:  0.00652,    Adjusted R-squared:  0.00567
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
##  1616 weights are ~= 1. The remaining 18270 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0171  0.8670  0.9480  0.9090  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          5.03e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                    psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##                   GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.008  1           1.004
## Year             1.008 16           1.000
```



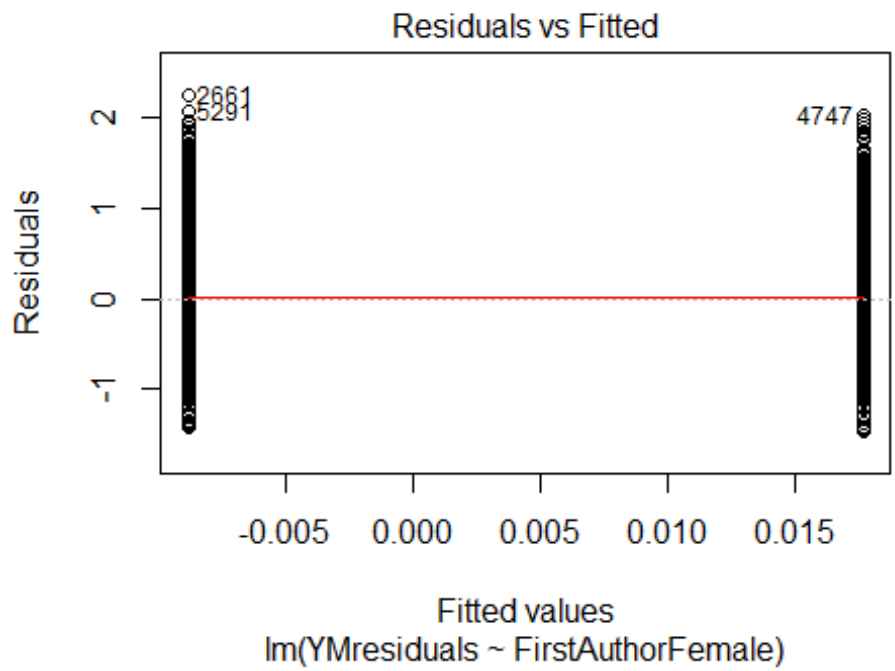

```
## [1] "List of  5 outliers with residuals above  2.5"
##          ScopusId  NLCS Year OneField Fields residuals
## 1262   0031381525 3.993 1997     1203      3     2.825
## 16265 60950608143 3.692 2008     1203      3     2.513
## 18103 70949095633 4.199 2009     1203      2     3.025
## 25949 84857205310 3.786 2012     1203      3     2.552
## 26486 84856523526 3.906 2012     1203      3     2.765
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.3073 -0.4713  0.0419  0.4787  3.0155
##
## Coefficients:
##                   Estimate Std. Error t value Pr(>|t|)
## (Intercept)        1.17544    0.02319   50.68  < 2e-16 ***
## LastAuthorFemale1  0.07368    0.01027    7.17  7.6e-13 ***
## Year1997           0.00140    0.03295    0.04   0.9661
## Year1998           0.00794    0.03247    0.24   0.8068
## Year1999          -0.06724    0.03204   -2.10   0.0359 *
## Year2000          -0.05264    0.03042   -1.73   0.0836 .
## Year2001          -0.10472    0.03294   -3.18   0.0015 **
## Year2002           0.00488    0.03254    0.15   0.8807
```

```
## Year2003          -0.00862    0.03299   -0.26   0.7939
## Year2004           0.00897    0.03386    0.26   0.7912
## Year2005           0.05820    0.03072    1.89   0.0582 .
## Year2006           0.01761    0.03036    0.58   0.5620
## Year2007           0.02805    0.03018    0.93   0.3528
## Year2008          -0.02192    0.02980   -0.74   0.4620
## Year2009           0.00803    0.02859    0.28   0.7788
## Year2010          -0.02326    0.02837   -0.82   0.4122
## Year2011           0.05819    0.02939    1.98   0.0477 *
## Year2012          -0.02356    0.03163   -0.74   0.4565
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.691
## Multiple R-squared:  0.00594,    Adjusted R-squared:  0.00509
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
##  1655 weights are ~= 1. The remaining 18231 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0178  0.8670  0.9480  0.9080  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          5.03e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Sample size for the above analysis:  19886"
## [1] ""
## [1] ""
## [1] "##################################"
## [1] "Analysis of AJSC 1204"
## [1] "##################################"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  136  162  189  180  164  231  268  280  276  288  305  334  367  315  357
## 2011 2012
##  400  454
```

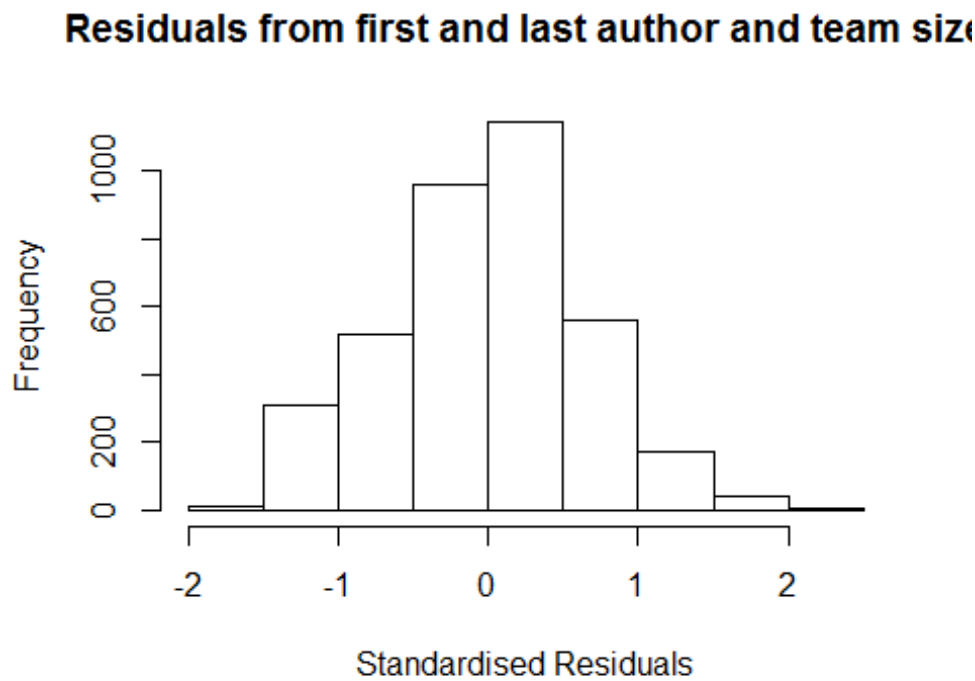
```
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##   98  128  154  150  127  155  234  241  235  237  254  265  310  274  292
## 2011 2012
##  337  379
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##   89  125  150  142  124  151  226  230  231  232  243  256  296  262  277
## 2011 2012
##  320  367
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
##  Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 170, df = 16, p-value <2e-16
```




```
##
##  Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 3.2, df = 1, p-value = 0.08
```

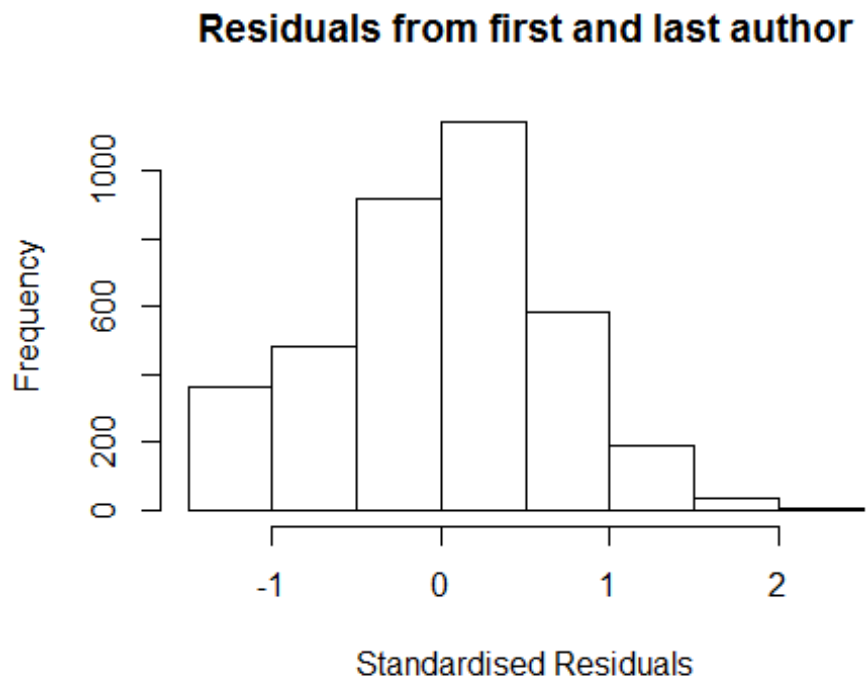



```
## [1] "Female first author team size 2018 geometric mean: 1.66344357903151"
## [1] "Male first author team size 2018 geometric mean: 1.75766357951674"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 23000, p-value = 0.6
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 1.50959394146698"
## [1] "Male last author team size 2018 geometric mean: 1.86466313215361"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 19000, p-value = 0.001
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.617  1           1.272
## LastAuthorFemale  1.622  1           1.274
## UniqueAuthors     1.126  4           1.015
## Year              1.156 16           1.005
```



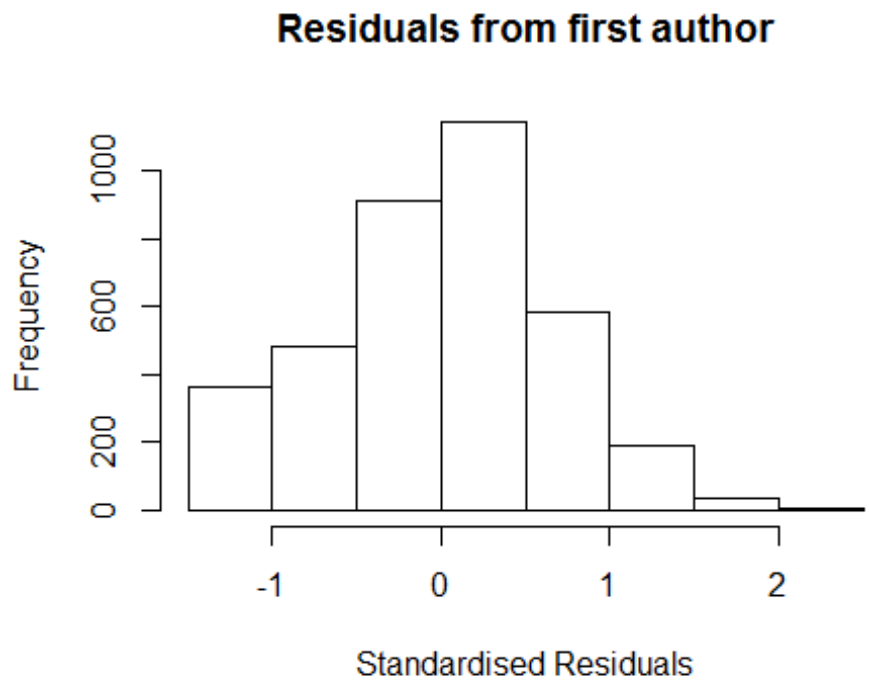

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
UniqueAuthors +
##     Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##     k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.6526 -0.4360  0.0265  0.4302  2.1798
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.24774    0.06299   19.81  < 2e-16 ***
## FirstAuthorFemale1  0.01072    0.02895    0.37  0.71110
## LastAuthorFemale1   0.02648    0.02966    0.89  0.37199
## UniqueAuthors2      0.27220    0.02833    9.61  < 2e-16 ***
## UniqueAuthors3      0.26990    0.03582    7.53  6.1e-14 ***
## UniqueAuthors4      0.31675    0.04515    7.02  2.7e-12 ***
## UniqueAuthors5      0.39807    0.04356    9.14  < 2e-16 ***
## Year1997           -0.04733    0.07874   -0.60  0.54786
## Year1998           -0.03005    0.07270   -0.41  0.67939
## Year1999           -0.18204    0.07456   -2.44  0.01467 *
```

```
## Year2000           -0.16198    0.07611   -2.13  0.03338 *
## Year2001           -0.08650    0.07577   -1.14  0.25368
## Year2002            0.05622    0.07863    0.71  0.47471
## Year2003           -0.00437    0.07722   -0.06  0.95488
## Year2004            0.06167    0.07704    0.80  0.42349
## Year2005            0.07527    0.07525    1.00  0.31725
## Year2006            0.04471    0.07447    0.60  0.54833
## Year2007            0.01273    0.07502    0.17  0.86523
## Year2008           -0.18917    0.07454   -2.54  0.01120 *
## Year2009           -0.12582    0.07439   -1.69  0.09088 .
## Year2010           -0.21265    0.07079   -3.00  0.00268 **
## Year2011           -0.25376    0.07519   -3.37  0.00075 ***
## Year2012           -0.11297    0.07736   -1.46  0.14426
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.631
## Multiple R-squared:  0.0686, Adjusted R-squared:  0.063
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
##  313 weights are ~= 1. The remaining 3408 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.209   0.857   0.950   0.901   0.985   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          2.69e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.620  1           1.273
## LastAuthorFemale  1.617  1           1.271
## Year              1.037 16           1.001
```



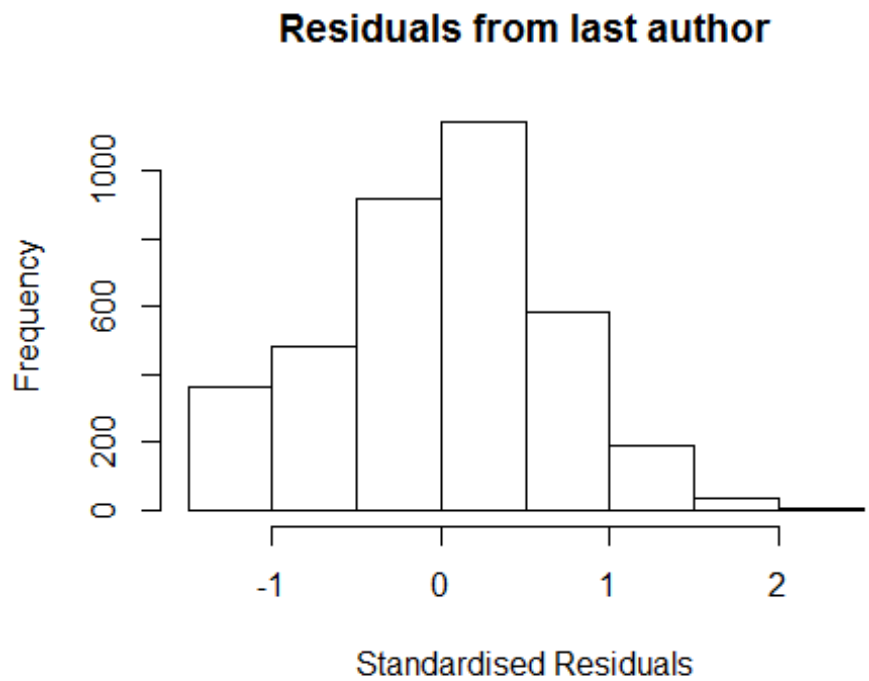

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.4474 -0.4332  0.0347  0.4352  2.2379
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)          1.34984    0.06212   21.73   <2e-16 ***
## FirstAuthorFemale1   0.01782    0.02969    0.60   0.5484
## LastAuthorFemale1    0.00164    0.03034    0.05   0.9568
## Year1997            -0.03821    0.07817   -0.49   0.6250
## Year1998            -0.03990    0.07155   -0.56   0.5771
## Year1999            -0.16817    0.07445   -2.26   0.0240 *
## Year2000            -0.12955    0.07605   -1.70   0.0885 .
## Year2001            -0.06815    0.07393   -0.92   0.3567
## Year2002             0.07814    0.07990    0.98   0.3282
## Year2003            -0.00565    0.07837   -0.07   0.9425
## Year2004             0.06942    0.07741    0.90   0.3699
## Year2005             0.07257    0.07655    0.95   0.3432
```

```
## Year2006            0.07026    0.07435    0.95   0.3447
## Year2007            0.03031    0.07563    0.40   0.6886
## Year2008           -0.17779    0.07586   -2.34   0.0191 *
## Year2009           -0.08017    0.07461   -1.07   0.2826
## Year2010           -0.17940    0.07130   -2.52   0.0119 *
## Year2011           -0.23292    0.07630   -3.05   0.0023 **
## Year2012           -0.09171    0.07809   -1.17   0.2403
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.64
## Multiple R-squared:  0.0235, Adjusted R-squared:  0.0187
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
##  324 weights are ~= 1. The remaining 3397 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.197   0.862   0.949   0.899   0.986   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          2.69e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.026  1           1.013
## Year              1.026 16           1.001
```




```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.4471 -0.4328  0.0337  0.4349  2.2377
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)           1.3501     0.0618   21.84   <2e-16 ***
## FirstAuthorFemale1    0.0189     0.0237    0.80   0.4253
## Year1997             -0.0382     0.0782   -0.49   0.6248
## Year1998             -0.0400     0.0715   -0.56   0.5760
## Year1999             -0.1682     0.0744   -2.26   0.0238 *
## Year2000             -0.1296     0.0760   -1.70   0.0885 .
## Year2001             -0.0682     0.0739   -0.92   0.3561
## Year2002              0.0781     0.0799    0.98   0.3282
## Year2003             -0.0057     0.0783   -0.07   0.9419
## Year2004              0.0693     0.0773    0.90   0.3701
## Year2005              0.0725     0.0766    0.95   0.3435
## Year2006              0.0702     0.0743    0.94   0.3448
```

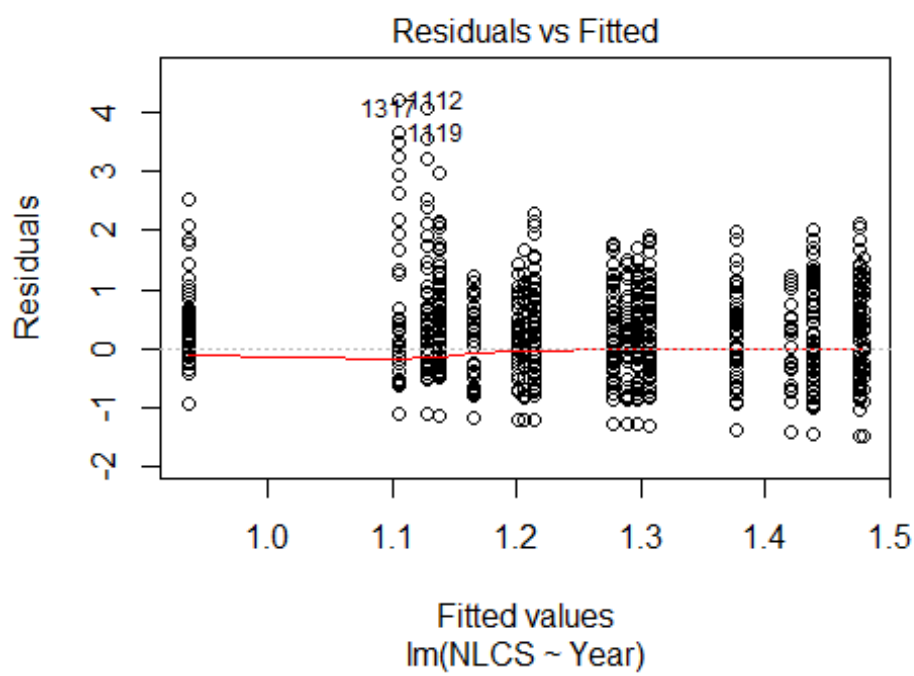
```
## Year2007             0.0303     0.0756    0.40   0.6889
## Year2008            -0.1778     0.0758   -2.35   0.0191 *
## Year2009            -0.0802     0.0746   -1.08   0.2820
## Year2010            -0.1795     0.0713   -2.52   0.0118 *
## Year2011            -0.2329     0.0763   -3.05   0.0023 **
## Year2012            -0.0917     0.0781   -1.18   0.2400
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.64
## Multiple R-squared:  0.0235, Adjusted R-squared:  0.019
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
##  324 weights are ~= 1. The remaining 3397 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.197   0.862   0.949   0.899   0.986   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          2.69e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##                   GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.023  1           1.012
## Year             1.023 16           1.001
```



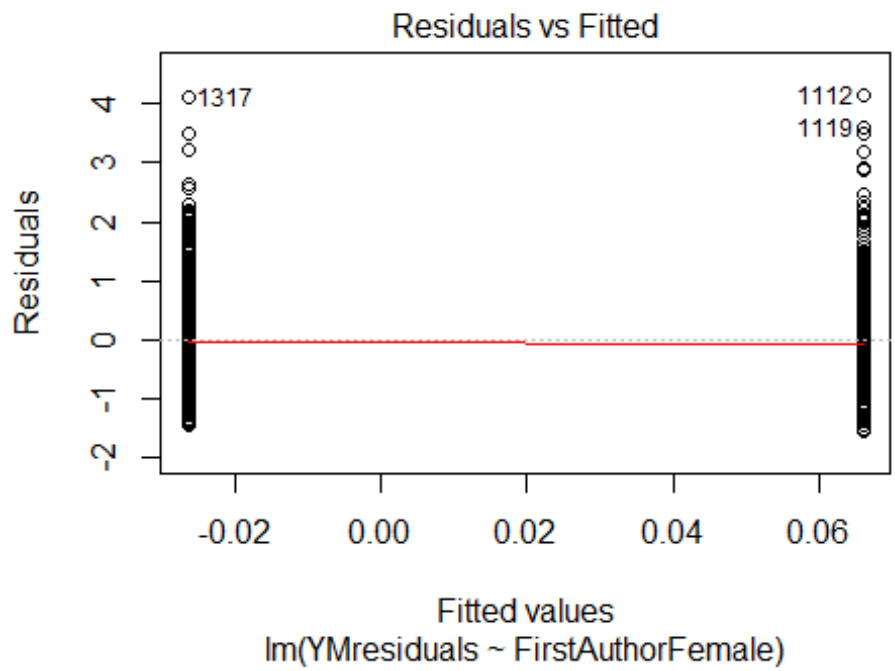

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.4435 -0.4331  0.0352  0.4378  2.2362
##
## Coefficients:
##                   Estimate Std. Error t value Pr(>|t|)
## (Intercept)        1.35195    0.06193   21.83   <2e-16 ***
## LastAuthorFemale1  0.01351    0.02417    0.56   0.5761
## Year1997          -0.03938    0.07804   -0.50   0.6139
## Year1998          -0.03961    0.07153   -0.55   0.5798
## Year1999          -0.16801    0.07445   -2.26   0.0241 *
## Year2000          -0.13016    0.07598   -1.71   0.0868 .
## Year2001          -0.06744    0.07386   -0.91   0.3613
## Year2002           0.07805    0.07987    0.98   0.3285
## Year2003          -0.00571    0.07833   -0.07   0.9419
## Year2004           0.06918    0.07742    0.89   0.3716
## Year2005           0.07296    0.07656    0.95   0.3406
## Year2006           0.06988    0.07434    0.94   0.3473
```

```
## Year2007           0.03038    0.07560    0.40   0.6878
## Year2008          -0.17759    0.07588   -2.34   0.0193 *
## Year2009          -0.07962    0.07457   -1.07   0.2857
## Year2010          -0.17924    0.07126   -2.52   0.0119 *
## Year2011          -0.23279    0.07631   -3.05   0.0023 **
## Year2012          -0.09150    0.07808   -1.17   0.2414
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.64
## Multiple R-squared:  0.0234, Adjusted R-squared:  0.0189
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
##  323 weights are ~= 1. The remaining 3398 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.197   0.862   0.948   0.899   0.986   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          2.69e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Sample size for the above analysis:  3721"
## [1] ""
## [1] ""
## [1] "####################################"
## [1] "Analysis of AJSC 1205"
## [1] "####################################"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##   32   57   53   57   53   53   86   92  100  111  116  135  113   68   94
## 2011 2012
##  131  179
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##   25   45   42   49   39   47   77   79   91   98   91  115   93   63   82
## 2011 2012
```

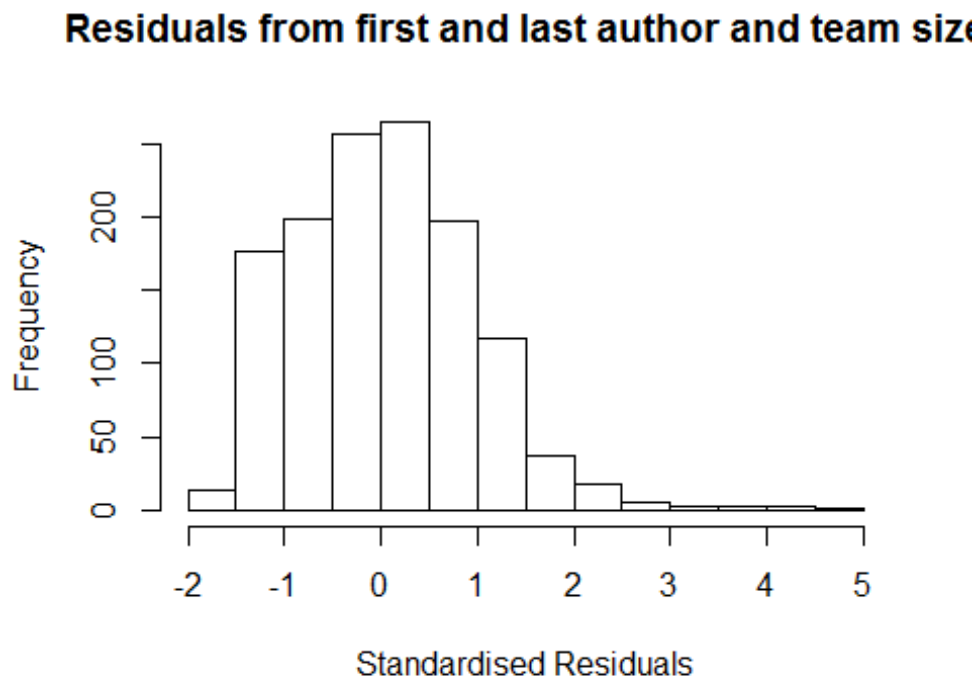
```
##  112  149
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##   25   45   42   49   38   47   77   78   91   97   91  115   92   63   81
## 2011 2012
##  112  149
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
##  Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 69, df = 16, p-value = 1e-08
```




```
##
##  Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 7.5, df = 1, p-value = 0.006
```



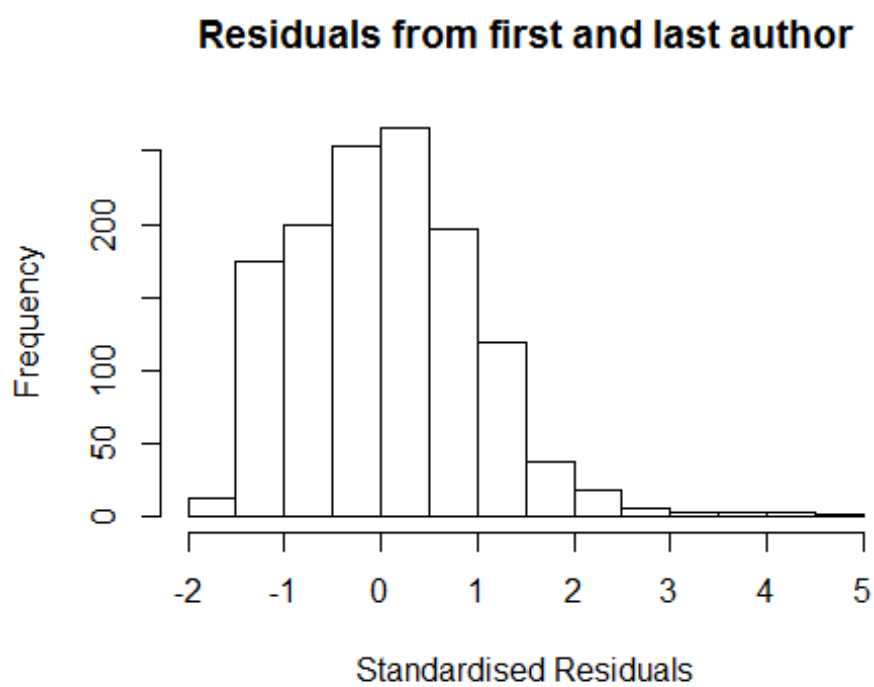

```
## [1] "Female first author team size 2018 geometric mean: 1.01944064370214"
## [1] "Male first author team size 2018 geometric mean: 1.0255256926759"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 980, p-value = 0.8
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 1.03817816766679"
## [1] "Male last author team size 2018 geometric mean: 1.01291879472495"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 1000, p-value = 0.4
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##                     GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 13.35  1           3.653
## LastAuthorFemale  13.70  1           3.702
## UniqueAuthors     28.97  3           1.752
## Year              26.63 16           1.108
```




```
## [1] "List of  14 outliers with residuals above  2.5"
##         ScopusId  NLCS Year OneField Fields residuals
## 1093 70450263771 3.724 2009     1205      1     3.049
## 1095 70450199674 3.298 2009     1205      1     2.567
## 1112 69949087596 5.315 2009     1205      1     4.584
## 1113 69949093703 4.348 2009     1205      1     3.617
## 1114 69949095259 4.064 2009     1205      1     3.333
## 1115 69949097607 4.590 2009     1205      1     3.915
## 1119 77950181710 4.785 2009     1205      2     4.054
## 1204 77950212744 3.462 2010     1205      1     2.587
## 1292 84864342269 3.662 2011     1205      1     2.653
## 1315 79958036692 4.676 2011     1205      1     3.667
## 1317 79958056051 5.217 2011     1205      1     4.264
## 1323 80052769954 3.662 2011     1205      1     2.653
## 1331 84855278391 4.333 2011     1205      1     3.380
## 1493 84863489169 4.115 2012     1205      1     2.983
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
UniqueAuthors +
##     Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##     k.max = 1000))
##  \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.56039 -0.68677  0.00788  0.61997  4.58423
```

```
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)          1.4126     0.1806    7.82  1.1e-14 ***
## FirstAuthorFemale1   0.0381     0.1786    0.21  0.83098
## LastAuthorFemale1    0.0176     0.1778    0.10  0.92122
## UniqueAuthors2       0.1749     0.1439    1.22  0.22428
## UniqueAuthors4       1.0771     0.1182    9.12  < 2e-16 ***
## UniqueAuthors5       1.3285     0.1847    7.19  1.1e-12 ***
## Year1997            -0.2662     0.2096   -1.27  0.20430
## Year1998            -0.2264     0.2189   -1.03  0.30108
## Year1999            -0.2338     0.1978   -1.18  0.23751
## Year2000             0.0920     0.2357    0.39  0.69632
## Year2001            -0.1344     0.2118   -0.63  0.52598
## Year2002            -0.1736     0.2042   -0.85  0.39532
## Year2003            -0.0776     0.2136   -0.36  0.71651
## Year2004            -0.0135     0.2043   -0.07  0.94723
## Year2005             0.0452     0.2137    0.21  0.83236
## Year2006            -0.1326     0.2011   -0.66  0.50974
## Year2007            -0.1328     0.1975   -0.67  0.50164
## Year2008            -0.2500     0.2057   -1.22  0.22433
## Year2009            -0.7376     0.2141   -3.44  0.00059 ***
## Year2010            -0.5378     0.2034   -2.64  0.00828 **
## Year2011            -0.4596     0.2062   -2.23  0.02599 *
## Year2012            -0.3360     0.1960   -1.71  0.08678 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.941
## Multiple R-squared:  0.0535, Adjusted R-squared:  0.0378
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  observation 918 is an outlier with |weight| = 0 ( < 7.7e-05);
##  96 weights are ~= 1. The remaining 1195 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0042  0.8790  0.9480  0.9150  0.9870  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          7.74e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
```

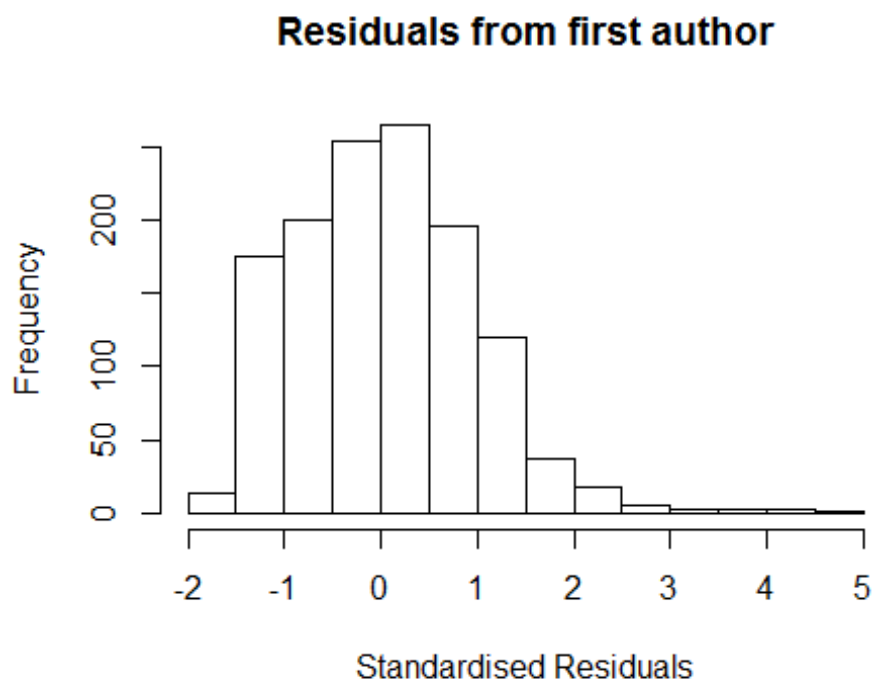
```
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##                     GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 13.451  1           3.668
## LastAuthorFemale  13.464  1           3.669
## Year               1.059 16           1.002
```




```
## [1] "List of  14 outliers with residuals above  2.5"
##         ScopusId  NLCS Year OneField Fields residuals
## 1093 70450263771 3.724 2009     1205      1     3.018
## 1095 70450199674 3.298 2009     1205      1     2.537
## 1112 69949087596 5.315 2009     1205      1     4.554
## 1113 69949093703 4.348 2009     1205      1     3.587
## 1114 69949095259 4.064 2009     1205      1     3.303
## 1115 69949097607 4.590 2009     1205      1     3.884
## 1119 77950181710 4.785 2009     1205      2     4.024
## 1204 77950212744 3.462 2010     1205      1     2.584
## 1292 84864342269 3.662 2011     1205      1     2.650
## 1315 79958036692 4.676 2011     1205      1     3.664
## 1317 79958056051 5.217 2011     1205      1     4.259
## 1323 80052769954 3.662 2011     1205      1     2.650
## 1331 84855278391 4.333 2011     1205      1     3.375
## 1493 84863489169 4.115 2012     1205      1     2.973
##
```

```
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.56356 -0.70648  0.00901  0.63248  4.55449
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.41309    0.18040    7.83    1e-14 ***
## FirstAuthorFemale1 -0.06030    0.21213   -0.28   0.7763
## LastAuthorFemale1   0.11433    0.21147    0.54   0.5888
## Year1997           -0.26612    0.20936   -1.27   0.2039
## Year1998           -0.22629    0.21868   -1.03   0.3010
## Year1999           -0.23203    0.19772   -1.17   0.2408
## Year2000            0.09645    0.23518    0.41   0.6818
## Year2001           -0.13438    0.21163   -0.63   0.5256
## Year2002           -0.16373    0.20378   -0.80   0.4219
## Year2003           -0.07123    0.21314   -0.33   0.7383
## Year2004            0.00603    0.20474    0.03   0.9765
## Year2005            0.06195    0.21346    0.29   0.7717
## Year2006           -0.12561    0.20043   -0.63   0.5310
## Year2007           -0.12591    0.19716   -0.64   0.5232
## Year2008           -0.24935    0.20543   -1.21   0.2250
## Year2009           -0.70661    0.21607   -3.27   0.0011 **
## Year2010           -0.53470    0.20320   -2.63   0.0086 **
## Year2011           -0.45508    0.20583   -2.21   0.0272 *
## Year2012           -0.32557    0.19623   -1.66   0.0973 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.947
## Multiple R-squared:  0.0464, Adjusted R-squared:  0.0329
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  observation 918 is an outlier with |weight| = 0 ( < 7.7e-05);
##  97 weights are ~= 1. The remaining 1194 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0061  0.8800  0.9470  0.9150  0.9860  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          7.74e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
```

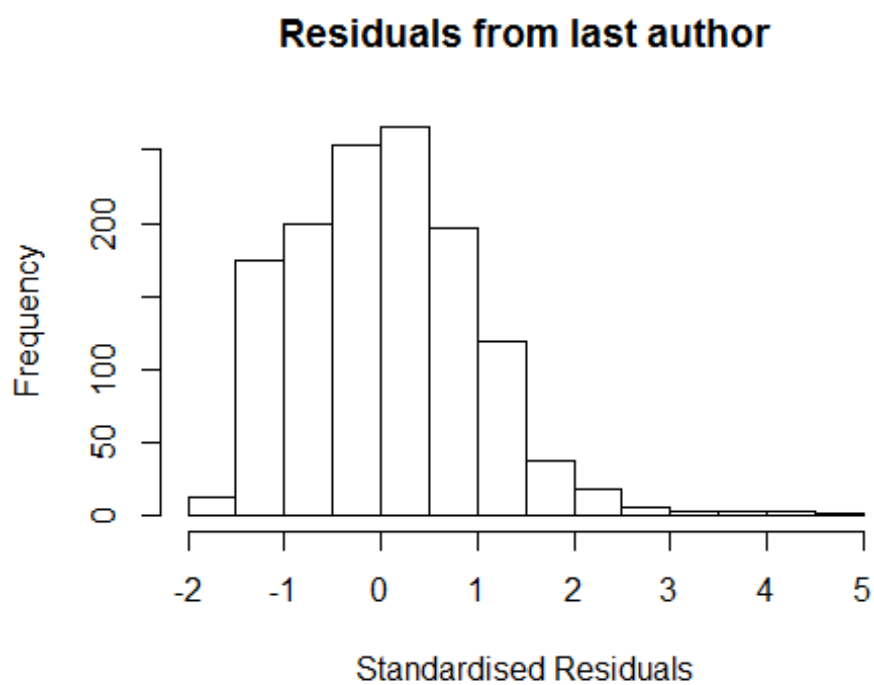
```
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.035  1           1.017
## Year              1.035 16           1.001
```




```
## [1] "List of  14 outliers with residuals above  2.5"
##         ScopusId  NLCS Year OneField Fields residuals
## 1093 70450263771 3.724 2009     1205      1     3.018
## 1095 70450199674 3.298 2009     1205      1     2.537
## 1112 69949087596 5.315 2009     1205      1     4.554
## 1113 69949093703 4.348 2009     1205      1     3.587
## 1114 69949095259 4.064 2009     1205      1     3.303
## 1115 69949097607 4.590 2009     1205      1     3.884
## 1119 77950181710 4.785 2009     1205      2     4.024
## 1204 77950212744 3.462 2010     1205      1     2.584
## 1292 84864342269 3.662 2011     1205      1     2.650
## 1315 79958036692 4.676 2011     1205      1     3.664
## 1317 79958056051 5.217 2011     1205      1     4.259
## 1323 80052769954 3.662 2011     1205      1     2.650
```

```
## 1331 84855278391 4.333 2011     1205      1     3.375
## 1493 84863489169 4.115 2012     1205      1     2.973
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.56792 -0.70693  0.00817  0.63198  4.55737
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.41416    0.18033    7.84  9.3e-15 ***
## FirstAuthorFemale1  0.05071    0.05882    0.86   0.3888
## Year1997           -0.26633    0.20933   -1.27   0.2035
## Year1998           -0.22623    0.21873   -1.03   0.3012
## Year1999           -0.22979    0.19789   -1.16   0.2458
## Year2000            0.10305    0.23436    0.44   0.6602
## Year2001           -0.13446    0.21161   -0.64   0.5253
## Year2002           -0.16367    0.20374   -0.80   0.4219
## Year2003           -0.07311    0.21323   -0.34   0.7317
## Year2004            0.00443    0.20458    0.02   0.9827
## Year2005            0.06070    0.21329    0.28   0.7760
## Year2006           -0.12583    0.20033   -0.63   0.5300
## Year2007           -0.12564    0.19721   -0.64   0.5242
## Year2008           -0.24955    0.20541   -1.21   0.2246
## Year2009           -0.70723    0.21596   -3.27   0.0011 **
## Year2010           -0.53495    0.20316   -2.63   0.0086 **
## Year2011           -0.45413    0.20579   -2.21   0.0275 *
## Year2012           -0.32520    0.19621   -1.66   0.0977 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.946
## Multiple R-squared:  0.0462, Adjusted R-squared:  0.0335
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  observation 918 is an outlier with |weight| = 0 ( < 7.7e-05);
##  95 weights are ~= 1. The remaining 1196 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0059  0.8800  0.9470  0.9150  0.9860  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          7.74e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
```

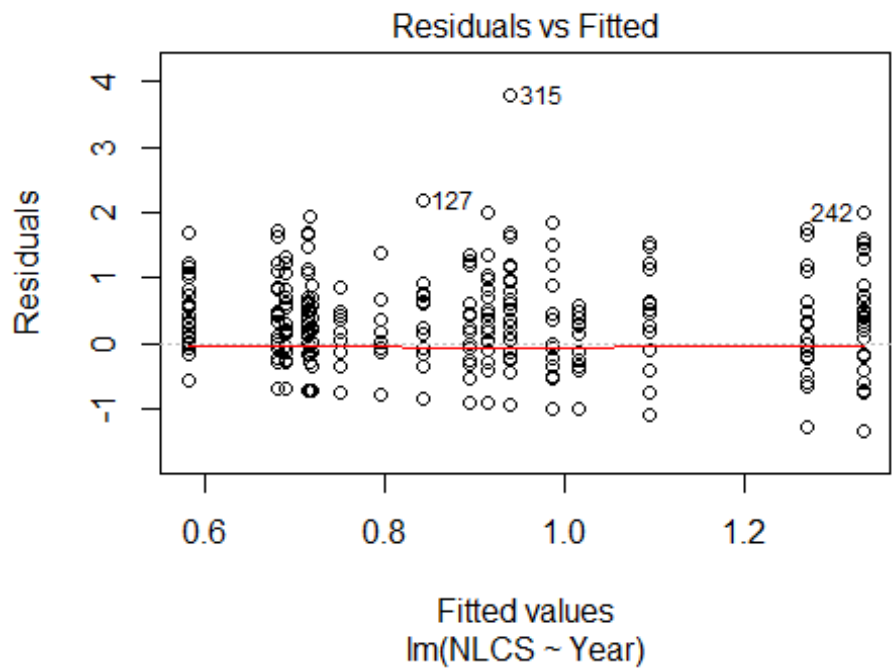
```
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##                   GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.036  1           1.018
## Year             1.036 16           1.001
```



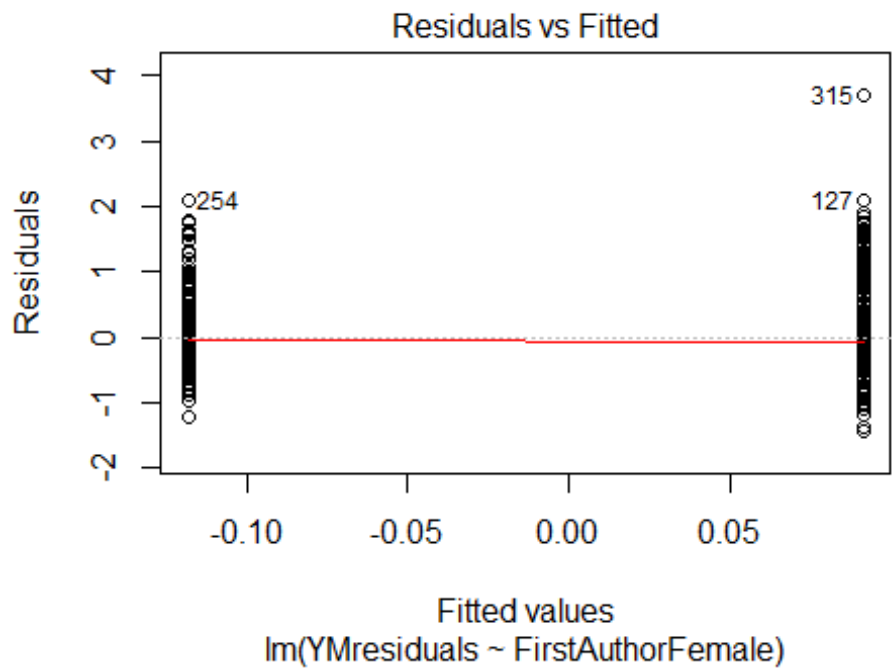

```
## [1] "List of  14 outliers with residuals above  2.5"
##         ScopusId  NLCS Year OneField Fields residuals
## 1093 70450263771 3.724 2009     1205      1     3.018
## 1095 70450199674 3.298 2009     1205      1     2.537
## 1112 69949087596 5.315 2009     1205      1     4.554
## 1113 69949093703 4.348 2009     1205      1     3.587
## 1114 69949095259 4.064 2009     1205      1     3.303
## 1115 69949097607 4.590 2009     1205      1     3.884
## 1119 77950181710 4.785 2009     1205      2     4.024
## 1204 77950212744 3.462 2010     1205      1     2.584
## 1292 84864342269 3.662 2011     1205      1     2.650
## 1315 79958036692 4.676 2011     1205      1     3.664
```

```
## 1317 79958056051 5.217 2011     1205      1     4.259
## 1323 80052769954 3.662 2011     1205      1     2.650
## 1331 84855278391 4.333 2011     1205      1     3.375
## 1493 84863489169 4.115 2012     1205      1     2.973
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.56850 -0.70548  0.00639  0.63309  4.55356
##
## Coefficients:
##                   Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.4125     0.1804    7.83    1e-14 ***
## LastAuthorFemale1   0.0560     0.0587    0.95   0.3406
## Year1997           -0.2661     0.2094   -1.27   0.2042
## Year1998           -0.2263     0.2187   -1.03   0.3009
## Year1999           -0.2309     0.1978   -1.17   0.2434
## Year2000            0.1001     0.2343    0.43   0.6694
## Year2001           -0.1343     0.2117   -0.63   0.5258
## Year2002           -0.1638     0.2039   -0.80   0.4218
## Year2003           -0.0719     0.2133   -0.34   0.7361
## Year2004            0.0056     0.2048    0.03   0.9782
## Year2005            0.0614     0.2135    0.29   0.7738
## Year2006           -0.1255     0.2005   -0.63   0.5313
## Year2007           -0.1253     0.1973   -0.64   0.5252
## Year2008           -0.2494     0.2055   -1.21   0.2251
## Year2009           -0.7070     0.2161   -3.27   0.0011 **
## Year2010           -0.5347     0.2033   -2.63   0.0086 **
## Year2011           -0.4545     0.2058   -2.21   0.0274 *
## Year2012           -0.3254     0.1963   -1.66   0.0976 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.946
## Multiple R-squared:  0.0464, Adjusted R-squared:  0.0336
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  observation 918 is an outlier with |weight| = 0 ( < 7.7e-05);
##  97 weights are ~= 1. The remaining 1194 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0058  0.8800  0.9470  0.9150  0.9860  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
```

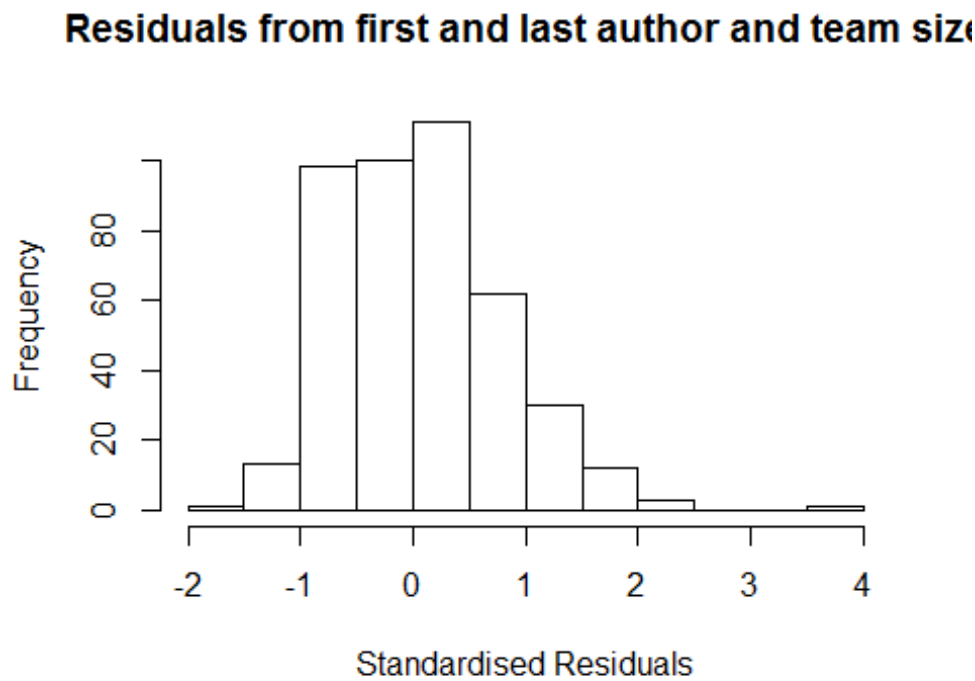
```
##          1.00e-07          1.00e-07          7.74e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                    psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Sample size for the above analysis:  1292"
## [1] ""
## [1] ""
## [1] "####################################"
## [1] "Analysis of AJSC 1206"
## [1] "####################################"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##   13    9   14   13   17   27   29   25   28   44   45   39   35   54   40
## 2011 2012
##   47   56
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##   10    9   11   11   16   18   25   23   27   38   38   36   29   49   34
## 2011 2012
##   38   42
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    9    8   11   10   14   18   23   20   24   35   38   36   27   49   32
## 2011 2012
##   38   39
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
##  Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 31, df = 16, p-value = 0.01
```




```
##
##  Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 17, df = 1, p-value = 4e-05
```

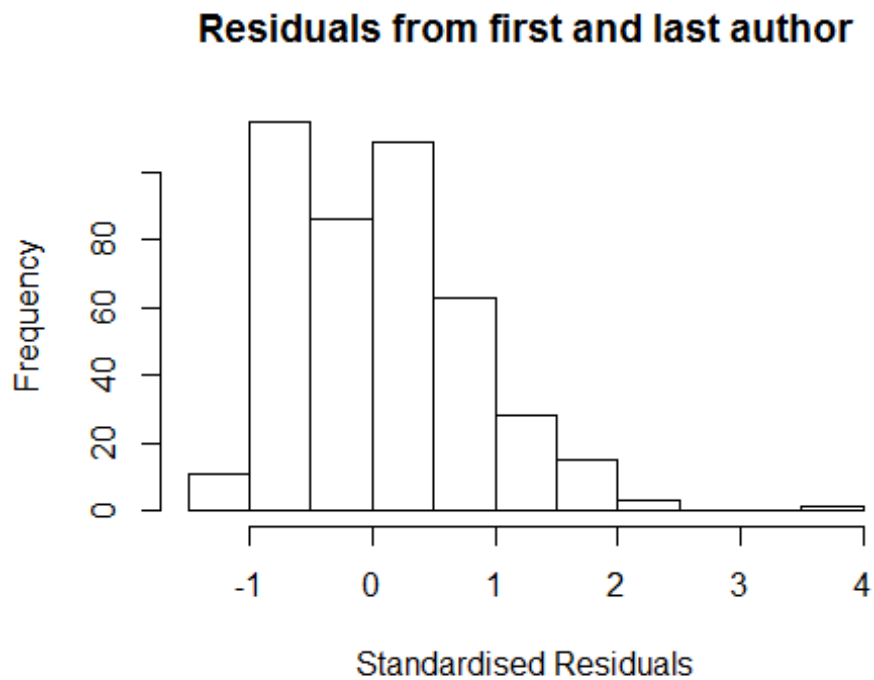



```
## [1] "Female first author team size 2018 geometric mean: 1.69803548691214"
## [1] "Male first author team size 2018 geometric mean: 1.22776392870837"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 1200, p-value = 0.01
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 1.43503100757191"
## [1] "Male last author team size 2018 geometric mean: 1.65753455409486"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 910, p-value = 0.3
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 7.166  1           2.677
## LastAuthorFemale  7.374  1           2.716
## UniqueAuthors     2.144  4           1.100
## Year              2.434 16           1.028
```



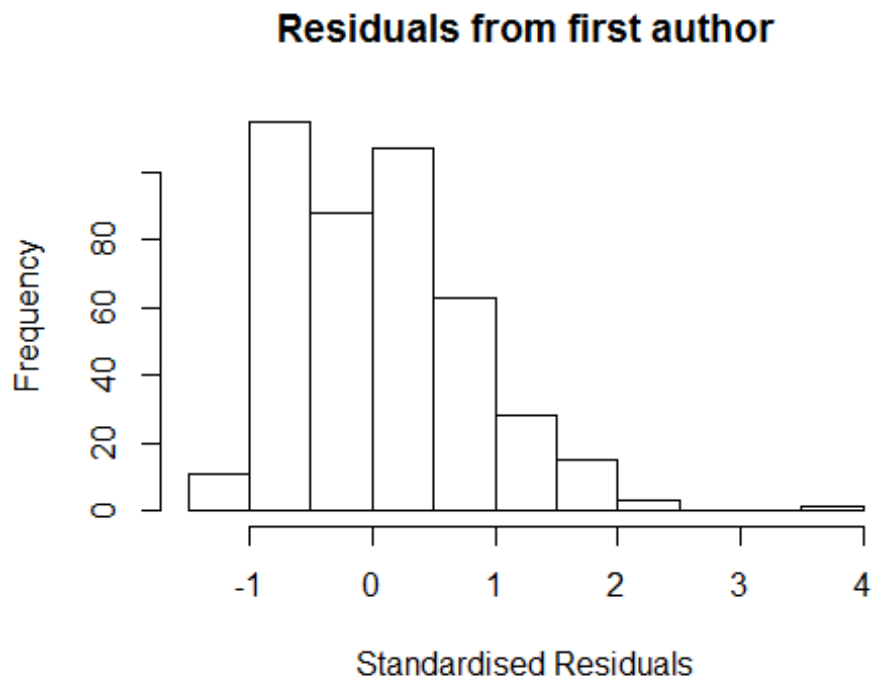

```
## [1] "List of  1 outliers with residuals above  2.5"
##        ScopusId NLCS Year OneField Fields residuals
## 315 84973631259 4.73 2007     1206      3     3.941
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
UniqueAuthors +
##     Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##     k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.5684 -0.5150  0.0249  0.4900  3.9409
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         0.78711    0.21704    3.63  0.00032 ***
## FirstAuthorFemale1 -0.22805    0.17292   -1.32  0.18798
## LastAuthorFemale1   0.08817    0.17330    0.51  0.61121
## UniqueAuthors2      0.24100    0.14252    1.69  0.09160 .
## UniqueAuthors3      0.61693    0.23865    2.59  0.01008 *
## UniqueAuthors4      0.58254    0.19419    3.00  0.00287 **
## UniqueAuthors5      0.51341    0.19561    2.62  0.00900 **
## Year1997           -0.10129    0.32505   -0.31  0.75550
## Year1998            0.16432    0.26815    0.61  0.54035
## Year1999           -0.12961    0.25261   -0.51  0.60817
```

```
## Year2000           -0.12446    0.24554   -0.51  0.61249
## Year2001            0.09512    0.28721    0.33  0.74066
## Year2002           -0.09486    0.25784   -0.37  0.71313
## Year2003           -0.00824    0.31096   -0.03  0.97886
## Year2004            0.31845    0.25562    1.25  0.21354
## Year2005            0.43598    0.27168    1.60  0.10932
## Year2006            0.11319    0.26676    0.42  0.67155
## Year2007            0.00202    0.27480    0.01  0.99412
## Year2008            0.14897    0.24046    0.62  0.53593
## Year2009           -0.27209    0.23593   -1.15  0.24948
## Year2010           -0.15511    0.24765   -0.63  0.53145
## Year2011           -0.21002    0.24305   -0.86  0.38805
## Year2012           -0.23348    0.25734   -0.91  0.36480
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.69
## Multiple R-squared:  0.133,  Adjusted R-squared:  0.0868
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
##  observation 230 is an outlier with |weight| = 0 ( < 0.00023);
##  37 weights are ~= 1. The remaining 393 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.313   0.884   0.946   0.905   0.981   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          2.32e-04          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 8.202  1           2.864
## LastAuthorFemale  7.881  1           2.807
## Year              1.348 16           1.009
```



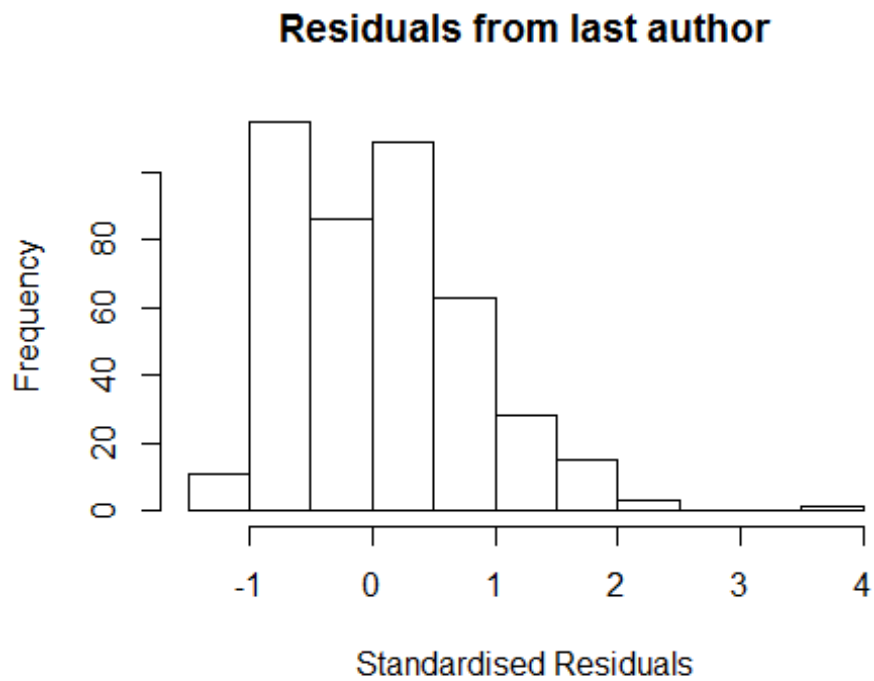

```
## [1] "List of  1 outliers with residuals above  2.5"
##        ScopusId NLCS Year OneField Fields residuals
## 315 84973631259 4.73 2007     1206      3      3.91
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.29046 -0.52720  0.00132  0.50988  3.90966
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)           0.8326     0.2549    3.27   0.0012 **
## FirstAuthorFemale1   -0.1812     0.1945   -0.93   0.3522
## LastAuthorFemale1     0.0204     0.1916    0.11   0.9154
## Year1997             -0.1343     0.3532   -0.38   0.7039
## Year1998              0.2415     0.2900    0.83   0.4055
## Year1999             -0.1134     0.2912   -0.39   0.6972
## Year2000             -0.0986     0.2933   -0.34   0.7369
## Year2001              0.0983     0.3163    0.31   0.7562
## Year2002             -0.0436     0.2962   -0.15   0.8829
## Year2003              0.0959     0.3444    0.28   0.7809
## Year2004              0.3960     0.2898    1.37   0.1725
## Year2005              0.4579     0.3037    1.51   0.1324
```

```
## Year2006             0.0868     0.2949    0.29   0.7686
## Year2007            -0.0122     0.3021   -0.04   0.9677
## Year2008             0.1541     0.2730    0.56   0.5728
## Year2009            -0.1960     0.2688   -0.73   0.4664
## Year2010            -0.1091     0.2818   -0.39   0.6989
## Year2011            -0.1679     0.2763   -0.61   0.5437
## Year2012            -0.1446     0.2818   -0.51   0.6082
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.712
## Multiple R-squared:  0.0811, Adjusted R-squared:  0.041
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
##  observation 230 is an outlier with |weight| = 0 ( < 0.00023);
##  45 weights are ~= 1. The remaining 385 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.297   0.883   0.944   0.903   0.977   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          2.32e-04          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.224  1           1.107
## Year              1.224 16           1.006
```




```
## [1] "List of  1 outliers with residuals above  2.5"
##        ScopusId NLCS Year OneField Fields residuals
## 315 84973631259 4.73 2007     1206      3      3.91
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.29157 -0.52694  0.00203  0.50922  3.90784
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)           0.8337     0.2547    3.27   0.0012 **
## FirstAuthorFemale1   -0.1630     0.0756   -2.16   0.0315 *
## Year1997             -0.1337     0.3535   -0.38   0.7055
## Year1998              0.2409     0.2899    0.83   0.4065
## Year1999             -0.1119     0.2909   -0.38   0.7007
## Year2000             -0.0997     0.2929   -0.34   0.7337
## Year2001              0.0988     0.3162    0.31   0.7548
## Year2002             -0.0440     0.2958   -0.15   0.8817
## Year2003              0.0960     0.3433    0.28   0.7798
## Year2004              0.3943     0.2891    1.36   0.1734
## Year2005              0.4579     0.3033    1.51   0.1319
## Year2006              0.0879     0.2959    0.30   0.7666
```

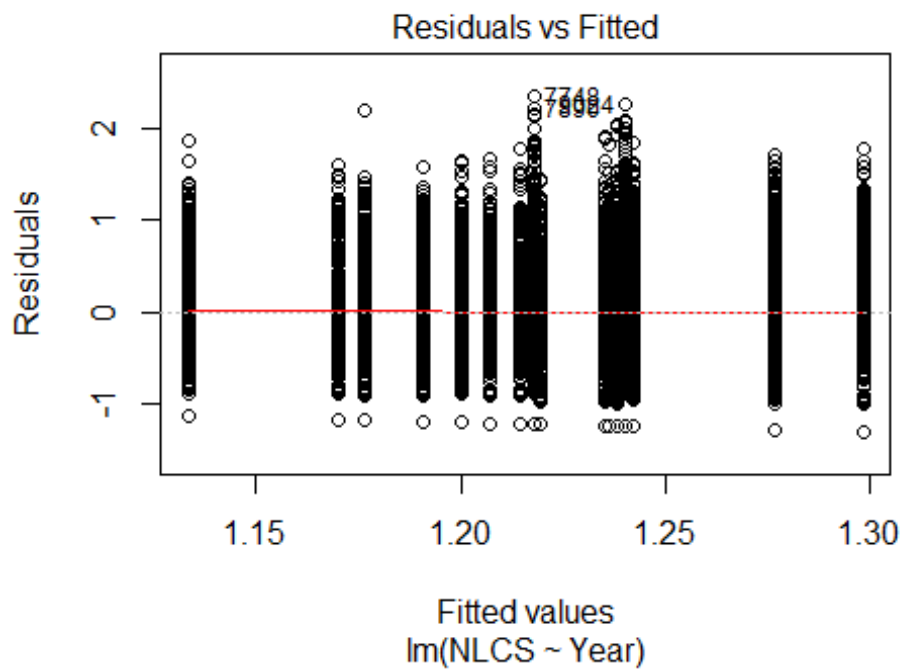
```
## Year2007            -0.0115     0.3016   -0.04   0.9696
## Year2008             0.1536     0.2725    0.56   0.5732
## Year2009            -0.1954     0.2686   -0.73   0.4673
## Year2010            -0.1093     0.2814   -0.39   0.6979
## Year2011            -0.1669     0.2764   -0.60   0.5463
## Year2012            -0.1437     0.2816   -0.51   0.6101
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.715
## Multiple R-squared:  0.0809, Adjusted R-squared:  0.0431
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
##  observation 230 is an outlier with |weight| = 0 ( < 0.00023);
##  44 weights are ~= 1. The remaining 386 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.302   0.883   0.945   0.904   0.977   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          2.32e-04          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##                   GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.198  1           1.095
## Year             1.198 16           1.006
```



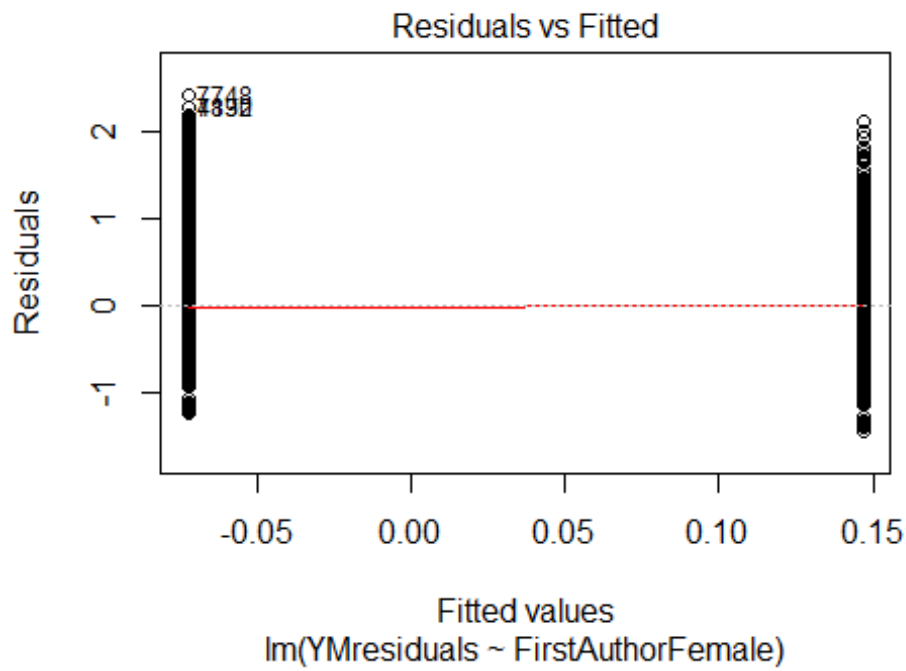

```
## [1] "List of  1 outliers with residuals above  2.5"
##        ScopusId NLCS Year OneField Fields residuals
## 315 84973631259 4.73 2007     1206      3      3.91
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.2724 -0.5346  0.0072  0.5106  3.9217
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         0.8249     0.2546    3.24   0.0013 **
## LastAuthorFemale1  -0.1396     0.0747   -1.87   0.0622 .
## Year1997           -0.1308     0.3518   -0.37   0.7102
## Year1998            0.2432     0.2897    0.84   0.4018
## Year1999           -0.0980     0.2943   -0.33   0.7393
## Year2000           -0.1087     0.2902   -0.37   0.7081
## Year2001            0.0994     0.3212    0.31   0.7571
## Year2002           -0.0574     0.2957   -0.19   0.8462
## Year2003            0.0963     0.3444    0.28   0.7799
## Year2004            0.3766     0.2872    1.31   0.1906
## Year2005            0.4476     0.3024    1.48   0.1396
## Year2006            0.0886     0.2939    0.30   0.7633
```

```
## Year2007           -0.0165     0.3000   -0.06   0.9561
## Year2008            0.1372     0.2713    0.51   0.6133
## Year2009           -0.1999     0.2691   -0.74   0.4580
## Year2010           -0.1181     0.2803   -0.42   0.6737
## Year2011           -0.1595     0.2780   -0.57   0.5665
## Year2012           -0.1506     0.2816   -0.53   0.5930
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.71
## Multiple R-squared:  0.0785, Adjusted R-squared:  0.0406
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
##  observation 230 is an outlier with |weight| = 0 ( < 0.00023);
##  44 weights are ~= 1. The remaining 386 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.285   0.885   0.943   0.903   0.977   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          2.32e-04          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                    psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Sample size for the above analysis:  431"
## [1] ""
## [1] ""
## [1] "####################################"
## [1] "Analysis of AJSC 1207"
## [1] "####################################"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  392  334  376  389  403  405  369  350  333  390  495  541  662  759  779
## 2011 2012
##  860  899
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  316  274  329  326  327  310  325  303  292  321  425  479  566  645  679
```

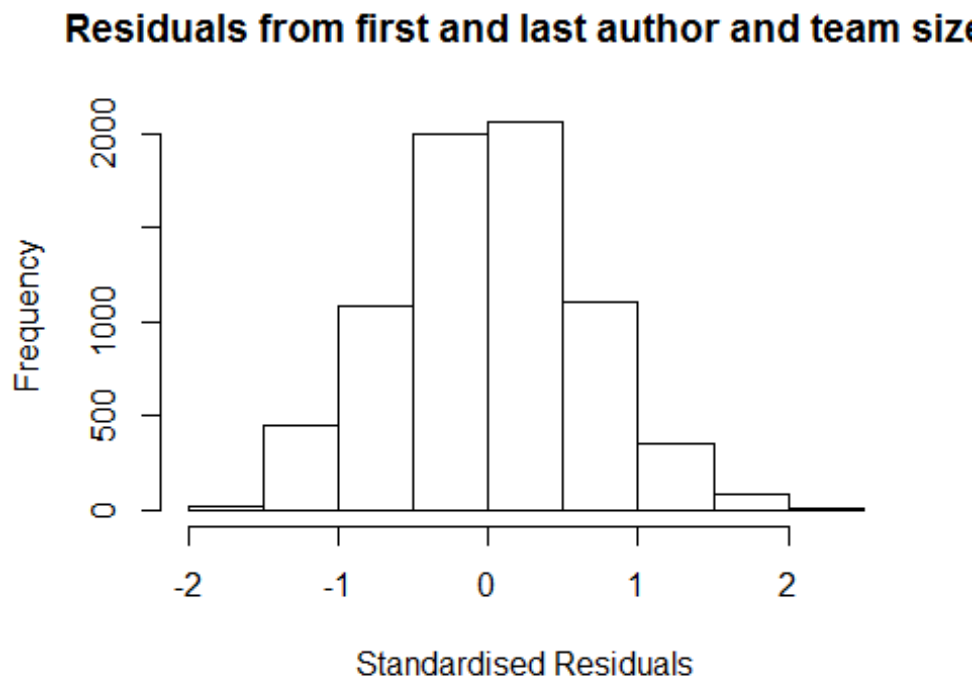
```
## 2011 2012
##  739  794
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  301  265  315  321  318  301  316  287  280  308  397  457  544  613  648
## 2011 2012
##  716  768
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
##  Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 170, df = 16, p-value <2e-16
```




```
##
##  Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 1.2, df = 1, p-value = 0.3
```

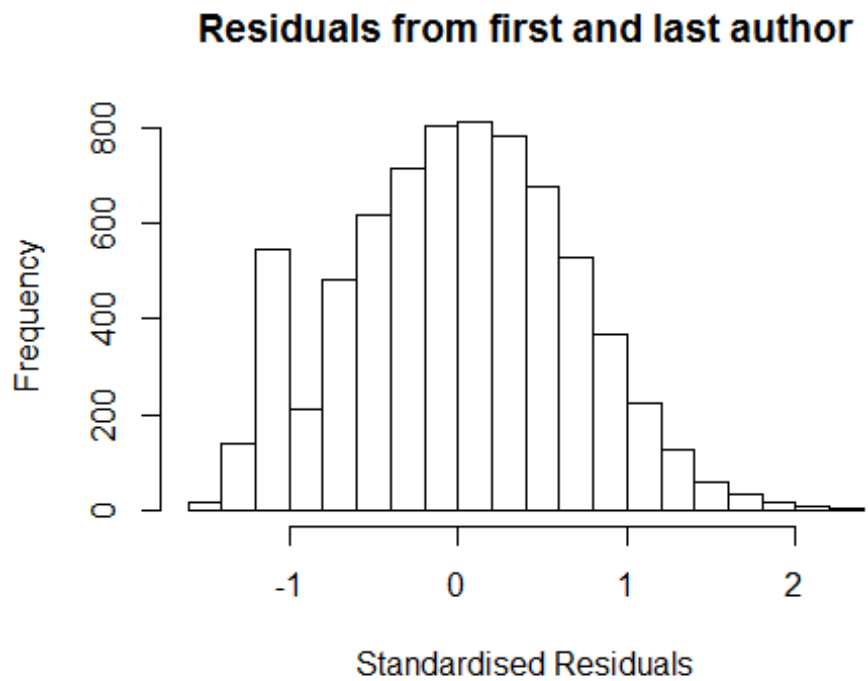



```
## [1] "Female first author team size 2018 geometric mean: 1.74136013073575"
## [1] "Male first author team size 2018 geometric mean: 1.29452486303445"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 80000, p-value = 9e-12
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 1.67125727591238"
## [1] "Male last author team size 2018 geometric mean: 1.33874834695544"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 74000, p-value = 4e-07
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##                     GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.819  1           1.349
## LastAuthorFemale  1.773  1           1.332
## UniqueAuthors     1.130  4           1.015
## Year              1.080 16           1.002
```



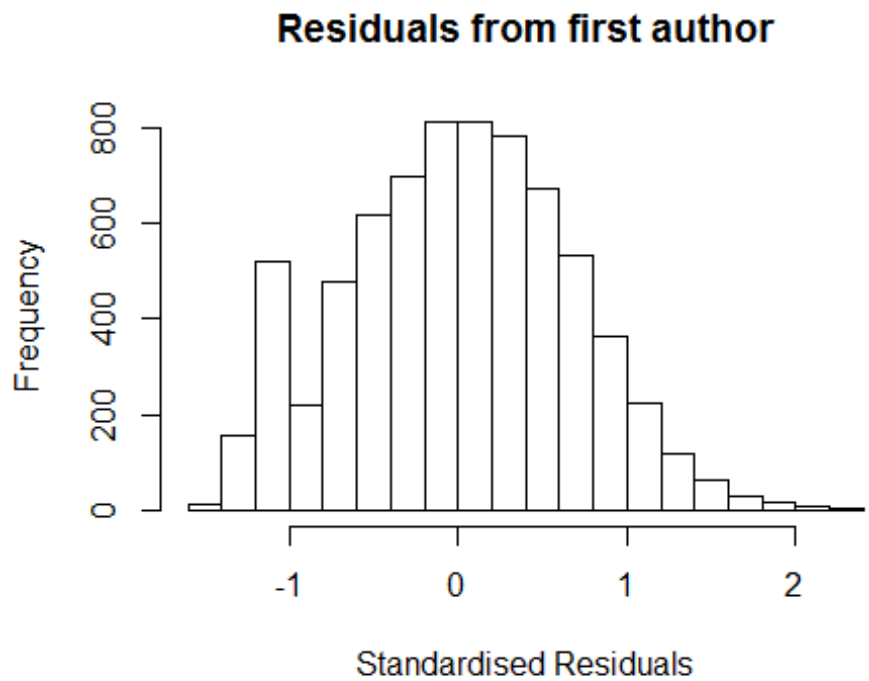

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
UniqueAuthors +
##     Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##     k.max = 1000))
##  \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.74141 -0.42473  0.00561  0.42793  2.45101
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)          1.0953     0.0355   30.86  < 2e-16 ***
## FirstAuthorFemale1   0.1039     0.0216    4.80  1.6e-06 ***
## LastAuthorFemale1    0.0452     0.0216    2.09  0.03629 *
## UniqueAuthors2       0.4322     0.0216   19.99  < 2e-16 ***
## UniqueAuthors3       0.5384     0.0260   20.71  < 2e-16 ***
## UniqueAuthors4       0.6235     0.0352   17.69  < 2e-16 ***
## UniqueAuthors5       0.6185     0.0321   19.28  < 2e-16 ***
## Year1997            -0.0180     0.0514   -0.35  0.72575
## Year1998            -0.0743     0.0485   -1.53  0.12550
## Year1999            -0.0306     0.0485   -0.63  0.52749
```

```
## Year2000            -0.0269     0.0497   -0.54  0.58896
## Year2001            -0.1119     0.0495   -2.26  0.02380 *
## Year2002            -0.1601     0.0512   -3.13  0.00176 **
## Year2003            -0.0922     0.0512   -1.80  0.07169 .
## Year2004            -0.0731     0.0492   -1.49  0.13748
## Year2005            -0.1105     0.0502   -2.20  0.02770 *
## Year2006            -0.1653     0.0451   -3.67  0.00025 ***
## Year2007            -0.1370     0.0448   -3.06  0.00223 **
## Year2008            -0.1047     0.0435   -2.41  0.01617 *
## Year2009            -0.1128     0.0424   -2.66  0.00786 **
## Year2010            -0.0414     0.0434   -0.95  0.34003
## Year2011            -0.0897     0.0454   -1.97  0.04834 *
## Year2012            -0.1037     0.0475   -2.18  0.02906 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.635
## Multiple R-squared:  0.14,   Adjusted R-squared:  0.137
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  629 weights are ~= 1. The remaining 6526 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.104   0.869   0.950   0.907   0.985   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.40e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 2.080  1           1.442
## LastAuthorFemale  2.062  1           1.436
## Year              1.026 16           1.001
```



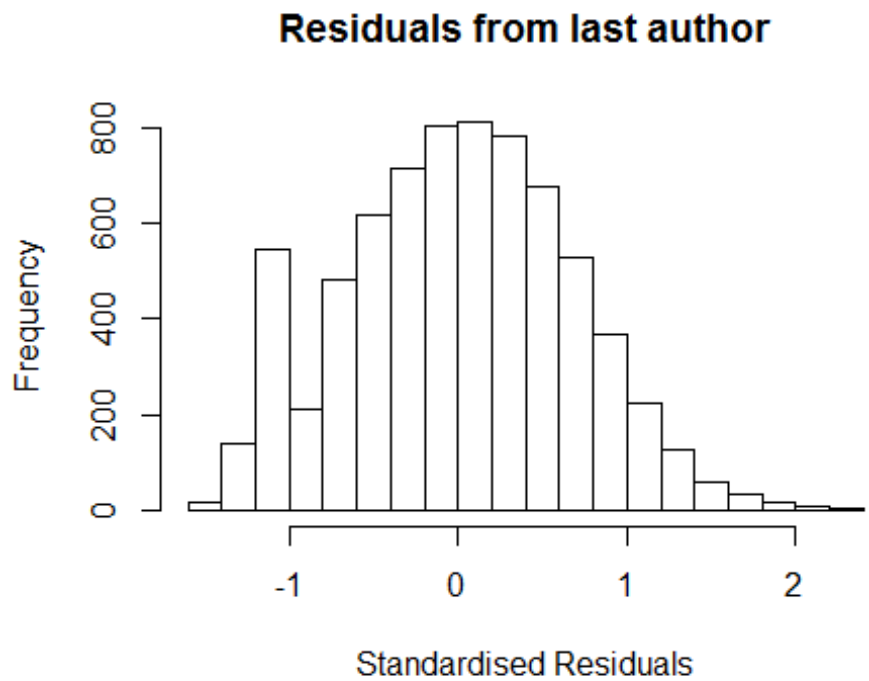

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.4313 -0.4634  0.0185  0.4632  2.3180
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)           1.1999     0.0383   31.34  < 2e-16 ***
## FirstAuthorFemale1    0.1955     0.0254    7.70  1.6e-14 ***
## LastAuthorFemale1     0.0548     0.0258    2.12   0.0337 *
## Year1997             -0.0415     0.0556   -0.75   0.4550
## Year1998             -0.0752     0.0518   -1.45   0.1469
## Year1999             -0.0583     0.0522   -1.12   0.2641
## Year2000             -0.0354     0.0517   -0.68   0.4943
## Year2001             -0.1093     0.0530   -2.06   0.0393 *
## Year2002             -0.1432     0.0546   -2.62   0.0087 **
## Year2003             -0.0579     0.0562   -1.03   0.3023
## Year2004             -0.0664     0.0538   -1.23   0.2173
## Year2005             -0.0644     0.0528   -1.22   0.2224
```

```
## Year2006            -0.1368     0.0490   -2.79   0.0053 **
## Year2007            -0.1057     0.0485   -2.18   0.0294 *
## Year2008            -0.0621     0.0474   -1.31   0.1906
## Year2009            -0.1045     0.0459   -2.28   0.0228 *
## Year2010            -0.0189     0.0471   -0.40   0.6888
## Year2011            -0.0844     0.0494   -1.71   0.0879 .
## Year2012            -0.0741     0.0515   -1.44   0.1503
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.689
## Multiple R-squared:  0.0286, Adjusted R-squared:  0.0262
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
##  573 weights are ~= 1. The remaining 6582 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.235   0.874   0.952   0.912   0.985   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.40e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.022  1           1.011
## Year              1.022 16           1.001
```




```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.4209 -0.4644  0.0189  0.4635  2.3146
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)           1.2032     0.0383   31.43   <2e-16 ***
## FirstAuthorFemale1    0.2353     0.0177   13.26   <2e-16 ***
## Year1997             -0.0391     0.0557   -0.70   0.4833
## Year1998             -0.0737     0.0519   -1.42   0.1557
## Year1999             -0.0567     0.0522   -1.09   0.2772
## Year2000             -0.0344     0.0519   -0.66   0.5069
## Year2001             -0.1090     0.0530   -2.06   0.0398 *
## Year2002             -0.1418     0.0547   -2.59   0.0095 **
## Year2003             -0.0588     0.0562   -1.05   0.2955
## Year2004             -0.0675     0.0537   -1.26   0.2093
## Year2005             -0.0641     0.0528   -1.22   0.2244
## Year2006             -0.1367     0.0491   -2.78   0.0054 **
```

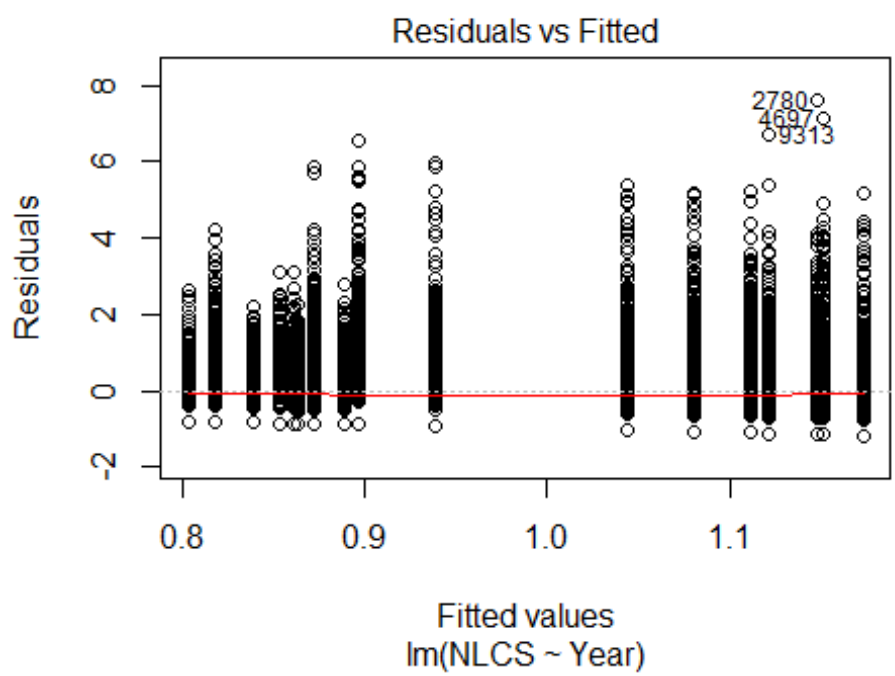
```
## Year2007            -0.1032     0.0485   -2.13   0.0334 *
## Year2008            -0.0610     0.0474   -1.29   0.1985
## Year2009            -0.1037     0.0459   -2.26   0.0239 *
## Year2010            -0.0176     0.0472   -0.37   0.7091
## Year2011            -0.0842     0.0495   -1.70   0.0891 .
## Year2012            -0.0734     0.0515   -1.43   0.1541
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.69
## Multiple R-squared:  0.0279, Adjusted R-squared:  0.0256
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  583 weights are ~= 1. The remaining 6572 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.238   0.875   0.952   0.912   0.985   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.40e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##                   GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.014  1           1.007
## Year             1.014 16           1.000
```



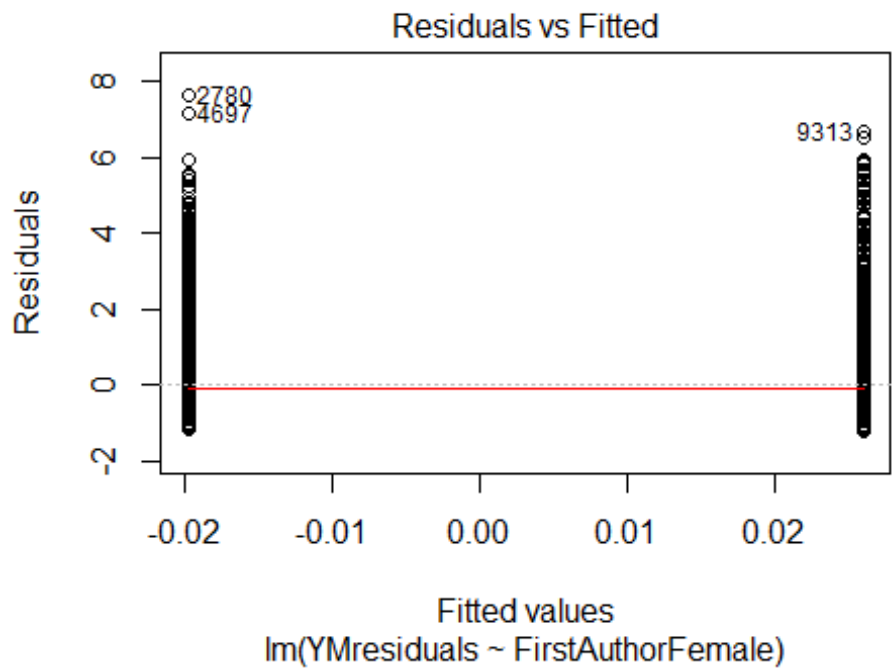

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.4011 -0.4664  0.0189  0.4670  2.2941
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.2152     0.0388   31.29   <2e-16 ***
## LastAuthorFemale1   0.1996     0.0179   11.16   <2e-16 ***
## Year1997           -0.0470     0.0555   -0.85   0.3973
## Year1998           -0.0786     0.0520   -1.51   0.1310
## Year1999           -0.0613     0.0526   -1.17   0.2439
## Year2000           -0.0360     0.0521   -0.69   0.4901
## Year2001           -0.1108     0.0535   -2.07   0.0383 *
## Year2002           -0.1472     0.0549   -2.68   0.0074 **
## Year2003           -0.0481     0.0565   -0.85   0.3951
## Year2004           -0.0581     0.0548   -1.06   0.2896
## Year2005           -0.0569     0.0533   -1.07   0.2855
## Year2006           -0.1259     0.0496   -2.54   0.0111 *
```

```
## Year2007           -0.1072     0.0492   -2.18   0.0295 *
## Year2008           -0.0585     0.0480   -1.22   0.2226
## Year2009           -0.1003     0.0465   -2.16   0.0309 *
## Year2010           -0.0138     0.0476   -0.29   0.7726
## Year2011           -0.0813     0.0499   -1.63   0.1031
## Year2012           -0.0701     0.0521   -1.35   0.1784
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.692
## Multiple R-squared:  0.0206, Adjusted R-squared:  0.0183
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
##  564 weights are ~= 1. The remaining 6591 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.249   0.874   0.951   0.912   0.986   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.40e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Sample size for the above analysis:  7155"
## [1] ""
## [1] ""
## [1] "####################################"
## [1] "Analysis of AJSC 1208"
## [1] "####################################"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  344  409  439  451  406  483 1320 1389 1419 1565 1496 1612 1649 1176 1051
## 2011 2012
## 1482 2104
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  308  344  382  396  372  431 1173 1230 1246 1379 1336 1429 1435 1033  924
## 2011 2012
```

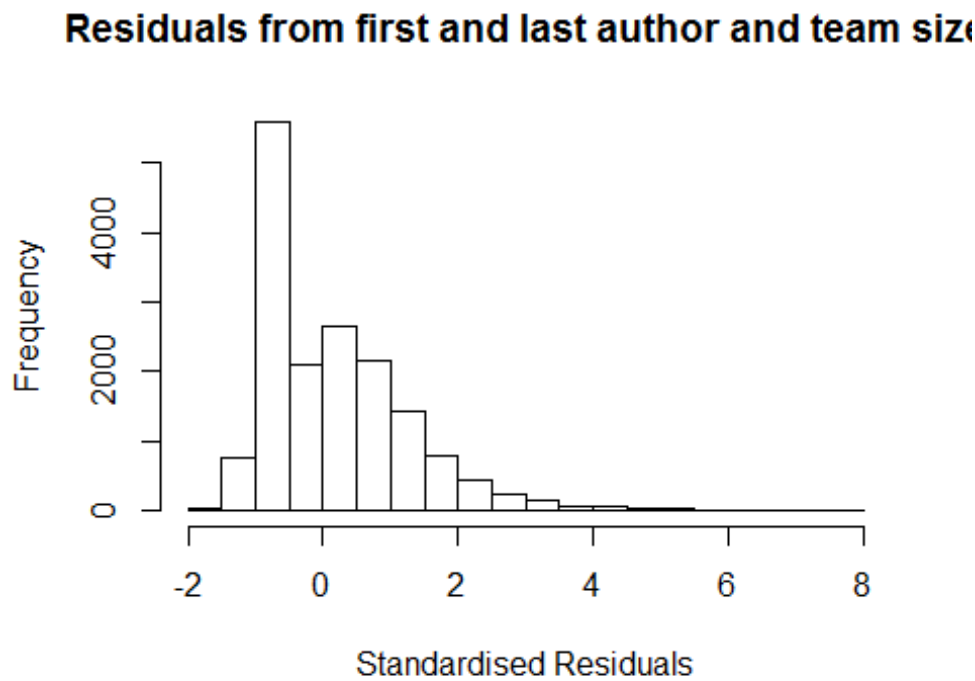
```
## 1289 1870
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  307  342  381  395  370  427 1173 1225 1241 1376 1325 1420 1426 1029  916
## 2011 2012
## 1282 1859
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
##  Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 470, df = 16, p-value <2e-16
```




```
##
##  Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 1.7, df = 1, p-value = 0.2
```




```
## [1] "Female first author team size 2018 geometric mean: 1.06592648525968"
## [1] "Male first author team size 2018 geometric mean: 1.05585819545849"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 260000, p-value = 0.5
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 1.06929617767031"
## [1] "Male last author team size 2018 geometric mean: 1.05266151708374"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 260000, p-value = 0.3
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##                     GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 20.022  1           4.475
## LastAuthorFemale  20.075  1           4.481
## UniqueAuthors      1.065  4           1.008
## Year               1.027 16           1.001
```




```
## [1] "List of  516 outliers with residuals above  2.5"
##          ScopusId  NLCS Year OneField Fields residuals
## 94    67651037634 3.400 1996     1208      1     2.637
## 138   61449406492 3.938 1996     1208      1     3.175
## 140   61949400015 3.345 1996     1208      1     2.527
## 376    0009431359 3.532 1997     1208      2     2.700
## 412   79959433454 3.987 1997     1208      2     3.209
## 952    0000490567 3.695 1998     1208      2     2.895
## 1207  60949584988 3.410 1999     1208      1     2.682
## 1485   0013039289 3.420 1999     1208      3     2.638
## 2557  10844245985 5.274 2002     1208      1     4.301
## 2562  33750620950 3.855 2002     1208      1     2.827
## 2563  33947361251 3.629 2002     1208      1     2.601
## 2565  34547299083 3.746 2002     1208      1     2.773
## 2566  47749108055 4.222 2002     1208      1     3.249
## 2581  60949458107 4.222 2002     1208      1     3.249
## 2605  60950002681 3.502 2002     1208      1     2.529
## 2607  60950027813 3.746 2002     1208      1     2.773
## 2619  60950346555 4.918 2002     1208      1     3.945
## 2732  65849154249 3.629 2002     1208      1     2.601
## 2733  65849216639 3.629 2002     1208      1     2.656
## 2734  65849259075 3.629 2002     1208      1     2.601
## 2736  65849326957 3.955 2002     1208      1     2.927
## 2748  67650096797 3.746 2002     1208      1     2.773
## 2751  70249100469 4.301 2002     1208      1     3.328
## 2752  70349438078 3.502 2002     1208      1     2.529
## 2780   3142697241 8.757 2002     1208      2     7.784
```

```
## 2781  33646911853 5.146 2002     1208      2     4.118
## 2783  33748905673 5.092 2002     1208      2     4.119
## 2789  44149119239 3.798 2002     1208      2     2.770
## 2825  60949601992 4.417 2002     1208      2     3.389
## 2860  60950259244 4.417 2002     1208      2     3.389
## 3194  61949270383 3.955 2002     1208      1     2.982
## 3322  34248720247 3.855 2002     1208      1     2.882
## 3323  34248734400 3.855 2002     1208      1     2.882
## 3326  38749133094 4.050 2002     1208      1     3.077
## 3327  58449103137 3.502 2002     1208      1     2.529
## 3332  60949291781 3.629 2002     1208      1     2.601
## 3342  60950305286 3.746 2002     1208      1     2.773
## 3345  60950448412 3.629 2002     1208      1     2.656
## 3347  60950498847 3.746 2002     1208      1     2.773
## 3348  60950501375 3.955 2002     1208      1     2.927
## 3352  60950608189 3.955 2002     1208      1     2.927
## 3354  60950639390 3.855 2002     1208      1     2.827
## 3355  60950641233 4.515 2002     1208      1     3.487
## 3358  60950680913 3.629 2002     1208      1     2.656
## 3361  60950699119 3.855 2002     1208      1     2.827
## 3380  61249427664 3.502 2002     1208      1     2.529
## 3382  61249520644 3.955 2002     1208      1     2.927
## 3439  70349348163 4.050 2002     1208      1     3.077
## 3440  70349378641 3.629 2002     1208      1     2.601
## 3460  70450000901 3.629 2002     1208      1     2.656
## 3504  61249583521 3.963 2002     1208      2     2.990
## 3509  61449442102 3.533 2002     1208      2     2.560
## 3515  62449109752 4.033 2002     1208      2     3.005
## 3516  62449174089 3.672 2002     1208      2     2.699
## 3517  62449296550 4.063 2002     1208      2     3.035
## 3657  60950279634 3.963 2002     1208      3     2.935
## 3793  85012444559 3.746 2002     1202      3     2.773
## 3843   0036526749 3.670 2002     1202      6     2.642
## 3845   1442335800 3.694 2002     1202      6     2.721
## 3884  34047093694 4.436 2003     1208      1     3.407
## 3887  46749133770 4.109 2003     1208      1     3.134
## 3912  60949464452 3.780 2003     1208      1     2.805
## 3918  60949658016 5.148 2003     1208      1     4.119
## 3921  60949700952 4.436 2003     1208      1     3.407
## 3960  60950414028 4.651 2003     1208      1     3.676
## 3963  60950438763 4.034 2003     1208      1     3.059
## 3977  60950500372 4.249 2003     1208      1     3.274
## 3993  60950588015 4.314 2003     1208      1     3.339
## 3999  60950625661 5.403 2003     1208      1     4.428
## 4036  61249262249 4.034 2003     1208      1     3.059
## 4094  70449761302 3.580 2003     1208      1     2.605
## 4141   8544244308 4.547 2003     1208      1     3.572
## 4142  26944502435 4.044 2003     1208      2     3.015
## 4143  33646681740 6.083 2003     1208      2     5.054
## 4146  34547600256 4.341 2003     1208      2     3.312
```

```
## 4197  60950066571 4.869 2003     1208      2     3.840
## 4230  61149570569 3.528 2003     1208      2     2.553
## 4489  45449108735 5.113 2003     1208      1     4.084
## 4583  61949186737 4.181 2003     1208      1     3.206
## 4600  34748861400 4.651 2003     1208      1     3.622
## 4621  41549119242 4.109 2003     1208      1     3.134
## 4691  36248970780 3.954 2003     1208      1     2.925
## 4697  60949457140 8.312 2003     1208      1     7.337
## 4710  60950639325 3.780 2003     1208      1     2.751
## 4713  60950694241 3.954 2003     1208      1     2.979
## 4736  61249635138 4.109 2003     1208      1     3.080
## 4742  61449197953 4.922 2003     1208      1     3.947
## 4747  61449337582 4.837 2003     1208      1     3.808
## 4782  64949161384 3.580 2003     1208      1     2.551
## 4822  70449843968 3.683 2003     1208      1     2.654
## 4863  29744456134 5.218 2003     1208      2     4.243
## 4872  34547980224 3.959 2003     1208      2     2.984
## 4873  48549096636 5.652 2003     1208      2     4.623
## 4888  61249716500 4.045 2003     1208      2     3.070
## 4990  70449856706 3.528 2003     1208      2     2.553
## 5092  61949103780 3.552 2003     1208      3     2.523
## 5274  28244501538 5.482 2004     1208      1     4.078
## 5280  34748922675 4.292 2004     1208      1     3.198
## 5307  60949771564 3.622 2004     1208      1     2.583
## 5445  61949116307 3.947 2004     1208      1     2.908
## 5463  67049083067 4.804 2004     1208      1     3.765
## 5468  67649743610 4.213 2004     1208      1     3.119
## 5514  34250312985 4.933 2004     1208      2     3.839
## 5562  60950391613 4.521 2004     1208      2     3.427
## 5581  60950719934 3.827 2004     1208      2     2.788
## 6108  33750435368 5.377 2004     1208      1     4.338
## 6112  33947671157 4.506 2004     1208      1     3.412
## 6126  60949382516 6.369 2004     1208      1     5.330
## 6152  61049449507 5.223 2004     1208      1     4.184
## 6178  61449455372 3.739 2004     1208      1     2.700
## 6196  62549140156 3.622 2004     1208      1     2.583
## 6270  71249085963 3.622 2004     1208      1     2.583
## 6285  60950338479 5.580 2004     1208      2     4.486
## 6299  61849165077 3.827 2004     1208      2     2.733
## 6350  70449786556 4.699 2004     1208      2     3.660
## 6352  70449824679 4.933 2004     1208      2     3.894
## 6699  34248708340 3.945 2005     1208      1     3.012
## 6705  34548494655 3.822 2005     1208      1     2.834
## 6729  60949655783 3.688 2005     1208      1     2.755
## 6765  60950409349 3.945 2005     1208      1     3.012
## 6790  60950596472 3.541 2005     1208      1     2.553
## 6800  60950708929 3.541 2005     1208      1     2.608
## 6850  61149580054 4.683 2005     1208      1     3.330
## 6901  65449143734 4.358 2005     1208      1     3.370
## 6902  65449165806 4.683 2005     1208      1     3.695
```

```
## 6909  67649784613 4.445 2005     1208      1     3.512
## 6947  70449899952 3.822 2005     1208      1     2.889
## 6955  70449970464 3.541 2005     1208      1     2.608
## 6966  70450058332 3.688 2005     1208      1     2.755
## 6985  33750587792 3.538 2005     1208      2     2.550
## 6989  34250008387 4.562 2005     1208      2     3.629
## 6994  38649134273 6.094 2005     1208      2     5.106
## 6999  60949254410 4.465 2005     1208      2     3.532
## 7041  60950274865 3.944 2005     1208      2     3.011
## 7043  60950350045 5.477 2005     1208      2     4.544
## 7070  60950562957 3.861 2005     1208      2     2.928
## 7078  60950642207 3.861 2005     1208      2     2.928
## 7177  33845250551 4.010 2005     1202      4     3.077
## 7380  62749191854 3.744 2005     1202      3     2.756
## 7383  63849109891 3.549 2005     1202      3     2.561
## 7553  60950504855 3.650 2005     1208      3     2.662
## 7604  61949241615 3.822 2005     1208      1     2.834
## 7630  61049327976 3.822 2005     1208      1     2.834
## 7650  21544433512 4.003 2005     1208      2     2.868
## 7723  33846337496 6.348 2005     1208      1     5.360
## 7724  46249120022 4.059 2005     1208      1     3.071
## 7725  46249120887 3.945 2005     1208      1     3.012
## 7734  60949594699 3.822 2005     1208      1     2.834
## 7742  60950233081 4.059 2005     1208      1     3.126
## 7743  60950413511 3.541 2005     1208      1     2.608
## 7755  60950658414 3.541 2005     1208      1     2.553
## 7804  61449502250 3.688 2005     1208      1     2.700
## 7832  67649792074 3.822 2005     1208      1     2.889
## 7837  67650088057 3.541 2005     1208      1     2.608
## 7844  67650178186 4.059 2005     1208      1     3.126
## 7845  68549117379 3.822 2005     1208      1     2.889
## 7868  70450064684 3.822 2005     1208      1     2.889
## 7876  70450093197 3.541 2005     1208      1     2.608
## 7898  44449106909 4.250 2005     1208      2     3.262
## 7902  60950423747 5.118 2005     1208      2     4.130
## 7914  61249595569 3.788 2005     1208      2     2.855
## 7915  61449306806 3.654 2005     1208      2     2.519
## 7919  61449561847 3.582 2005     1208      2     2.594
## 7937  62449266342 3.723 2005     1208      2     2.790
## 8040  33846333428 3.605 2005     1208      3     2.672
## 8313  50949119108 4.710 2006     1208      1     3.773
## 8336  60949594614 3.653 2006     1208      1     2.661
## 8347  60950223477 3.786 2006     1208      1     2.794
## 8356  60950442490 3.908 2006     1208      1     2.971
## 8368  60950493268 4.224 2006     1208      1     3.287
## 8369  60950493735 3.653 2006     1208      1     2.661
## 8373  60950504606 3.786 2006     1208      1     2.794
## 8394  60950726337 4.317 2006     1208      1     3.325
## 8444  61449369584 3.508 2006     1208      1     2.571
## 8484  67650109778 3.786 2006     1208      1     2.794
```

```
## 8486  67650128316 3.653 2006     1208      1     2.716
## 8487  67650135684 3.786 2006     1208      1     2.849
## 8490  67650164706 4.317 2006     1208      1     3.325
## 8493  68849095704 3.508 2006     1208      1     2.516
## 8529  70450105045 4.224 2006     1208      1     3.287
## 9045  60950518936 3.508 2006     1208      1     2.516
## 9237  61949128551 4.404 2006     1208      1     3.467
## 9296  84870103615 4.317 2006     1208      1     3.380
## 9311  42449161018 5.130 2006     1208      1     4.138
## 9313  53549130671 7.825 2006     1208      1     6.833
## 9314  55349128031 5.280 2006     1208      1     4.288
## 9325  60950413512 4.777 2006     1208      1     3.785
## 9326  60950445899 3.908 2006     1208      1     2.916
## 9327  60950622690 4.126 2006     1208      1     3.189
## 9342  61249099688 4.021 2006     1208      1     3.029
## 9344  61249305083 4.317 2006     1208      1     3.380
## 9351  61449427822 3.653 2006     1208      1     2.661
## 9352  61449427823 4.126 2006     1208      1     3.134
## 9361  61949197756 4.021 2006     1208      1     3.084
## 9365  62349122821 3.653 2006     1208      1     2.661
## 9367  62449189059 4.777 2006     1208      1     3.840
## 9371  62749148771 4.777 2006     1208      1     3.840
## 9381  64949139647 4.224 2006     1208      1     3.232
## 9426  67650079640 3.908 2006     1208      1     2.916
## 9440  70449824307 3.786 2006     1208      1     2.849
## 9460  70450001260 3.508 2006     1208      1     2.516
## 9468  70450045620 4.126 2006     1208      1     3.134
## 9469  70450058396 3.786 2006     1208      1     2.849
## 9479  70749096336 4.021 2006     1208      1     3.029
## 9491  33846152354 3.582 2006     1208      2     2.645
## 9497  34249084654 5.314 2006     1208      2     4.322
## 9503  44449151884 6.523 2006     1208      2     5.531
## 9504  60949270586 3.627 2006     1208      2     2.690
## 9524  61449282187 3.627 2006     1208      2     2.690
## 9541  61449530625 6.523 2006     1208      2     5.586
## 9574  70349296952 3.627 2006     1208      2     2.690
## 9577  70449775017 4.234 2006     1208      2     3.297
## 9589  70450078889 3.627 2006     1208      2     2.635
## 9803  57749157632 6.254 2007     1208      1     5.292
## 9867  60950549890 3.610 2007     1208      1     2.648
## 9899  61149182825 4.555 2007     1208      1     3.648
## 9949  62449306099 4.555 2007     1208      1     3.648
## 9978  67650155463 3.783 2007     1208      1     2.821
## 9981  67650162724 3.610 2007     1208      1     2.648
## 9982  67650168953 3.939 2007     1208      1     2.977
## 10006 70449930640 3.610 2007     1208      1     2.703
## 10040 84868448556 3.610 2007     1208      1     2.703
## 10075 43249135021 4.018 2007     1208      2     3.056
## 10078 43249146684 4.726 2007     1208      2     3.764
## 10107 60949438530 4.726 2007     1208      2     3.819
```

```
## 10155 60950643835 4.839 2007     1208      2     3.932
## 10158 60950662433 3.834 2007     1208      2     2.927
## 10415 61949141429 3.441 2007     1208      3     2.534
## 10670 61049446035 5.641 2007     1208      1     4.734
## 10676 70450032239 3.610 2007     1208      1     2.703
## 10800 61049441123 3.416 2007     1208      1     2.509
## 10809 70450064287 5.151 2007     1208      1     4.244
## 10853 34147141318 3.409 2007     1208      2     2.502
## 10866 60949326226 3.610 2007     1208      1     2.648
## 10867 60949370742 5.976 2007     1208      1     5.069
## 10880 60950131470 3.939 2007     1208      1     2.977
## 10891 60950529591 3.416 2007     1208      1     2.509
## 10892 60950536451 4.449 2007     1208      1     3.487
## 10925 62449170699 5.841 2007     1208      1     4.879
## 10968 67649683064 4.214 2007     1208      1     3.307
## 10984 67650096553 4.082 2007     1208      1     3.175
## 10985 67650143129 5.474 2007     1208      1     4.567
## 10986 67650925409 3.939 2007     1208      1     3.032
## 10988 70249093068 3.610 2007     1208      1     2.703
## 10992 70449859265 3.610 2007     1208      1     2.703
## 10994 70449877375 3.610 2007     1208      1     2.703
## 11007 70450015708 6.179 2007     1208      1     5.272
## 11014 70450064757 3.783 2007     1208      1     2.821
## 11015 70450077380 4.654 2007     1208      1     3.747
## 11112 62449186372 3.857 2007     1208      2     2.895
## 11165 70450078906 4.018 2007     1208      2     3.111
## 11217 61249239835 4.188 2007     1208      3     3.281
## 11228 61249660537 3.741 2007     1208      3     2.834
## 11409 60949485276 3.367 2008     1208      1     2.522
## 11434 60950624831 3.802 2008     1208      1     2.957
## 11478 60950719981 5.997 2008     1208      1     5.098
## 11532 60950299886 3.598 2008     1208      1     2.753
## 11540 57749125280 4.798 2008     1208      1     3.953
## 11548 60949270646 4.686 2008     1208      1     3.841
## 11573 61049554358 3.802 2008     1208      1     2.957
## 11605 61949182548 3.598 2008     1208      1     2.699
## 11611 61949371494 3.367 2008     1208      1     2.522
## 11618 62449098267 5.426 2008     1208      1     4.581
## 11624 62449159076 3.598 2008     1208      1     2.753
## 11626 62449175332 5.942 2008     1208      1     5.043
## 11633 62449269964 3.984 2008     1208      1     3.139
## 11641 62949147900 4.300 2008     1208      1     3.401
## 11644 62949210835 3.367 2008     1208      1     2.522
## 11672 67650099391 3.984 2008     1208      1     3.085
## 11676 67650112946 3.984 2008     1208      1     3.085
## 11681 67650132635 4.149 2008     1208      1     3.250
## 11684 67650151239 3.984 2008     1208      1     3.085
## 11687 67650155476 3.367 2008     1208      1     2.522
## 11688 67650162729 4.798 2008     1208      1     3.899
## 11691 67650178260 5.001 2008     1208      1     4.102
```

```
## 11695 70449752752 4.300 2008     1208      1     3.455
## 11696 70449780196 3.367 2008     1208      1     2.522
## 11723 70450068192 3.367 2008     1208      1     2.522
## 11792 61149209901 3.427 2008     1208      2     2.582
## 11796 61149472770 4.222 2008     1208      2     3.323
## 12102 65749261392 3.585 2008     1202      3     2.686
## 12226 70450003031 5.184 2008     1208      1     4.285
## 12362 70450082589 3.516 2008     1205      3     2.617
## 12399 45349087542 3.582 2008     1208      2     2.737
## 12435 61249636505 6.426 2008     1208      1     5.581
## 12468 60949164169 3.720 2008     1208      3     2.875
## 12498 70449866030 3.802 2008     1208      1     2.903
## 12500 70450060571 4.686 2008     1208      1     3.787
## 12551 61049119505 3.367 2008     1208      1     2.522
## 12629 57749127836 3.984 2008     1208      1     3.139
## 12658 61249639675 3.598 2008     1208      1     2.699
## 12661 61449513631 3.984 2008     1208      1     3.139
## 12664 61949197772 4.149 2008     1208      1     3.304
## 12671 62449262219 4.300 2008     1208      1     3.455
## 12675 62749110813 6.152 2008     1208      1     5.307
## 12677 62749122811 5.268 2008     1208      1     4.423
## 12679 62749136269 3.984 2008     1208      1     3.139
## 12682 62749151909 4.798 2008     1208      1     3.953
## 12683 62749155048 4.149 2008     1208      1     3.250
## 12684 62749156172 5.942 2008     1208      1     5.097
## 12691 64949093382 3.598 2008     1208      1     2.699
## 12696 64949107038 3.802 2008     1208      1     2.903
## 12700 64949153413 4.300 2008     1208      1     3.455
## 12727 67650070793 4.149 2008     1208      1     3.304
## 12735 67650090943 5.570 2008     1208      1     4.725
## 12739 70449931160 5.268 2008     1208      1     4.369
## 12746 78650710041 3.802 2008     1208      1     2.903
## 12795 61249468411 4.222 2008     1208      2     3.377
## 12797 61249548069 3.427 2008     1208      2     2.582
## 12861 40449115608 3.564 2008     1202      4     2.517
## 13002 85014946308 3.363 2008     1203      4     2.518
## 13097 77950310142 4.054 2009     1208      1     3.314
## 13214 84884186648 4.054 2009     1208      1     3.368
## 13249 61949462044 3.777 2009     1208      3     3.091
## 13400 73949133087 4.425 2009     1208      1     3.739
## 13418 70349840389 4.249 2009     1208      1     3.563
## 13474 68549123389 3.591 2009     1208      1     2.905
## 13559 73649148390 3.360 2009     1202      4     2.674
## 13613 67649716605 3.591 2009     1208      1     2.851
## 13663 67649950267 4.054 2009     1208      1     3.368
## 13679 77950325600 4.997 2009     1208      1     4.311
## 13680 77950788847 3.306 2009     1208      1     2.620
## 13683 77950809870 4.054 2009     1208      1     2.786
## 13684 77950815002 6.760 2009     1208      1     6.020
## 13685 77950822756 5.116 2009     1208      1     4.376
```

```
## 13686 77950831009 4.585 2009     1208      1     3.899
## 13688 77950841010 6.560 2009     1208      1     5.820
## 13693 68949083793 3.825 2009     1208      2     3.139
## 13695 68949090683 4.540 2009     1208      2     3.489
## 13697 68949092620 3.727 2009     1208      2     3.041
## 13700 68949123188 3.423 2009     1208      2     2.737
## 13702 68949135917 3.636 2009     1208      2     2.950
## 13893 79958822732 4.054 2009     1208      1     3.314
## 13895 80054949947 3.837 2009     1208      1     3.097
## 13939 77950229849 3.223 2009     1208      2     2.537
## 13940 77950253669 3.223 2009     1208      2     2.537
## 14020 65849111025 3.591 2009     1208      1     2.851
## 14021 65849141252 3.306 2009     1208      1     2.620
## 14039 77956351978 3.837 2009     1208      1     3.097
## 14092 77950181710 4.785 2009     1205      2     4.045
## 14275 77950904394 3.302 2010     1208      1     2.650
## 14278 78649401790 3.586 2010     1208      1     2.880
## 14282 78649430807 3.586 2010     1208      1     2.880
## 14293 84855406055 3.832 2010     1208      1     3.126
## 14542 77957559973 4.049 2010     1208      1     3.397
## 14543 77957584700 3.302 2010     1208      1     2.596
## 14575 78649656075 4.243 2010     1208      1     3.537
## 14591 77956331789 4.376 2010     1208      2     3.724
## 14712 79959202296 3.634 2010     1208      2     2.982
## 14810 77950211324 3.302 2010     1208      1     2.650
## 14880 79956372974 3.668 2010     1202      4     3.016
## 14898 77949582311 4.243 2010     1208      1     3.537
## 14902 77949615907 3.586 2010     1208      1     2.934
## 14927 77950140620 3.832 2010     1208      1     3.180
## 14928 77950253361 3.302 2010     1208      1     2.650
## 14934 77950886825 3.586 2010     1208      1     2.880
## 14935 77950888966 3.586 2010     1208      1     2.880
## 15039 76349106441 3.586 2010     1208      1     2.880
## 15050 77954443733 3.302 2010     1208      1     2.596
## 15094 77950287584 3.586 2010     1208      1     2.934
## 15120 79955418712 4.243 2010     1208      1     3.537
## 15143 73249128238 4.062 2010     1208      2     2.704
## 15226 77956500865 4.776 2010     1208      3     3.255
## 15228 78649472664 5.064 2010     1208      3     3.253
## 15279 82455228764 3.393 2010     1208      3     2.741
## 15298 77951908318 3.254 2010     1202      6     2.602
## 15358 84855209572 3.491 2011     1208      1     2.778
## 15359 84855210092 3.491 2011     1208      1     2.724
## 15365 84855762177 3.887 2011     1208      1     3.174
## 15368 84855797090 3.491 2011     1208      1     2.778
## 15380 84859358430 3.491 2011     1208      1     2.778
## 15382 84860778742 5.724 2011     1208      1     5.011
## 15403 84870329240 3.491 2011     1208      1     2.778
## 15411 84872712403 3.491 2011     1208      1     2.724
## 15419 84873283995 3.491 2011     1208      1     2.724
```

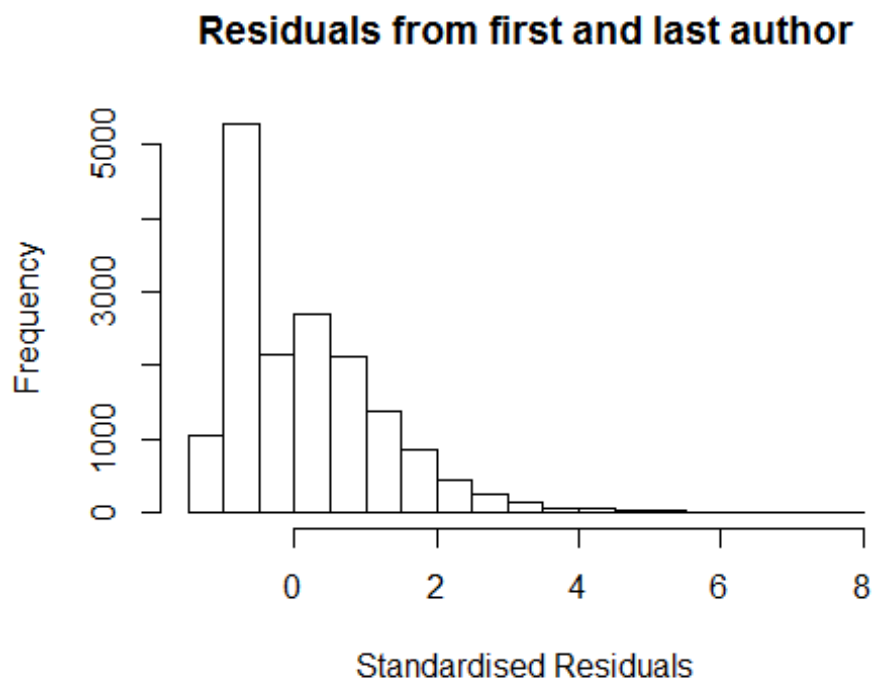
```
## 15424 84873302367 3.887 2011     1208      1     3.174
## 15447 84884198571 3.887 2011     1208      1     3.174
## 15465 84887967106 3.491 2011     1208      1     2.778
## 15474 84855781328 3.573 2011     1208      2     2.860
## 15475 84855791481 5.406 2011     1208      2     4.639
## 15674 80155194058 3.906 2011     1202      5     3.139
## 15694 80054789940 4.221 2011     1208      1     3.454
## 15698 80054708194 3.887 2011     1208      1     3.174
## 15717 80053615731 4.511 2011     1208      1     3.798
## 15803 80052975515 3.243 2011     1205      2     2.530
## 15826 80052542479 3.887 2011     1208      1     3.174
## 15845 84055192224 6.801 2011     1208      1     6.034
## 15847 84856196552 3.887 2011     1208      1     2.809
## 15853 84862322836 3.887 2011     1208      1     3.120
## 15869 84883480879 4.995 2011     1208      1     4.228
## 15899 84861070464 5.091 2011     1208      2     4.378
## 16068 79959952724 4.221 2011     1208      1     3.454
## 16069 79959977747 3.491 2011     1208      1     2.724
## 16100 82755175921 3.270 2011     1208      2     2.503
## 16212 79955377318 6.894 2011     1208      1     6.127
## 16216 79955439141 4.221 2011     1208      1     3.508
## 16219 79957888983 3.887 2011     1208      1     3.174
## 16290 79958846842 3.887 2011     1208      1     2.809
## 16360 79955674091 3.677 2011     1203      6     2.910
## 16379 79955677215 6.146 2011     1208      1     5.379
## 16383 80054915853 3.887 2011     1208      1     3.120
## 16498 79955368617 3.375 2011     1203      3     2.608
## 16516 78650843435 5.564 2011     1208      1     4.851
## 16517 78650845690 4.221 2011     1208      1     3.508
## 16541 79960860760 4.221 2011     1208      1     3.454
## 16542 79960871745 4.221 2011     1208      1     3.508
## 16550 81355124072 3.887 2011     1208      1     3.120
## 16567 84055198510 3.887 2011     1208      1     3.120
## 16586 85011484631 3.887 2011     1208      1     3.120
## 16615 80555156002 3.648 2011     1208      2     2.881
## 16620 82455186258 3.290 2011     1208      2     2.523
## 16630 84555179123 4.538 2011     1208      2     3.825
## 16639 84855481533 3.290 2011     1208      2     2.523
## 16726 82455232963 4.204 2011     1208      3     3.072
## 16727 82455244076 4.997 2011     1208      3     3.919
## 16835 84871206897 3.953 2012     1208      2     3.336
## 16837 84871229041 3.842 2012     1208      2     3.225
## 16880 84870331366 3.391 2012     1208      1     2.774
## 16882 84870350788 6.781 2012     1208      1     6.164
## 16892 84855231775 3.391 2012     1208      1     2.719
## 16896 84856905745 3.391 2012     1208      1     2.774
## 16898 84856945735 4.382 2012     1208      1     3.765
## 16903 84871438787 3.937 2012     1208      1     3.265
## 16909 84871985774 3.391 2012     1208      1     2.774
## 16914 84872706828 3.391 2012     1208      1     2.774
```

```
## 16926 84873281046 3.937 2012     1208      1     3.320
## 16949 84874079832 7.447 2012     1208      1     6.775
## 16950 84874095385 3.937 2012     1208      1     3.265
## 16957 84876994217 5.086 2012     1208      1     4.414
## 16958 84876998823 3.391 2012     1208      1     2.719
## 16959 84877000419 3.937 2012     1208      1     3.265
## 16960 84878432148 4.382 2012     1208      1     3.765
## 17050 84871269332 3.156 2012     1208      2     2.539
## 17170 84871367509 3.245 2012     1202      3     2.573
## 17171 84871394290 5.615 2012     1202      3     4.943
## 17215 84874341423 3.664 2012     1202      3     3.047
## 17318 84869003069 5.086 2012     1208      1     4.469
## 17320 84869033091 3.937 2012     1208      1     3.265
## 17326 84868366032 3.394 2012     1208      2     2.777
## 17327 84868368393 3.394 2012     1208      2     2.777
## 17375 84874056751 3.264 2012     1202      4     2.592
## 17431 84871011871 3.937 2012     1208      1     3.320
## 17464 84868631017 3.559 2012     1208      2     2.887
## 17465 84868671128 3.962 2012     1208      2     3.345
## 17569 84865697715 3.939 2012     1208      2     3.322
## 17570 84865700895 4.859 2012     1208      2     4.187
## 17571 84865204088 3.937 2012     1208      1     3.320
## 17574 84866596965 4.759 2012     1208      1     4.142
## 17578 84866635016 3.937 2012     1208      1     3.320
## 17594 84867391849 3.391 2012     1208      1     2.719
## 17604 84870451817 3.937 2012     1208      1     3.320
## 17607 84870509187 4.382 2012     1208      1     3.765
## 17616 84871215069 3.391 2012     1208      1     2.719
## 17617 84871229501 3.937 2012     1208      1     3.320
## 17622 84871342985 5.374 2012     1208      1     4.337
## 17624 84871749299 3.391 2012     1208      1     2.719
## 17631 84873941703 3.937 2012     1208      1     3.320
## 17632 84873952317 3.391 2012     1208      1     2.774
## 17633 84873956808 3.937 2012     1208      1     3.265
## 17668 84885611499 5.086 2012     1208      1     4.469
## 17727 84865713540 3.451 2012     1202      3     2.834
## 17750 84866640267 5.127 2012     1208      3     4.455
## 17762 84866982605 3.164 2012     1203      3     2.547
## 17828 84864711707 3.258 2012     1202      5     2.641
## 17849 84865371557 5.086 2012     1208      1     4.414
## 17850 84865386536 6.781 2012     1208      1     6.164
## 17858 84865347957 3.391 2012     1208      1     2.719
## 17863 84865160135 5.086 2012     1208      1     4.469
## 17864 84865162714 5.086 2012     1208      1     4.414
## 17874 84864926904 4.302 2012     1208      2     3.685
## 17928 84870409111 3.654 2012     1203      3     2.835
## 17953 84864349568 5.374 2012     1208      1     4.702
## 17968 84874377033 4.759 2012     1208      1     4.142
## 17987 84864035512 4.620 2012     1208      2     4.003
## 18033 84864201531 3.937 2012     1208      1     3.265
```

```
## 18034 84864215638 6.455 2012     1208      1     5.783
## 18035 84864219524 5.086 2012     1208      1     4.414
## 18036 84864248050 5.374 2012     1208      1     4.702
## 18037 84864268571 6.455 2012     1208      1     5.838
## 18043 84865177201 3.937 2012     1208      1     3.320
## 18054 84875659820 5.632 2012     1208      1     5.015
## 18091 84866718537 4.477 2012     1208      2     3.860
## 18215 84859330957 5.086 2012     1208      1     4.469
## 18233 84861746194 3.412 2012     1208      2     2.795
## 18247 84883394654 3.596 2012     1208      2     2.924
## 18259 84872852785 3.240 2012     1203      3     2.623
## 18266 84872871724 3.721 2012     1203      3     3.049
## 18281 84859745370 4.759 2012     1208      1     4.087
## 18284 84859760309 3.391 2012     1208      1     2.774
## 18296 84876226058 4.759 2012     1208      1     4.087
## 18414 84865691275 3.391 2012     1208      1     2.774
## 18427 84884668586 3.391 2012     1208      1     2.774
## 18448 84861350433 4.598 2012     1208      2     3.926
## 18449 84861359249 3.559 2012     1208      2     2.887
## 18452 84861395909 3.559 2012     1208      2     2.887
## 18480 84860734606 3.999 2012     1202      3     3.327
## 18489 84862115131 3.380 2012     1202      4     2.708
## 18490 84862116557 3.380 2012     1202      4     2.763
## 18555 84856799214 3.937 2012     1208      1     3.265
## 18557 84856642915 4.382 2012     1208      1     3.710
## 18558 84856653530 4.382 2012     1208      1     3.710
## 18565 84856448236 3.227 2012     1202      3     2.610
## 18566 84856468758 4.331 2012     1202      3     3.659
## 18567 84856478984 3.344 2012     1202      3     2.672
## 18576 84864376206 4.361 2012     1208      2     3.379
## 18632 84856135031 3.937 2012     1208      1     3.265
## 18636 84856185796 4.759 2012     1208      1     4.142
## 18637 84856189103 4.759 2012     1208      1     4.142
## 18638 84856197118 6.781 2012     1208      1     6.164
## 18648 84863619762 3.391 2012     1208      1     2.774
## 18664 84867394653 3.937 2012     1208      1     3.320
## 18737 84865710516 3.559 2012     1208      2     2.887
## 18740 84867763226 3.284 2012     1205      2     2.667
## 18745 84870312908 3.962 2012     1208      2     3.345
## 18747 84870334955 6.510 2012     1208      2     5.838
## 18751 84870345747 6.391 2012     1208      2     5.774
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
UniqueAuthors +
##     Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##     k.max = 1000))
##  \--> method = "MM"
## Residuals:
```

```
##     Min      1Q  Median      3Q     Max
## -1.8964 -0.7403 -0.0354  0.7411  7.7840
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         0.76337    0.04960   15.39  < 2e-16 ***
## FirstAuthorFemale1 -0.16294    0.07115   -2.29  0.02203 *
## LastAuthorFemale1   0.21748    0.07116    3.06  0.00225 **
## UniqueAuthors2      0.36506    0.05218    7.00  2.7e-12 ***
## UniqueAuthors3      0.86887    0.10964    7.92  2.4e-15 ***
## UniqueAuthors4      0.95105    0.20142    4.72  2.4e-06 ***
## UniqueAuthors5      0.94174    0.44221    2.13  0.03322 *
## Year1997            0.01417    0.06591    0.22  0.82976
## Year1998            0.03707    0.06524    0.57  0.56983
## Year1999           -0.03549    0.06156   -0.58  0.56434
## Year2000            0.03561    0.06354    0.56  0.57518
## Year2001            0.00414    0.06137    0.07  0.94619
## Year2002            0.20966    0.05901    3.55  0.00038 ***
## Year2003            0.21149    0.05829    3.63  0.00029 ***
## Year2004            0.27574    0.05814    4.74  2.1e-06 ***
## Year2005            0.16972    0.05781    2.94  0.00333 **
## Year2006            0.17363    0.05771    3.01  0.00263 **
## Year2007            0.14400    0.05676    2.54  0.01119 *
## Year2008            0.08117    0.05618    1.44  0.14855
## Year2009           -0.07756    0.05678   -1.37  0.17193
## Year2010           -0.11162    0.05674   -1.97  0.04916 *
## Year2011           -0.05047    0.05621   -0.90  0.36927
## Year2012           -0.14622    0.05427   -2.69  0.00706 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.913
## Multiple R-squared:  0.03,   Adjusted R-squared:  0.0287
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
##  66 observations
c(2223,2421,3496,3615,4107,4269,5346,5362,5507,6122,6167,6744,8136,8137,8301,
8307,8344,8580,9339,9506,9559,9611,9629,10043,10161,10169,10679,10858,11065,1
1067,11072,11116,11120,11938,11941,11942,11945,13395,13479,13798,13846,14110,
14249,14369,14684,14748,14755,14952,15080,15341,15378,15452,15539,15540,15551
,15552,15630,15696,15697,15698,15699,15715,15863,16233,16332,16334)
##   are outliers with |weight| = 0 ( < 6.1e-06);
##  1163 weights are ~= 1. The remaining 15265 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0002  0.8890  0.9370  0.8910  0.9700  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
```

```
##          1.00e-07          1.00e-07          6.06e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##                     GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 22.570  1           4.751
## LastAuthorFemale  22.558  1           4.749
## Year               1.009 16           1.000
```




```
## [1] "List of  519 outliers with residuals above  2.5"
##          ScopusId  NLCS Year OneField Fields residuals
## 94    67651037634 3.400 1996     1208      1     2.626
## 138   61449406492 3.938 1996     1208      1     3.164
## 140   61949400015 3.345 1996     1208      1     2.514
## 376    0009431359 3.532 1997     1208      2     2.687
## 412   79959433454 3.987 1997     1208      2     3.200
## 952    0000490567 3.695 1998     1208      2     2.872
```

```
## 1207  60949584988 3.410 1999     1208      1     2.664
## 1485   0013039289 3.420 1999     1208      3     2.617
## 2557  10844245985 5.274 2002     1208      1     4.289
## 2562  33750620950 3.855 2002     1208      1     2.813
## 2563  33947361251 3.629 2002     1208      1     2.587
## 2565  34547299083 3.746 2002     1208      1     2.761
## 2566  47749108055 4.222 2002     1208      1     3.237
## 2581  60949458107 4.222 2002     1208      1     3.237
## 2605  60950002681 3.502 2002     1208      1     2.517
## 2607  60950027813 3.746 2002     1208      1     2.761
## 2619  60950346555 4.918 2002     1208      1     3.933
## 2732  65849154249 3.629 2002     1208      1     2.587
## 2733  65849216639 3.629 2002     1208      1     2.644
## 2734  65849259075 3.629 2002     1208      1     2.587
## 2736  65849326957 3.955 2002     1208      1     2.913
## 2748  67650096797 3.746 2002     1208      1     2.761
## 2751  70249100469 4.301 2002     1208      1     3.316
## 2752  70349438078 3.502 2002     1208      1     2.517
## 2780   3142697241 8.757 2002     1208      2     7.772
## 2781  33646911853 5.146 2002     1208      2     4.104
## 2783  33748905673 5.092 2002     1208      2     4.107
## 2789  44149119239 3.798 2002     1208      2     2.756
## 2825  60949601992 4.417 2002     1208      2     3.375
## 2860  60950259244 4.417 2002     1208      2     3.375
## 3194  61949270383 3.955 2002     1208      1     2.970
## 3322  34248720247 3.855 2002     1208      1     2.870
## 3323  34248734400 3.855 2002     1208      1     2.870
## 3326  38749133094 4.050 2002     1208      1     3.065
## 3327  58449103137 3.502 2002     1208      1     2.517
## 3332  60949291781 3.629 2002     1208      1     2.587
## 3342  60950305286 3.746 2002     1208      1     2.761
## 3345  60950448412 3.629 2002     1208      1     2.644
## 3347  60950498847 3.746 2002     1208      1     2.761
## 3348  60950501375 3.955 2002     1208      1     2.913
## 3352  60950608189 3.955 2002     1208      1     2.913
## 3354  60950639390 3.855 2002     1208      1     2.813
## 3355  60950641233 4.515 2002     1208      1     3.473
## 3358  60950680913 3.629 2002     1208      1     2.644
## 3361  60950699119 3.855 2002     1208      1     2.813
## 3380  61249427664 3.502 2002     1208      1     2.517
## 3382  61249520644 3.955 2002     1208      1     2.913
## 3439  70349348163 4.050 2002     1208      1     3.065
## 3440  70349378641 3.629 2002     1208      1     2.587
## 3460  70450000901 3.629 2002     1208      1     2.644
## 3504  61249583521 3.963 2002     1208      2     2.978
## 3509  61449442102 3.533 2002     1208      2     2.548
## 3515  62449109752 4.033 2002     1208      2     2.991
## 3516  62449174089 3.672 2002     1208      2     2.687
## 3517  62449296550 4.063 2002     1208      2     3.021
## 3657  60950279634 3.963 2002     1208      3     2.921
```

```
## 3793  85012444559 3.746 2002     1202      3     2.761
## 3843   0036526749 3.670 2002     1202      6     2.628
## 3845   1442335800 3.694 2002     1202      6     2.709
## 3884  34047093694 4.436 2003     1208      1     3.392
## 3887  46749133770 4.109 2003     1208      1     3.123
## 3912  60949464452 3.780 2003     1208      1     2.794
## 3918  60949658016 5.148 2003     1208      1     4.104
## 3921  60949700952 4.436 2003     1208      1     3.392
## 3960  60950414028 4.651 2003     1208      1     3.665
## 3963  60950438763 4.034 2003     1208      1     3.048
## 3977  60950500372 4.249 2003     1208      1     3.263
## 3993  60950588015 4.314 2003     1208      1     3.328
## 3999  60950625661 5.403 2003     1208      1     4.417
## 4036  61249262249 4.034 2003     1208      1     3.048
## 4094  70449761302 3.580 2003     1208      1     2.594
## 4141   8544244308 4.547 2003     1208      1     3.561
## 4142  26944502435 4.044 2003     1208      2     3.000
## 4143  33646681740 6.083 2003     1208      2     5.039
## 4146  34547600256 4.341 2003     1208      2     3.297
## 4197  60950066571 4.869 2003     1208      2     3.825
## 4230  61149570569 3.528 2003     1208      2     2.542
## 4489  45449108735 5.113 2003     1208      1     4.069
## 4583  61949186737 4.181 2003     1208      1     3.195
## 4600  34748861400 4.651 2003     1208      1     3.607
## 4621  41549119242 4.109 2003     1208      1     3.123
## 4691  36248970780 3.954 2003     1208      1     2.910
## 4697  60949457140 8.312 2003     1208      1     7.326
## 4710  60950639325 3.780 2003     1208      1     2.736
## 4713  60950694241 3.954 2003     1208      1     2.968
## 4736  61249635138 4.109 2003     1208      1     3.065
## 4742  61449197953 4.922 2003     1208      1     3.936
## 4747  61449337582 4.837 2003     1208      1     3.793
## 4782  64949161384 3.580 2003     1208      1     2.536
## 4822  70449843968 3.683 2003     1208      1     2.639
## 4863  29744456134 5.218 2003     1208      2     4.232
## 4872  34547980224 3.959 2003     1208      2     2.973
## 4873  48549096636 5.652 2003     1208      2     4.608
## 4888  61249716500 4.045 2003     1208      2     3.059
## 4990  70449856706 3.528 2003     1208      2     2.542
## 5092  61949103780 3.552 2003     1208      3     2.508
## 5274  28244501538 5.482 2004     1208      1     4.428
## 5280  34748922675 4.292 2004     1208      1     3.180
## 5307  60949771564 3.622 2004     1208      1     2.568
## 5445  61949116307 3.947 2004     1208      1     2.893
## 5463  67049083067 4.804 2004     1208      1     3.750
## 5468  67649743610 4.213 2004     1208      1     3.101
## 5514  34250312985 4.933 2004     1208      2     3.821
## 5562  60950391613 4.521 2004     1208      2     3.409
## 5581  60950719934 3.827 2004     1208      2     2.773
## 5935  60949318761 3.739 2004     1208      1     2.918
```

```
## 6108  33750435368 5.377 2004     1208      1     4.323
## 6112  33947671157 4.506 2004     1208      1     3.394
## 6126  60949382516 6.369 2004     1208      1     5.315
## 6152  61049449507 5.223 2004     1208      1     4.169
## 6178  61449455372 3.739 2004     1208      1     2.685
## 6196  62549140156 3.622 2004     1208      1     2.568
## 6270  71249085963 3.622 2004     1208      1     2.568
## 6285  60950338479 5.580 2004     1208      2     4.468
## 6299  61849165077 3.827 2004     1208      2     2.715
## 6350  70449786556 4.699 2004     1208      2     3.645
## 6352  70449824679 4.933 2004     1208      2     3.879
## 6699  34248708340 3.945 2005     1208      1     2.997
## 6705  34548494655 3.822 2005     1208      1     2.817
## 6729  60949655783 3.688 2005     1208      1     2.740
## 6765  60950409349 3.945 2005     1208      1     2.997
## 6790  60950596472 3.541 2005     1208      1     2.536
## 6800  60950708929 3.541 2005     1208      1     2.593
## 6850  61149580054 4.683 2005     1208      1     3.678
## 6901  65449143734 4.358 2005     1208      1     3.353
## 6902  65449165806 4.683 2005     1208      1     3.678
## 6909  67649784613 4.445 2005     1208      1     3.497
## 6947  70449899952 3.822 2005     1208      1     2.874
## 6955  70449970464 3.541 2005     1208      1     2.593
## 6966  70450058332 3.688 2005     1208      1     2.740
## 6985  33750587792 3.538 2005     1208      2     2.533
## 6989  34250008387 4.562 2005     1208      2     3.614
## 6994  38649134273 6.094 2005     1208      2     5.089
## 6999  60949254410 4.465 2005     1208      2     3.517
## 7041  60950274865 3.944 2005     1208      2     2.996
## 7043  60950350045 5.477 2005     1208      2     4.529
## 7070  60950562957 3.861 2005     1208      2     2.913
## 7078  60950642207 3.861 2005     1208      2     2.913
## 7177  33845250551 4.010 2005     1202      4     3.062
## 7380  62749191854 3.744 2005     1202      3     2.739
## 7383  63849109891 3.549 2005     1202      3     2.544
## 7553  60950504855 3.650 2005     1208      3     2.645
## 7604  61949241615 3.822 2005     1208      1     2.817
## 7630  61049327976 3.822 2005     1208      1     2.817
## 7650  21544433512 4.003 2005     1208      2     3.289
## 7723  33846337496 6.348 2005     1208      1     5.343
## 7724  46249120022 4.059 2005     1208      1     3.054
## 7725  46249120887 3.945 2005     1208      1     2.997
## 7734  60949594699 3.822 2005     1208      1     2.817
## 7742  60950233081 4.059 2005     1208      1     3.111
## 7743  60950413511 3.541 2005     1208      1     2.593
## 7755  60950658414 3.541 2005     1208      1     2.536
## 7804  61449502250 3.688 2005     1208      1     2.683
## 7832  67649792074 3.822 2005     1208      1     2.874
## 7837  67650088057 3.541 2005     1208      1     2.593
## 7844  67650178186 4.059 2005     1208      1     3.111
```

```
## 7845  68549117379 3.822 2005     1208      1     2.874
## 7868  70450064684 3.822 2005     1208      1     2.874
## 7876  70450093197 3.541 2005     1208      1     2.593
## 7898  44449106909 4.250 2005     1208      2     3.245
## 7902  60950423747 5.118 2005     1208      2     4.113
## 7914  61249595569 3.788 2005     1208      2     2.840
## 7915  61449306806 3.654 2005     1208      2     2.940
## 7919  61449561847 3.582 2005     1208      2     2.577
## 7937  62449266342 3.723 2005     1208      2     2.775
## 8040  33846333428 3.605 2005     1208      3     2.657
## 8313  50949119108 4.710 2006     1208      1     3.759
## 8336  60949594614 3.653 2006     1208      1     2.645
## 8347  60950223477 3.786 2006     1208      1     2.778
## 8356  60950442490 3.908 2006     1208      1     2.957
## 8368  60950493268 4.224 2006     1208      1     3.273
## 8369  60950493735 3.653 2006     1208      1     2.645
## 8373  60950504606 3.786 2006     1208      1     2.778
## 8394  60950726337 4.317 2006     1208      1     3.309
## 8444  61449369584 3.508 2006     1208      1     2.557
## 8484  67650109778 3.786 2006     1208      1     2.778
## 8486  67650128316 3.653 2006     1208      1     2.702
## 8487  67650135684 3.786 2006     1208      1     2.835
## 8490  67650164706 4.317 2006     1208      1     3.309
## 8529  70450105045 4.224 2006     1208      1     3.273
## 9018  60949478073 3.653 2006     1208      1     2.702
## 9237  61949128551 4.404 2006     1208      1     3.453
## 9296  84870103615 4.317 2006     1208      1     3.366
## 9311  42449161018 5.130 2006     1208      1     4.122
## 9313  53549130671 7.825 2006     1208      1     6.817
## 9314  55349128031 5.280 2006     1208      1     4.272
## 9325  60950413512 4.777 2006     1208      1     3.769
## 9326  60950445899 3.908 2006     1208      1     2.900
## 9327  60950622690 4.126 2006     1208      1     3.175
## 9342  61249099688 4.021 2006     1208      1     3.013
## 9344  61249305083 4.317 2006     1208      1     3.366
## 9351  61449427822 3.653 2006     1208      1     2.645
## 9352  61449427823 4.126 2006     1208      1     3.118
## 9361  61949197756 4.021 2006     1208      1     3.070
## 9365  62349122821 3.653 2006     1208      1     2.645
## 9367  62449189059 4.777 2006     1208      1     3.826
## 9371  62749148771 4.777 2006     1208      1     3.826
## 9381  64949139647 4.224 2006     1208      1     3.216
## 9426  67650079640 3.908 2006     1208      1     2.900
## 9440  70449824307 3.786 2006     1208      1     2.835
## 9465  70450032962 3.786 2006     1208      1     2.544
## 9468  70450045620 4.126 2006     1208      1     3.118
## 9469  70450058396 3.786 2006     1208      1     2.835
## 9479  70749096336 4.021 2006     1208      1     3.013
## 9491  33846152354 3.582 2006     1208      2     2.631
## 9497  34249084654 5.314 2006     1208      2     4.306
```

```
## 9503  44449151884 6.523 2006     1208      2     5.515
## 9504  60949270586 3.627 2006     1208      2     2.676
## 9524  61449282187 3.627 2006     1208      2     2.676
## 9541  61449530625 6.523 2006     1208      2     5.572
## 9574  70349296952 3.627 2006     1208      2     2.676
## 9577  70449775017 4.234 2006     1208      2     3.283
## 9589  70450078889 3.627 2006     1208      2     2.619
## 9803  57749157632 6.254 2007     1208      1     5.277
## 9867  60950549890 3.610 2007     1208      1     2.633
## 9899  61149182825 4.555 2007     1208      1     3.635
## 9949  62449306099 4.555 2007     1208      1     3.635
## 9967  66049152167 4.214 2007     1208      1     3.237
## 9978  67650155463 3.783 2007     1208      1     2.806
## 9981  67650162724 3.610 2007     1208      1     2.633
## 9982  67650168953 3.939 2007     1208      1     2.962
## 10006 70449930640 3.610 2007     1208      1     2.690
## 10040 84868448556 3.610 2007     1208      1     2.690
## 10075 43249135021 4.018 2007     1208      2     3.041
## 10078 43249146684 4.726 2007     1208      2     3.749
## 10107 60949438530 4.726 2007     1208      2     3.806
## 10155 60950643835 4.839 2007     1208      2     3.919
## 10158 60950662433 3.834 2007     1208      2     2.914
## 10415 61949141429 3.441 2007     1208      3     2.521
## 10670 61049446035 5.641 2007     1208      1     4.721
## 10676 70450032239 3.610 2007     1208      1     2.690
## 10809 70450064287 5.151 2007     1208      1     4.231
## 10866 60949326226 3.610 2007     1208      1     2.633
## 10867 60949370742 5.976 2007     1208      1     5.056
## 10880 60950131470 3.939 2007     1208      1     2.962
## 10892 60950536451 4.449 2007     1208      1     3.472
## 10925 62449170699 5.841 2007     1208      1     4.864
## 10968 67649683064 4.214 2007     1208      1     3.294
## 10984 67650096553 4.082 2007     1208      1     3.162
## 10985 67650143129 5.474 2007     1208      1     4.554
## 10986 67650925409 3.939 2007     1208      1     3.019
## 10988 70249093068 3.610 2007     1208      1     2.690
## 10992 70449859265 3.610 2007     1208      1     2.690
## 10994 70449877375 3.610 2007     1208      1     2.690
## 11007 70450015708 6.179 2007     1208      1     5.259
## 11014 70450064757 3.783 2007     1208      1     2.806
## 11015 70450077380 4.654 2007     1208      1     3.734
## 11112 62449186372 3.857 2007     1208      2     2.880
## 11165 70450078906 4.018 2007     1208      2     3.098
## 11217 61249239835 4.188 2007     1208      3     3.268
## 11228 61249660537 3.741 2007     1208      3     2.821
## 11409 60949485276 3.367 2008     1208      1     2.509
## 11434 60950624831 3.802 2008     1208      1     2.944
## 11478 60950719981 5.997 2008     1208      1     5.081
## 11494 60950608143 3.692 2008     1203      3     2.543
## 11532 60950299886 3.598 2008     1208      1     2.740
```

```
## 11540 57749125280 4.798 2008     1208      1     3.940
## 11548 60949270646 4.686 2008     1208      1     3.828
## 11573 61049554358 3.802 2008     1208      1     2.944
## 11605 61949182548 3.598 2008     1208      1     2.682
## 11611 61949371494 3.367 2008     1208      1     2.509
## 11618 62449098267 5.426 2008     1208      1     4.568
## 11624 62449159076 3.598 2008     1208      1     2.740
## 11626 62449175332 5.942 2008     1208      1     5.026
## 11633 62449269964 3.984 2008     1208      1     3.126
## 11641 62949147900 4.300 2008     1208      1     3.384
## 11644 62949210835 3.367 2008     1208      1     2.509
## 11672 67650099391 3.984 2008     1208      1     3.068
## 11676 67650112946 3.984 2008     1208      1     3.068
## 11681 67650132635 4.149 2008     1208      1     3.233
## 11684 67650151239 3.984 2008     1208      1     3.068
## 11687 67650155476 3.367 2008     1208      1     2.509
## 11688 67650162729 4.798 2008     1208      1     3.882
## 11691 67650178260 5.001 2008     1208      1     4.085
## 11695 70449752752 4.300 2008     1208      1     3.442
## 11696 70449780196 3.367 2008     1208      1     2.509
## 11723 70450068192 3.367 2008     1208      1     2.509
## 11792 61149209901 3.427 2008     1208      2     2.569
## 11796 61149472770 4.222 2008     1208      2     3.306
## 12102 65749261392 3.585 2008     1202      3     2.669
## 12226 70450003031 5.184 2008     1208      1     4.268
## 12362 70450082589 3.516 2008     1205      3     2.600
## 12399 45349087542 3.582 2008     1208      2     2.724
## 12435 61249636505 6.426 2008     1208      1     5.568
## 12468 60949164169 3.720 2008     1208      3     2.862
## 12498 70449866030 3.802 2008     1208      1     2.886
## 12500 70450060571 4.686 2008     1208      1     3.770
## 12551 61049119505 3.367 2008     1208      1     2.509
## 12629 57749127836 3.984 2008     1208      1     3.126
## 12658 61249639675 3.598 2008     1208      1     2.682
## 12661 61449513631 3.984 2008     1208      1     3.126
## 12664 61949197772 4.149 2008     1208      1     3.291
## 12671 62449262219 4.300 2008     1208      1     3.442
## 12675 62749110813 6.152 2008     1208      1     5.294
## 12677 62749122811 5.268 2008     1208      1     4.410
## 12679 62749136269 3.984 2008     1208      1     3.126
## 12682 62749151909 4.798 2008     1208      1     3.940
## 12683 62749155048 4.149 2008     1208      1     3.233
## 12684 62749156172 5.942 2008     1208      1     5.084
## 12691 64949093382 3.598 2008     1208      1     2.682
## 12696 64949107038 3.802 2008     1208      1     2.886
## 12700 64949153413 4.300 2008     1208      1     3.442
## 12727 67650070793 4.149 2008     1208      1     3.291
## 12735 67650090943 5.570 2008     1208      1     4.712
## 12739 70449931160 5.268 2008     1208      1     4.352
## 12746 78650710041 3.802 2008     1208      1     2.886
```

```
## 12795 61249468411 4.222 2008     1208      2     3.364
## 12797 61249548069 3.427 2008     1208      2     2.569
## 12861 40449115608 3.564 2008     1202      4     2.939
## 13002 85014946308 3.363 2008     1203      4     2.505
## 13097 77950310142 4.054 2009     1208      1     3.287
## 13214 84884186648 4.054 2009     1208      1     3.344
## 13249 61949462044 3.777 2009     1208      3     3.067
## 13400 73949133087 4.425 2009     1208      1     3.715
## 13418 70349840389 4.249 2009     1208      1     3.539
## 13474 68549123389 3.591 2009     1208      1     2.881
## 13490 84867054792 3.306 2009     1208      1     2.830
## 13559 73649148390 3.360 2009     1202      4     2.650
## 13613 67649716605 3.591 2009     1208      1     2.824
## 13663 67649950267 4.054 2009     1208      1     3.344
## 13679 77950325600 4.997 2009     1208      1     4.287
## 13680 77950788847 3.306 2009     1208      1     2.596
## 13683 77950809870 4.054 2009     1208      1     3.054
## 13684 77950815002 6.760 2009     1208      1     5.993
## 13685 77950822756 5.116 2009     1208      1     4.349
## 13686 77950831009 4.585 2009     1208      1     3.875
## 13688 77950841010 6.560 2009     1208      1     5.793
## 13693 68949083793 3.825 2009     1208      2     3.115
## 13695 68949090683 4.540 2009     1208      2     3.830
## 13697 68949092620 3.727 2009     1208      2     3.017
## 13700 68949123188 3.423 2009     1208      2     2.713
## 13702 68949135917 3.636 2009     1208      2     2.926
## 13893 79958822732 4.054 2009     1208      1     3.287
## 13895 80054949947 3.837 2009     1208      1     3.070
## 13939 77950229849 3.223 2009     1208      2     2.513
## 13940 77950253669 3.223 2009     1208      2     2.513
## 14020 65849111025 3.591 2009     1208      1     2.824
## 14021 65849141252 3.306 2009     1208      1     2.596
## 14039 77956351978 3.837 2009     1208      1     3.070
## 14092 77950181710 4.785 2009     1205      2     4.018
## 14275 77950904394 3.302 2010     1208      1     2.629
## 14278 78649401790 3.586 2010     1208      1     2.856
## 14282 78649430807 3.586 2010     1208      1     2.856
## 14293 84855406055 3.832 2010     1208      1     3.102
## 14542 77957559973 4.049 2010     1208      1     3.376
## 14543 77957584700 3.302 2010     1208      1     2.572
## 14572 77956632352 3.506 2010     1208      2     2.776
## 14575 78649656075 4.243 2010     1208      1     3.513
## 14591 77956331789 4.376 2010     1208      2     3.703
## 14712 79959202296 3.634 2010     1208      2     2.961
## 14810 77950211324 3.302 2010     1208      1     2.629
## 14880 79956372974 3.668 2010     1202      4     2.995
## 14898 77949582311 4.243 2010     1208      1     3.513
## 14902 77949615907 3.586 2010     1208      1     2.913
## 14927 77950140620 3.832 2010     1208      1     3.159
## 14928 77950253361 3.302 2010     1208      1     2.629
```

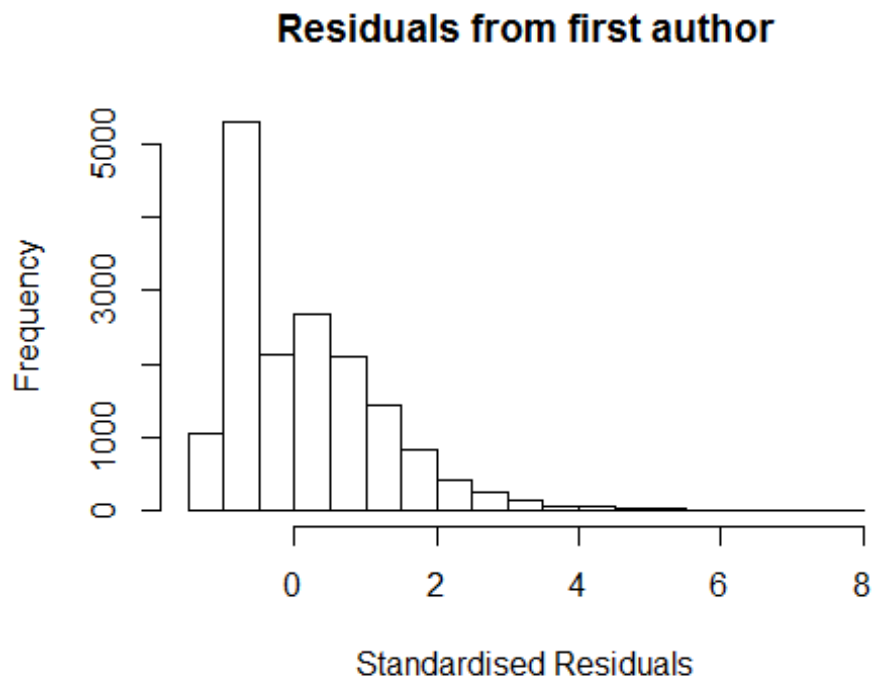
```
## 14934 77950886825 3.586 2010     1208      1     2.856
## 14935 77950888966 3.586 2010     1208      1     2.856
## 15039 76349106441 3.586 2010     1208      1     2.856
## 15050 77954443733 3.302 2010     1208      1     2.572
## 15094 77950287584 3.586 2010     1208      1     2.913
## 15120 79955418712 4.243 2010     1208      1     3.513
## 15143 73249128238 4.062 2010     1208      2     3.622
## 15226 77956500865 4.776 2010     1208      3     4.103
## 15228 78649472664 5.064 2010     1208      3     4.100
## 15279 82455228764 3.393 2010     1208      3     2.720
## 15298 77951908318 3.254 2010     1202      6     2.581
## 15358 84855209572 3.491 2011     1208      1     2.756
## 15359 84855210092 3.491 2011     1208      1     2.698
## 15365 84855762177 3.887 2011     1208      1     3.152
## 15368 84855797090 3.491 2011     1208      1     2.756
## 15380 84859358430 3.491 2011     1208      1     2.756
## 15382 84860778742 5.724 2011     1208      1     4.989
## 15403 84870329240 3.491 2011     1208      1     2.756
## 15411 84872712403 3.491 2011     1208      1     2.698
## 15419 84873283995 3.491 2011     1208      1     2.698
## 15424 84873302367 3.887 2011     1208      1     3.152
## 15447 84884198571 3.887 2011     1208      1     3.152
## 15465 84887967106 3.491 2011     1208      1     2.756
## 15474 84855781328 3.573 2011     1208      2     2.838
## 15475 84855791481 5.406 2011     1208      2     4.613
## 15484 84857779496 3.573 2011     1208      2     2.547
## 15615 84857935340 3.474 2011     1202      6     2.739
## 15674 80155194058 3.906 2011     1202      5     3.113
## 15694 80054789940 4.221 2011     1208      1     3.428
## 15698 80054708194 3.887 2011     1208      1     3.152
## 15717 80053615731 4.511 2011     1208      1     3.776
## 15803 80052975515 3.243 2011     1205      2     2.508
## 15826 80052542479 3.887 2011     1208      1     3.152
## 15845 84055192224 6.801 2011     1208      1     6.008
## 15847 84856196552 3.887 2011     1208      1     3.152
## 15853 84862322836 3.887 2011     1208      1     3.094
## 15869 84883480879 4.995 2011     1208      1     4.202
## 15899 84861070464 5.091 2011     1208      2     4.356
## 16068 79959952724 4.221 2011     1208      1     3.428
## 16069 79959977747 3.491 2011     1208      1     2.698
## 16212 79955377318 6.894 2011     1208      1     6.101
## 16216 79955439141 4.221 2011     1208      1     3.486
## 16219 79957888983 3.887 2011     1208      1     3.152
## 16290 79958846842 3.887 2011     1208      1     3.152
## 16360 79955674091 3.677 2011     1203      6     2.884
## 16379 79955677215 6.146 2011     1208      1     5.353
## 16383 80054915853 3.887 2011     1208      1     3.094
## 16498 79955368617 3.375 2011     1203      3     2.582
## 16516 78650843435 5.564 2011     1208      1     4.829
## 16517 78650845690 4.221 2011     1208      1     3.486
```

```
## 16541 79960860760 4.221 2011     1208      1     3.428
## 16542 79960871745 4.221 2011     1208      1     3.486
## 16550 81355124072 3.887 2011     1208      1     3.094
## 16567 84055198510 3.887 2011     1208      1     3.094
## 16586 85011484631 3.887 2011     1208      1     3.094
## 16615 80555156002 3.648 2011     1208      2     2.855
## 16630 84555179123 4.538 2011     1208      2     3.803
## 16726 82455232963 4.204 2011     1208      3     3.411
## 16727 82455244076 4.997 2011     1208      3     4.262
## 16835 84871206897 3.953 2012     1208      2     3.313
## 16837 84871229041 3.842 2012     1208      2     3.202
## 16880 84870331366 3.391 2012     1208      1     2.751
## 16882 84870350788 6.781 2012     1208      1     6.141
## 16892 84855231775 3.391 2012     1208      1     2.693
## 16896 84856905745 3.391 2012     1208      1     2.751
## 16898 84856945735 4.382 2012     1208      1     3.742
## 16903 84871438787 3.937 2012     1208      1     3.239
## 16909 84871985774 3.391 2012     1208      1     2.751
## 16914 84872706828 3.391 2012     1208      1     2.751
## 16926 84873281046 3.937 2012     1208      1     3.297
## 16949 84874079832 7.447 2012     1208      1     6.749
## 16950 84874095385 3.937 2012     1208      1     3.239
## 16957 84876994217 5.086 2012     1208      1     4.388
## 16958 84876998823 3.391 2012     1208      1     2.693
## 16959 84877000419 3.937 2012     1208      1     3.239
## 16960 84878432148 4.382 2012     1208      1     3.742
## 17050 84871269332 3.156 2012     1208      2     2.516
## 17170 84871367509 3.245 2012     1202      3     2.547
## 17171 84871394290 5.615 2012     1202      3     4.917
## 17215 84874341423 3.664 2012     1202      3     3.024
## 17318 84869003069 5.086 2012     1208      1     4.446
## 17320 84869033091 3.937 2012     1208      1     3.239
## 17326 84868366032 3.394 2012     1208      2     2.754
## 17327 84868368393 3.394 2012     1208      2     2.754
## 17375 84874056751 3.264 2012     1202      4     2.566
## 17431 84871011871 3.937 2012     1208      1     3.297
## 17462 84867923419 3.313 2012     1208      2     2.906
## 17464 84868631017 3.559 2012     1208      2     2.861
## 17465 84868671128 3.962 2012     1208      2     3.322
## 17569 84865697715 3.939 2012     1208      2     3.299
## 17570 84865700895 4.859 2012     1208      2     4.161
## 17571 84865204088 3.937 2012     1208      1     3.297
## 17574 84866596965 4.759 2012     1208      1     4.119
## 17578 84866635016 3.937 2012     1208      1     3.297
## 17594 84867391849 3.391 2012     1208      1     2.693
## 17604 84870451817 3.937 2012     1208      1     3.297
## 17607 84870509187 4.382 2012     1208      1     3.742
## 17616 84871215069 3.391 2012     1208      1     2.693
## 17617 84871229501 3.937 2012     1208      1     3.297
## 17622 84871342985 5.374 2012     1208      1     4.676
```

```
## 17624 84871749299 3.391 2012     1208      1     2.693
## 17631 84873941703 3.937 2012     1208      1     3.297
## 17632 84873952317 3.391 2012     1208      1     2.751
## 17633 84873956808 3.937 2012     1208      1     3.239
## 17668 84885611499 5.086 2012     1208      1     4.446
## 17727 84865713540 3.451 2012     1202      3     2.811
## 17750 84866640267 5.127 2012     1208      3     4.429
## 17762 84866982605 3.164 2012     1203      3     2.524
## 17828 84864711707 3.258 2012     1202      5     2.618
## 17849 84865371557 5.086 2012     1208      1     4.388
## 17850 84865386536 6.781 2012     1208      1     6.141
## 17858 84865347957 3.391 2012     1208      1     2.693
## 17863 84865160135 5.086 2012     1208      1     4.446
## 17864 84865162714 5.086 2012     1208      1     4.388
## 17874 84864926904 4.302 2012     1208      2     3.662
## 17897 84883323231 3.240 2012     1208      2     2.600
## 17928 84870409111 3.654 2012     1203      3     3.247
## 17953 84864349568 5.374 2012     1208      1     4.676
## 17968 84874377033 4.759 2012     1208      1     4.119
## 17987 84864035512 4.620 2012     1208      2     3.980
## 18033 84864201531 3.937 2012     1208      1     3.239
## 18034 84864215638 6.455 2012     1208      1     5.757
## 18035 84864219524 5.086 2012     1208      1     4.388
## 18036 84864248050 5.374 2012     1208      1     4.676
## 18037 84864268571 6.455 2012     1208      1     5.815
## 18043 84865177201 3.937 2012     1208      1     3.297
## 18054 84875659820 5.632 2012     1208      1     4.992
## 18091 84866718537 4.477 2012     1208      2     3.837
## 18215 84859330957 5.086 2012     1208      1     4.446
## 18233 84861746194 3.412 2012     1208      2     2.772
## 18242 84883352820 3.619 2012     1208      2     2.979
## 18247 84883394654 3.596 2012     1208      2     2.898
## 18259 84872852785 3.240 2012     1203      3     2.600
## 18266 84872871724 3.721 2012     1203      3     3.023
## 18281 84859745370 4.759 2012     1208      1     4.061
## 18284 84859760309 3.391 2012     1208      1     2.751
## 18296 84876226058 4.759 2012     1208      1     4.061
## 18414 84865691275 3.391 2012     1208      1     2.751
## 18427 84884668586 3.391 2012     1208      1     2.751
## 18448 84861350433 4.598 2012     1208      2     3.900
## 18449 84861359249 3.559 2012     1208      2     2.861
## 18452 84861395909 3.559 2012     1208      2     2.861
## 18480 84860734606 3.999 2012     1202      3     3.301
## 18489 84862115131 3.380 2012     1202      4     2.682
## 18490 84862116557 3.380 2012     1202      4     2.740
## 18555 84856799214 3.937 2012     1208      1     3.239
## 18557 84856642915 4.382 2012     1208      1     3.684
## 18558 84856653530 4.382 2012     1208      1     3.684
## 18565 84856448236 3.227 2012     1202      3     2.587
## 18566 84856468758 4.331 2012     1202      3     3.633
```

```
## 18567 84856478984 3.344 2012     1202      3     2.646
## 18576 84864376206 4.361 2012     1208      2     3.721
## 18632 84856135031 3.937 2012     1208      1     3.239
## 18636 84856185796 4.759 2012     1208      1     4.119
## 18637 84856189103 4.759 2012     1208      1     4.119
## 18638 84856197118 6.781 2012     1208      1     6.141
## 18648 84863619762 3.391 2012     1208      1     2.751
## 18664 84867394653 3.937 2012     1208      1     3.297
## 18737 84865710516 3.559 2012     1208      2     2.861
## 18740 84867763226 3.284 2012     1205      2     2.644
## 18745 84870312908 3.962 2012     1208      2     3.322
## 18747 84870334955 6.510 2012     1208      2     5.812
## 18751 84870345747 6.391 2012     1208      2     5.751
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.2770 -0.7533 -0.0364  0.7443  7.7722
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         0.77393    0.04953   15.63  < 2e-16 ***
## FirstAuthorFemale1 -0.23333    0.07605   -3.07  0.00216 **
## LastAuthorFemale1   0.29089    0.07599    3.83  0.00013 ***
## Year1997            0.01345    0.06602    0.20  0.83853
## Year1998            0.04867    0.06540    0.74  0.45680
## Year1999           -0.02822    0.06162   -0.46  0.64691
## Year2000            0.04535    0.06348    0.71  0.47494
## Year2001            0.00952    0.06161    0.15  0.87715
## Year2002            0.21084    0.05891    3.58  0.00035 ***
## Year2003            0.21216    0.05813    3.65  0.00026 ***
## Year2004            0.28004    0.05811    4.82  1.5e-06 ***
## Year2005            0.17368    0.05776    3.01  0.00264 **
## Year2006            0.17670    0.05769    3.06  0.00220 **
## Year2007            0.14574    0.05667    2.57  0.01014 *
## Year2008            0.08431    0.05616    1.50  0.13333
## Year2009           -0.06434    0.05698   -1.13  0.25880
## Year2010           -0.10101    0.05689   -1.78  0.07582 .
## Year2011           -0.03848    0.05632   -0.68  0.49448
## Year2012           -0.13373    0.05435   -2.46  0.01389 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.919
## Multiple R-squared:  0.0202, Adjusted R-squared:  0.0191
## Convergence in 12 IRWLS iterations
```

```
##
## Robustness weights:
##  63 observations
c(2421,3496,3615,4107,4269,4621,5346,5362,5507,6122,6167,6744,8136,8301,8307,
8344,8580,9339,9506,9559,9611,9629,10043,10161,10169,10858,11065,11067,11072,
11116,11120,11941,11942,11945,13395,13479,13798,13846,14110,14249,14369,14684
,14748,14755,14952,15080,15341,15378,15452,15539,15540,15551,15552,15630,1569
6,15697,15698,15699,15715,15863,16233,16332,16334)
##   are outliers with |weight| <= 1.4e-07 ( < 6.1e-06);
##  1187 weights are ~= 1. The remaining 15244 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0001  0.8930  0.9340  0.8910  0.9710  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          6.06e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                    psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.005  1           1.003
## Year              1.005 16           1.000
```




```
## [1] "List of  519 outliers with residuals above  2.5"
##          ScopusId  NLCS Year OneField Fields residuals
## 94    67651037634 3.400 1996     1208      1     2.626
## 138   61449406492 3.938 1996     1208      1     3.164
## 140   61949400015 3.345 1996     1208      1     2.514
## 376    0009431359 3.532 1997     1208      2     2.687
## 412   79959433454 3.987 1997     1208      2     3.200
## 952    0000490567 3.695 1998     1208      2     2.872
## 1207  60949584988 3.410 1999     1208      1     2.664
## 1485   0013039289 3.420 1999     1208      3     2.617
## 2557  10844245985 5.274 2002     1208      1     4.289
## 2562  33750620950 3.855 2002     1208      1     2.813
## 2563  33947361251 3.629 2002     1208      1     2.587
## 2565  34547299083 3.746 2002     1208      1     2.761
## 2566  47749108055 4.222 2002     1208      1     3.237
## 2581  60949458107 4.222 2002     1208      1     3.237
## 2605  60950002681 3.502 2002     1208      1     2.517
## 2607  60950027813 3.746 2002     1208      1     2.761
## 2619  60950346555 4.918 2002     1208      1     3.933
## 2732  65849154249 3.629 2002     1208      1     2.587
## 2733  65849216639 3.629 2002     1208      1     2.644
## 2734  65849259075 3.629 2002     1208      1     2.587
## 2736  65849326957 3.955 2002     1208      1     2.913
## 2748  67650096797 3.746 2002     1208      1     2.761
## 2751  70249100469 4.301 2002     1208      1     3.316
## 2752  70349438078 3.502 2002     1208      1     2.517
## 2780   3142697241 8.757 2002     1208      2     7.772
```

```
## 2781  33646911853 5.146 2002     1208      2     4.104
## 2783  33748905673 5.092 2002     1208      2     4.107
## 2789  44149119239 3.798 2002     1208      2     2.756
## 2825  60949601992 4.417 2002     1208      2     3.375
## 2860  60950259244 4.417 2002     1208      2     3.375
## 3194  61949270383 3.955 2002     1208      1     2.970
## 3322  34248720247 3.855 2002     1208      1     2.870
## 3323  34248734400 3.855 2002     1208      1     2.870
## 3326  38749133094 4.050 2002     1208      1     3.065
## 3327  58449103137 3.502 2002     1208      1     2.517
## 3332  60949291781 3.629 2002     1208      1     2.587
## 3342  60950305286 3.746 2002     1208      1     2.761
## 3345  60950448412 3.629 2002     1208      1     2.644
## 3347  60950498847 3.746 2002     1208      1     2.761
## 3348  60950501375 3.955 2002     1208      1     2.913
## 3352  60950608189 3.955 2002     1208      1     2.913
## 3354  60950639390 3.855 2002     1208      1     2.813
## 3355  60950641233 4.515 2002     1208      1     3.473
## 3358  60950680913 3.629 2002     1208      1     2.644
## 3361  60950699119 3.855 2002     1208      1     2.813
## 3380  61249427664 3.502 2002     1208      1     2.517
## 3382  61249520644 3.955 2002     1208      1     2.913
## 3439  70349348163 4.050 2002     1208      1     3.065
## 3440  70349378641 3.629 2002     1208      1     2.587
## 3460  70450000901 3.629 2002     1208      1     2.644
## 3504  61249583521 3.963 2002     1208      2     2.978
## 3509  61449442102 3.533 2002     1208      2     2.548
## 3515  62449109752 4.033 2002     1208      2     2.991
## 3516  62449174089 3.672 2002     1208      2     2.687
## 3517  62449296550 4.063 2002     1208      2     3.021
## 3657  60950279634 3.963 2002     1208      3     2.921
## 3793  85012444559 3.746 2002     1202      3     2.761
## 3843   0036526749 3.670 2002     1202      6     2.628
## 3845   1442335800 3.694 2002     1202      6     2.709
## 3884  34047093694 4.436 2003     1208      1     3.392
## 3887  46749133770 4.109 2003     1208      1     3.123
## 3912  60949464452 3.780 2003     1208      1     2.794
## 3918  60949658016 5.148 2003     1208      1     4.104
## 3921  60949700952 4.436 2003     1208      1     3.392
## 3960  60950414028 4.651 2003     1208      1     3.665
## 3963  60950438763 4.034 2003     1208      1     3.048
## 3977  60950500372 4.249 2003     1208      1     3.263
## 3993  60950588015 4.314 2003     1208      1     3.328
## 3999  60950625661 5.403 2003     1208      1     4.417
## 4036  61249262249 4.034 2003     1208      1     3.048
## 4094  70449761302 3.580 2003     1208      1     2.594
## 4141   8544244308 4.547 2003     1208      1     3.561
## 4142  26944502435 4.044 2003     1208      2     3.000
## 4143  33646681740 6.083 2003     1208      2     5.039
## 4146  34547600256 4.341 2003     1208      2     3.297
```

```
## 4197  60950066571 4.869 2003     1208      2     3.825
## 4230  61149570569 3.528 2003     1208      2     2.542
## 4489  45449108735 5.113 2003     1208      1     4.069
## 4583  61949186737 4.181 2003     1208      1     3.195
## 4600  34748861400 4.651 2003     1208      1     3.607
## 4621  41549119242 4.109 2003     1208      1     3.123
## 4691  36248970780 3.954 2003     1208      1     2.910
## 4697  60949457140 8.312 2003     1208      1     7.326
## 4710  60950639325 3.780 2003     1208      1     2.736
## 4713  60950694241 3.954 2003     1208      1     2.968
## 4736  61249635138 4.109 2003     1208      1     3.065
## 4742  61449197953 4.922 2003     1208      1     3.936
## 4747  61449337582 4.837 2003     1208      1     3.793
## 4782  64949161384 3.580 2003     1208      1     2.536
## 4822  70449843968 3.683 2003     1208      1     2.639
## 4863  29744456134 5.218 2003     1208      2     4.232
## 4872  34547980224 3.959 2003     1208      2     2.973
## 4873  48549096636 5.652 2003     1208      2     4.608
## 4888  61249716500 4.045 2003     1208      2     3.059
## 4990  70449856706 3.528 2003     1208      2     2.542
## 5092  61949103780 3.552 2003     1208      3     2.508
## 5274  28244501538 5.482 2004     1208      1     4.428
## 5280  34748922675 4.292 2004     1208      1     3.180
## 5307  60949771564 3.622 2004     1208      1     2.568
## 5445  61949116307 3.947 2004     1208      1     2.893
## 5463  67049083067 4.804 2004     1208      1     3.750
## 5468  67649743610 4.213 2004     1208      1     3.101
## 5514  34250312985 4.933 2004     1208      2     3.821
## 5562  60950391613 4.521 2004     1208      2     3.409
## 5581  60950719934 3.827 2004     1208      2     2.773
## 5935  60949318761 3.739 2004     1208      1     2.918
## 6108  33750435368 5.377 2004     1208      1     4.323
## 6112  33947671157 4.506 2004     1208      1     3.394
## 6126  60949382516 6.369 2004     1208      1     5.315
## 6152  61049449507 5.223 2004     1208      1     4.169
## 6178  61449455372 3.739 2004     1208      1     2.685
## 6196  62549140156 3.622 2004     1208      1     2.568
## 6270  71249085963 3.622 2004     1208      1     2.568
## 6285  60950338479 5.580 2004     1208      2     4.468
## 6299  61849165077 3.827 2004     1208      2     2.715
## 6350  70449786556 4.699 2004     1208      2     3.645
## 6352  70449824679 4.933 2004     1208      2     3.879
## 6699  34248708340 3.945 2005     1208      1     2.997
## 6705  34548494655 3.822 2005     1208      1     2.817
## 6729  60949655783 3.688 2005     1208      1     2.740
## 6765  60950409349 3.945 2005     1208      1     2.997
## 6790  60950596472 3.541 2005     1208      1     2.536
## 6800  60950708929 3.541 2005     1208      1     2.593
## 6850  61149580054 4.683 2005     1208      1     3.678
## 6901  65449143734 4.358 2005     1208      1     3.353
```

```
## 6902  65449165806 4.683 2005     1208      1     3.678
## 6909  67649784613 4.445 2005     1208      1     3.497
## 6947  70449899952 3.822 2005     1208      1     2.874
## 6955  70449970464 3.541 2005     1208      1     2.593
## 6966  70450058332 3.688 2005     1208      1     2.740
## 6985  33750587792 3.538 2005     1208      2     2.533
## 6989  34250008387 4.562 2005     1208      2     3.614
## 6994  38649134273 6.094 2005     1208      2     5.089
## 6999  60949254410 4.465 2005     1208      2     3.517
## 7041  60950274865 3.944 2005     1208      2     2.996
## 7043  60950350045 5.477 2005     1208      2     4.529
## 7070  60950562957 3.861 2005     1208      2     2.913
## 7078  60950642207 3.861 2005     1208      2     2.913
## 7177  33845250551 4.010 2005     1202      4     3.062
## 7380  62749191854 3.744 2005     1202      3     2.739
## 7383  63849109891 3.549 2005     1202      3     2.544
## 7553  60950504855 3.650 2005     1208      3     2.645
## 7604  61949241615 3.822 2005     1208      1     2.817
## 7630  61049327976 3.822 2005     1208      1     2.817
## 7650  21544433512 4.003 2005     1208      2     3.289
## 7723  33846337496 6.348 2005     1208      1     5.343
## 7724  46249120022 4.059 2005     1208      1     3.054
## 7725  46249120887 3.945 2005     1208      1     2.997
## 7734  60949594699 3.822 2005     1208      1     2.817
## 7742  60950233081 4.059 2005     1208      1     3.111
## 7743  60950413511 3.541 2005     1208      1     2.593
## 7755  60950658414 3.541 2005     1208      1     2.536
## 7804  61449502250 3.688 2005     1208      1     2.683
## 7832  67649792074 3.822 2005     1208      1     2.874
## 7837  67650088057 3.541 2005     1208      1     2.593
## 7844  67650178186 4.059 2005     1208      1     3.111
## 7845  68549117379 3.822 2005     1208      1     2.874
## 7868  70450064684 3.822 2005     1208      1     2.874
## 7876  70450093197 3.541 2005     1208      1     2.593
## 7898  44449106909 4.250 2005     1208      2     3.245
## 7902  60950423747 5.118 2005     1208      2     4.113
## 7914  61249595569 3.788 2005     1208      2     2.840
## 7915  61449306806 3.654 2005     1208      2     2.940
## 7919  61449561847 3.582 2005     1208      2     2.577
## 7937  62449266342 3.723 2005     1208      2     2.775
## 8040  33846333428 3.605 2005     1208      3     2.657
## 8313  50949119108 4.710 2006     1208      1     3.759
## 8336  60949594614 3.653 2006     1208      1     2.645
## 8347  60950223477 3.786 2006     1208      1     2.778
## 8356  60950442490 3.908 2006     1208      1     2.957
## 8368  60950493268 4.224 2006     1208      1     3.273
## 8369  60950493735 3.653 2006     1208      1     2.645
## 8373  60950504606 3.786 2006     1208      1     2.778
## 8394  60950726337 4.317 2006     1208      1     3.309
## 8444  61449369584 3.508 2006     1208      1     2.557
```

```
## 8484  67650109778 3.786 2006     1208      1     2.778
## 8486  67650128316 3.653 2006     1208      1     2.702
## 8487  67650135684 3.786 2006     1208      1     2.835
## 8490  67650164706 4.317 2006     1208      1     3.309
## 8529  70450105045 4.224 2006     1208      1     3.273
## 9018  60949478073 3.653 2006     1208      1     2.702
## 9237  61949128551 4.404 2006     1208      1     3.453
## 9296  84870103615 4.317 2006     1208      1     3.366
## 9311  42449161018 5.130 2006     1208      1     4.122
## 9313  53549130671 7.825 2006     1208      1     6.817
## 9314  55349128031 5.280 2006     1208      1     4.272
## 9325  60950413512 4.777 2006     1208      1     3.769
## 9326  60950445899 3.908 2006     1208      1     2.900
## 9327  60950622690 4.126 2006     1208      1     3.175
## 9342  61249099688 4.021 2006     1208      1     3.013
## 9344  61249305083 4.317 2006     1208      1     3.366
## 9351  61449427822 3.653 2006     1208      1     2.645
## 9352  61449427823 4.126 2006     1208      1     3.118
## 9361  61949197756 4.021 2006     1208      1     3.070
## 9365  62349122821 3.653 2006     1208      1     2.645
## 9367  62449189059 4.777 2006     1208      1     3.826
## 9371  62749148771 4.777 2006     1208      1     3.826
## 9381  64949139647 4.224 2006     1208      1     3.216
## 9426  67650079640 3.908 2006     1208      1     2.900
## 9440  70449824307 3.786 2006     1208      1     2.835
## 9465  70450032962 3.786 2006     1208      1     2.544
## 9468  70450045620 4.126 2006     1208      1     3.118
## 9469  70450058396 3.786 2006     1208      1     2.835
## 9479  70749096336 4.021 2006     1208      1     3.013
## 9491  33846152354 3.582 2006     1208      2     2.631
## 9497  34249084654 5.314 2006     1208      2     4.306
## 9503  44449151884 6.523 2006     1208      2     5.515
## 9504  60949270586 3.627 2006     1208      2     2.676
## 9524  61449282187 3.627 2006     1208      2     2.676
## 9541  61449530625 6.523 2006     1208      2     5.572
## 9574  70349296952 3.627 2006     1208      2     2.676
## 9577  70449775017 4.234 2006     1208      2     3.283
## 9589  70450078889 3.627 2006     1208      2     2.619
## 9803  57749157632 6.254 2007     1208      1     5.277
## 9867  60950549890 3.610 2007     1208      1     2.633
## 9899  61149182825 4.555 2007     1208      1     3.635
## 9949  62449306099 4.555 2007     1208      1     3.635
## 9967  66049152167 4.214 2007     1208      1     3.237
## 9978  67650155463 3.783 2007     1208      1     2.806
## 9981  67650162724 3.610 2007     1208      1     2.633
## 9982  67650168953 3.939 2007     1208      1     2.962
## 10006 70449930640 3.610 2007     1208      1     2.690
## 10040 84868448556 3.610 2007     1208      1     2.690
## 10075 43249135021 4.018 2007     1208      2     3.041
## 10078 43249146684 4.726 2007     1208      2     3.749
```

```
## 10107 60949438530 4.726 2007     1208      2     3.806
## 10155 60950643835 4.839 2007     1208      2     3.919
## 10158 60950662433 3.834 2007     1208      2     2.914
## 10415 61949141429 3.441 2007     1208      3     2.521
## 10670 61049446035 5.641 2007     1208      1     4.721
## 10676 70450032239 3.610 2007     1208      1     2.690
## 10809 70450064287 5.151 2007     1208      1     4.231
## 10866 60949326226 3.610 2007     1208      1     2.633
## 10867 60949370742 5.976 2007     1208      1     5.056
## 10880 60950131470 3.939 2007     1208      1     2.962
## 10892 60950536451 4.449 2007     1208      1     3.472
## 10925 62449170699 5.841 2007     1208      1     4.864
## 10968 67649683064 4.214 2007     1208      1     3.294
## 10984 67650096553 4.082 2007     1208      1     3.162
## 10985 67650143129 5.474 2007     1208      1     4.554
## 10986 67650925409 3.939 2007     1208      1     3.019
## 10988 70249093068 3.610 2007     1208      1     2.690
## 10992 70449859265 3.610 2007     1208      1     2.690
## 10994 70449877375 3.610 2007     1208      1     2.690
## 11007 70450015708 6.179 2007     1208      1     5.259
## 11014 70450064757 3.783 2007     1208      1     2.806
## 11015 70450077380 4.654 2007     1208      1     3.734
## 11112 62449186372 3.857 2007     1208      2     2.880
## 11165 70450078906 4.018 2007     1208      2     3.098
## 11217 61249239835 4.188 2007     1208      3     3.268
## 11228 61249660537 3.741 2007     1208      3     2.821
## 11409 60949485276 3.367 2008     1208      1     2.509
## 11434 60950624831 3.802 2008     1208      1     2.944
## 11478 60950719981 5.997 2008     1208      1     5.081
## 11494 60950608143 3.692 2008     1203      3     2.543
## 11532 60950299886 3.598 2008     1208      1     2.740
## 11540 57749125280 4.798 2008     1208      1     3.940
## 11548 60949270646 4.686 2008     1208      1     3.828
## 11573 61049554358 3.802 2008     1208      1     2.944
## 11605 61949182548 3.598 2008     1208      1     2.682
## 11611 61949371494 3.367 2008     1208      1     2.509
## 11618 62449098267 5.426 2008     1208      1     4.568
## 11624 62449159076 3.598 2008     1208      1     2.740
## 11626 62449175332 5.942 2008     1208      1     5.026
## 11633 62449269964 3.984 2008     1208      1     3.126
## 11641 62949147900 4.300 2008     1208      1     3.384
## 11644 62949210835 3.367 2008     1208      1     2.509
## 11672 67650099391 3.984 2008     1208      1     3.068
## 11676 67650112946 3.984 2008     1208      1     3.068
## 11681 67650132635 4.149 2008     1208      1     3.233
## 11684 67650151239 3.984 2008     1208      1     3.068
## 11687 67650155476 3.367 2008     1208      1     2.509
## 11688 67650162729 4.798 2008     1208      1     3.882
## 11691 67650178260 5.001 2008     1208      1     4.085
## 11695 70449752752 4.300 2008     1208      1     3.442
```

```
## 11696 70449780196 3.367 2008     1208      1     2.509
## 11723 70450068192 3.367 2008     1208      1     2.509
## 11792 61149209901 3.427 2008     1208      2     2.569
## 11796 61149472770 4.222 2008     1208      2     3.306
## 12102 65749261392 3.585 2008     1202      3     2.669
## 12226 70450003031 5.184 2008     1208      1     4.268
## 12362 70450082589 3.516 2008     1205      3     2.600
## 12399 45349087542 3.582 2008     1208      2     2.724
## 12435 61249636505 6.426 2008     1208      1     5.568
## 12468 60949164169 3.720 2008     1208      3     2.862
## 12498 70449866030 3.802 2008     1208      1     2.886
## 12500 70450060571 4.686 2008     1208      1     3.770
## 12551 61049119505 3.367 2008     1208      1     2.509
## 12629 57749127836 3.984 2008     1208      1     3.126
## 12658 61249639675 3.598 2008     1208      1     2.682
## 12661 61449513631 3.984 2008     1208      1     3.126
## 12664 61949197772 4.149 2008     1208      1     3.291
## 12671 62449262219 4.300 2008     1208      1     3.442
## 12675 62749110813 6.152 2008     1208      1     5.294
## 12677 62749122811 5.268 2008     1208      1     4.410
## 12679 62749136269 3.984 2008     1208      1     3.126
## 12682 62749151909 4.798 2008     1208      1     3.940
## 12683 62749155048 4.149 2008     1208      1     3.233
## 12684 62749156172 5.942 2008     1208      1     5.084
## 12691 64949093382 3.598 2008     1208      1     2.682
## 12696 64949107038 3.802 2008     1208      1     2.886
## 12700 64949153413 4.300 2008     1208      1     3.442
## 12727 67650070793 4.149 2008     1208      1     3.291
## 12735 67650090943 5.570 2008     1208      1     4.712
## 12739 70449931160 5.268 2008     1208      1     4.352
## 12746 78650710041 3.802 2008     1208      1     2.886
## 12795 61249468411 4.222 2008     1208      2     3.364
## 12797 61249548069 3.427 2008     1208      2     2.569
## 12861 40449115608 3.564 2008     1202      4     2.939
## 13002 85014946308 3.363 2008     1203      4     2.505
## 13097 77950310142 4.054 2009     1208      1     3.287
## 13214 84884186648 4.054 2009     1208      1     3.344
## 13249 61949462044 3.777 2009     1208      3     3.067
## 13400 73949133087 4.425 2009     1208      1     3.715
## 13418 70349840389 4.249 2009     1208      1     3.539
## 13474 68549123389 3.591 2009     1208      1     2.881
## 13490 84867054792 3.306 2009     1208      1     2.830
## 13559 73649148390 3.360 2009     1202      4     2.650
## 13613 67649716605 3.591 2009     1208      1     2.824
## 13663 67649950267 4.054 2009     1208      1     3.344
## 13679 77950325600 4.997 2009     1208      1     4.287
## 13680 77950788847 3.306 2009     1208      1     2.596
## 13683 77950809870 4.054 2009     1208      1     3.054
## 13684 77950815002 6.760 2009     1208      1     5.993
## 13685 77950822756 5.116 2009     1208      1     4.349
```

```
## 13686 77950831009 4.585 2009     1208      1     3.875
## 13688 77950841010 6.560 2009     1208      1     5.793
## 13693 68949083793 3.825 2009     1208      2     3.115
## 13695 68949090683 4.540 2009     1208      2     3.830
## 13697 68949092620 3.727 2009     1208      2     3.017
## 13700 68949123188 3.423 2009     1208      2     2.713
## 13702 68949135917 3.636 2009     1208      2     2.926
## 13893 79958822732 4.054 2009     1208      1     3.287
## 13895 80054949947 3.837 2009     1208      1     3.070
## 13939 77950229849 3.223 2009     1208      2     2.513
## 13940 77950253669 3.223 2009     1208      2     2.513
## 14020 65849111025 3.591 2009     1208      1     2.824
## 14021 65849141252 3.306 2009     1208      1     2.596
## 14039 77956351978 3.837 2009     1208      1     3.070
## 14092 77950181710 4.785 2009     1205      2     4.018
## 14275 77950904394 3.302 2010     1208      1     2.629
## 14278 78649401790 3.586 2010     1208      1     2.856
## 14282 78649430807 3.586 2010     1208      1     2.856
## 14293 84855406055 3.832 2010     1208      1     3.102
## 14542 77957559973 4.049 2010     1208      1     3.376
## 14543 77957584700 3.302 2010     1208      1     2.572
## 14572 77956632352 3.506 2010     1208      2     2.776
## 14575 78649656075 4.243 2010     1208      1     3.513
## 14591 77956331789 4.376 2010     1208      2     3.703
## 14712 79959202296 3.634 2010     1208      2     2.961
## 14810 77950211324 3.302 2010     1208      1     2.629
## 14880 79956372974 3.668 2010     1202      4     2.995
## 14898 77949582311 4.243 2010     1208      1     3.513
## 14902 77949615907 3.586 2010     1208      1     2.913
## 14927 77950140620 3.832 2010     1208      1     3.159
## 14928 77950253361 3.302 2010     1208      1     2.629
## 14934 77950886825 3.586 2010     1208      1     2.856
## 14935 77950888966 3.586 2010     1208      1     2.856
## 15039 76349106441 3.586 2010     1208      1     2.856
## 15050 77954443733 3.302 2010     1208      1     2.572
## 15094 77950287584 3.586 2010     1208      1     2.913
## 15120 79955418712 4.243 2010     1208      1     3.513
## 15143 73249128238 4.062 2010     1208      2     3.622
## 15226 77956500865 4.776 2010     1208      3     4.103
## 15228 78649472664 5.064 2010     1208      3     4.100
## 15279 82455228764 3.393 2010     1208      3     2.720
## 15298 77951908318 3.254 2010     1202      6     2.581
## 15358 84855209572 3.491 2011     1208      1     2.756
## 15359 84855210092 3.491 2011     1208      1     2.698
## 15365 84855762177 3.887 2011     1208      1     3.152
## 15368 84855797090 3.491 2011     1208      1     2.756
## 15380 84859358430 3.491 2011     1208      1     2.756
## 15382 84860778742 5.724 2011     1208      1     4.989
## 15403 84870329240 3.491 2011     1208      1     2.756
## 15411 84872712403 3.491 2011     1208      1     2.698
```

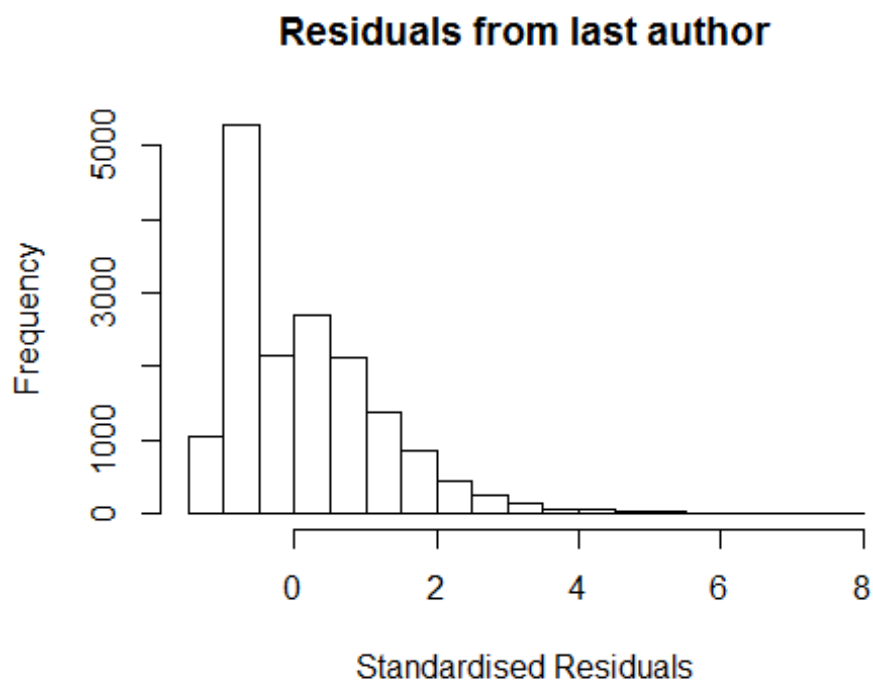
```
## 15419 84873283995 3.491 2011     1208      1     2.698
## 15424 84873302367 3.887 2011     1208      1     3.152
## 15447 84884198571 3.887 2011     1208      1     3.152
## 15465 84887967106 3.491 2011     1208      1     2.756
## 15474 84855781328 3.573 2011     1208      2     2.838
## 15475 84855791481 5.406 2011     1208      2     4.613
## 15484 84857779496 3.573 2011     1208      2     2.547
## 15615 84857935340 3.474 2011     1202      6     2.739
## 15674 80155194058 3.906 2011     1202      5     3.113
## 15694 80054789940 4.221 2011     1208      1     3.428
## 15698 80054708194 3.887 2011     1208      1     3.152
## 15717 80053615731 4.511 2011     1208      1     3.776
## 15803 80052975515 3.243 2011     1205      2     2.508
## 15826 80052542479 3.887 2011     1208      1     3.152
## 15845 84055192224 6.801 2011     1208      1     6.008
## 15847 84856196552 3.887 2011     1208      1     3.152
## 15853 84862322836 3.887 2011     1208      1     3.094
## 15869 84883480879 4.995 2011     1208      1     4.202
## 15899 84861070464 5.091 2011     1208      2     4.356
## 16068 79959952724 4.221 2011     1208      1     3.428
## 16069 79959977747 3.491 2011     1208      1     2.698
## 16212 79955377318 6.894 2011     1208      1     6.101
## 16216 79955439141 4.221 2011     1208      1     3.486
## 16219 79957888983 3.887 2011     1208      1     3.152
## 16290 79958846842 3.887 2011     1208      1     3.152
## 16360 79955674091 3.677 2011     1203      6     2.884
## 16379 79955677215 6.146 2011     1208      1     5.353
## 16383 80054915853 3.887 2011     1208      1     3.094
## 16498 79955368617 3.375 2011     1203      3     2.582
## 16516 78650843435 5.564 2011     1208      1     4.829
## 16517 78650845690 4.221 2011     1208      1     3.486
## 16541 79960860760 4.221 2011     1208      1     3.428
## 16542 79960871745 4.221 2011     1208      1     3.486
## 16550 81355124072 3.887 2011     1208      1     3.094
## 16567 84055198510 3.887 2011     1208      1     3.094
## 16586 85011484631 3.887 2011     1208      1     3.094
## 16615 80555156002 3.648 2011     1208      2     2.855
## 16630 84555179123 4.538 2011     1208      2     3.803
## 16726 82455232963 4.204 2011     1208      3     3.411
## 16727 82455244076 4.997 2011     1208      3     4.262
## 16835 84871206897 3.953 2012     1208      2     3.313
## 16837 84871229041 3.842 2012     1208      2     3.202
## 16880 84870331366 3.391 2012     1208      1     2.751
## 16882 84870350788 6.781 2012     1208      1     6.141
## 16892 84855231775 3.391 2012     1208      1     2.693
## 16896 84856905745 3.391 2012     1208      1     2.751
## 16898 84856945735 4.382 2012     1208      1     3.742
## 16903 84871438787 3.937 2012     1208      1     3.239
## 16909 84871985774 3.391 2012     1208      1     2.751
## 16914 84872706828 3.391 2012     1208      1     2.751
```

```
## 16926 84873281046 3.937 2012     1208      1     3.297
## 16949 84874079832 7.447 2012     1208      1     6.749
## 16950 84874095385 3.937 2012     1208      1     3.239
## 16957 84876994217 5.086 2012     1208      1     4.388
## 16958 84876998823 3.391 2012     1208      1     2.693
## 16959 84877000419 3.937 2012     1208      1     3.239
## 16960 84878432148 4.382 2012     1208      1     3.742
## 17050 84871269332 3.156 2012     1208      2     2.516
## 17170 84871367509 3.245 2012     1202      3     2.547
## 17171 84871394290 5.615 2012     1202      3     4.917
## 17215 84874341423 3.664 2012     1202      3     3.024
## 17318 84869003069 5.086 2012     1208      1     4.446
## 17320 84869033091 3.937 2012     1208      1     3.239
## 17326 84868366032 3.394 2012     1208      2     2.754
## 17327 84868368393 3.394 2012     1208      2     2.754
## 17375 84874056751 3.264 2012     1202      4     2.566
## 17431 84871011871 3.937 2012     1208      1     3.297
## 17462 84867923419 3.313 2012     1208      2     2.906
## 17464 84868631017 3.559 2012     1208      2     2.861
## 17465 84868671128 3.962 2012     1208      2     3.322
## 17569 84865697715 3.939 2012     1208      2     3.299
## 17570 84865700895 4.859 2012     1208      2     4.161
## 17571 84865204088 3.937 2012     1208      1     3.297
## 17574 84866596965 4.759 2012     1208      1     4.119
## 17578 84866635016 3.937 2012     1208      1     3.297
## 17594 84867391849 3.391 2012     1208      1     2.693
## 17604 84870451817 3.937 2012     1208      1     3.297
## 17607 84870509187 4.382 2012     1208      1     3.742
## 17616 84871215069 3.391 2012     1208      1     2.693
## 17617 84871229501 3.937 2012     1208      1     3.297
## 17622 84871342985 5.374 2012     1208      1     4.676
## 17624 84871749299 3.391 2012     1208      1     2.693
## 17631 84873941703 3.937 2012     1208      1     3.297
## 17632 84873952317 3.391 2012     1208      1     2.751
## 17633 84873956808 3.937 2012     1208      1     3.239
## 17668 84885611499 5.086 2012     1208      1     4.446
## 17727 84865713540 3.451 2012     1202      3     2.811
## 17750 84866640267 5.127 2012     1208      3     4.429
## 17762 84866982605 3.164 2012     1203      3     2.524
## 17828 84864711707 3.258 2012     1202      5     2.618
## 17849 84865371557 5.086 2012     1208      1     4.388
## 17850 84865386536 6.781 2012     1208      1     6.141
## 17858 84865347957 3.391 2012     1208      1     2.693
## 17863 84865160135 5.086 2012     1208      1     4.446
## 17864 84865162714 5.086 2012     1208      1     4.388
## 17874 84864926904 4.302 2012     1208      2     3.662
## 17897 84883323231 3.240 2012     1208      2     2.600
## 17928 84870409111 3.654 2012     1203      3     3.247
## 17953 84864349568 5.374 2012     1208      1     4.676
## 17968 84874377033 4.759 2012     1208      1     4.119
```

```
## 17987 84864035512 4.620 2012     1208      2     3.980
## 18033 84864201531 3.937 2012     1208      1     3.239
## 18034 84864215638 6.455 2012     1208      1     5.757
## 18035 84864219524 5.086 2012     1208      1     4.388
## 18036 84864248050 5.374 2012     1208      1     4.676
## 18037 84864268571 6.455 2012     1208      1     5.815
## 18043 84865177201 3.937 2012     1208      1     3.297
## 18054 84875659820 5.632 2012     1208      1     4.992
## 18091 84866718537 4.477 2012     1208      2     3.837
## 18215 84859330957 5.086 2012     1208      1     4.446
## 18233 84861746194 3.412 2012     1208      2     2.772
## 18242 84883352820 3.619 2012     1208      2     2.979
## 18247 84883394654 3.596 2012     1208      2     2.898
## 18259 84872852785 3.240 2012     1203      3     2.600
## 18266 84872871724 3.721 2012     1203      3     3.023
## 18281 84859745370 4.759 2012     1208      1     4.061
## 18284 84859760309 3.391 2012     1208      1     2.751
## 18296 84876226058 4.759 2012     1208      1     4.061
## 18414 84865691275 3.391 2012     1208      1     2.751
## 18427 84884668586 3.391 2012     1208      1     2.751
## 18448 84861350433 4.598 2012     1208      2     3.900
## 18449 84861359249 3.559 2012     1208      2     2.861
## 18452 84861395909 3.559 2012     1208      2     2.861
## 18480 84860734606 3.999 2012     1202      3     3.301
## 18489 84862115131 3.380 2012     1202      4     2.682
## 18490 84862116557 3.380 2012     1202      4     2.740
## 18555 84856799214 3.937 2012     1208      1     3.239
## 18557 84856642915 4.382 2012     1208      1     3.684
## 18558 84856653530 4.382 2012     1208      1     3.684
## 18565 84856448236 3.227 2012     1202      3     2.587
## 18566 84856468758 4.331 2012     1202      3     3.633
## 18567 84856478984 3.344 2012     1202      3     2.646
## 18576 84864376206 4.361 2012     1208      2     3.721
## 18632 84856135031 3.937 2012     1208      1     3.239
## 18636 84856185796 4.759 2012     1208      1     4.119
## 18637 84856189103 4.759 2012     1208      1     4.119
## 18638 84856197118 6.781 2012     1208      1     6.141
## 18648 84863619762 3.391 2012     1208      1     2.751
## 18664 84867394653 3.937 2012     1208      1     3.297
## 18737 84865710516 3.559 2012     1208      2     2.861
## 18740 84867763226 3.284 2012     1205      2     2.644
## 18745 84870312908 3.962 2012     1208      2     3.322
## 18747 84870334955 6.510 2012     1208      2     5.812
## 18751 84870345747 6.391 2012     1208      2     5.751
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
```

```
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.1058 -0.7642 -0.0326  0.7481  7.7684
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)          0.7756     0.0495   15.67  < 2e-16 ***
## FirstAuthorFemale1   0.0491     0.0161    3.06  0.00222 **
## Year1997             0.0140     0.0660    0.21  0.83200
## Year1998             0.0513     0.0654    0.79  0.43218
## Year1999            -0.0253     0.0617   -0.41  0.68176
## Year2000             0.0480     0.0637    0.75  0.45096
## Year2001             0.0159     0.0616    0.26  0.79639
## Year2002             0.2131     0.0589    3.62  0.00030 ***
## Year2003             0.2145     0.0581    3.69  0.00022 ***
## Year2004             0.2811     0.0581    4.84  1.3e-06 ***
## Year2005             0.1753     0.0577    3.04  0.00239 **
## Year2006             0.1805     0.0577    3.13  0.00175 **
## Year2007             0.1474     0.0567    2.60  0.00930 **
## Year2008             0.0875     0.0562    1.56  0.11931
## Year2009            -0.0606     0.0569   -1.06  0.28758
## Year2010            -0.0992     0.0569   -1.74  0.08126 .
## Year2011            -0.0362     0.0563   -0.64  0.51998
## Year2012            -0.1309     0.0544   -2.41  0.01607 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.919
## Multiple R-squared:  0.0189, Adjusted R-squared:  0.0178
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
##  64 observations
c(2421,3496,3615,4107,4269,4621,5346,5362,5507,6122,6167,6744,8136,8301,8307,
8344,8580,9339,9506,9559,9611,9629,10043,10161,10169,10858,11065,11067,11072,
11116,11120,11941,11942,11945,13274,13395,13479,13798,13846,14110,14249,14369
,14684,14748,14755,14952,15080,15341,15378,15452,15539,15540,15551,15552,1563
0,15696,15697,15698,15699,15715,15863,16233,16332,16334)
##   are outliers with |weight| = 0 ( < 6.1e-06);
##  1235 weights are ~= 1. The remaining 15195 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0001  0.8940  0.9340  0.8910  0.9700  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          6.06e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
```

```
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                    cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##                   GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.005  1           1.002
## Year             1.005 16           1.000
```




```
## [1] "List of  519 outliers with residuals above  2.5"
##          ScopusId  NLCS Year OneField Fields residuals
## 94    67651037634 3.400 1996     1208      1     2.626
## 138   61449406492 3.938 1996     1208      1     3.164
## 140   61949400015 3.345 1996     1208      1     2.514
## 376    0009431359 3.532 1997     1208      2     2.687
## 412   79959433454 3.987 1997     1208      2     3.200
## 952    0000490567 3.695 1998     1208      2     2.872
## 1207  60949584988 3.410 1999     1208      1     2.664
## 1485   0013039289 3.420 1999     1208      3     2.617
## 2557  10844245985 5.274 2002     1208      1     4.289
## 2562  33750620950 3.855 2002     1208      1     2.813
## 2563  33947361251 3.629 2002     1208      1     2.587
## 2565  34547299083 3.746 2002     1208      1     2.761
```

```
## 2566  47749108055 4.222 2002     1208      1     3.237
## 2581  60949458107 4.222 2002     1208      1     3.237
## 2605  60950002681 3.502 2002     1208      1     2.517
## 2607  60950027813 3.746 2002     1208      1     2.761
## 2619  60950346555 4.918 2002     1208      1     3.933
## 2732  65849154249 3.629 2002     1208      1     2.587
## 2733  65849216639 3.629 2002     1208      1     2.644
## 2734  65849259075 3.629 2002     1208      1     2.587
## 2736  65849326957 3.955 2002     1208      1     2.913
## 2748  67650096797 3.746 2002     1208      1     2.761
## 2751  70249100469 4.301 2002     1208      1     3.316
## 2752  70349438078 3.502 2002     1208      1     2.517
## 2780   3142697241 8.757 2002     1208      2     7.772
## 2781  33646911853 5.146 2002     1208      2     4.104
## 2783  33748905673 5.092 2002     1208      2     4.107
## 2789  44149119239 3.798 2002     1208      2     2.756
## 2825  60949601992 4.417 2002     1208      2     3.375
## 2860  60950259244 4.417 2002     1208      2     3.375
## 3194  61949270383 3.955 2002     1208      1     2.970
## 3322  34248720247 3.855 2002     1208      1     2.870
## 3323  34248734400 3.855 2002     1208      1     2.870
## 3326  38749133094 4.050 2002     1208      1     3.065
## 3327  58449103137 3.502 2002     1208      1     2.517
## 3332  60949291781 3.629 2002     1208      1     2.587
## 3342  60950305286 3.746 2002     1208      1     2.761
## 3345  60950448412 3.629 2002     1208      1     2.644
## 3347  60950498847 3.746 2002     1208      1     2.761
## 3348  60950501375 3.955 2002     1208      1     2.913
## 3352  60950608189 3.955 2002     1208      1     2.913
## 3354  60950639390 3.855 2002     1208      1     2.813
## 3355  60950641233 4.515 2002     1208      1     3.473
## 3358  60950680913 3.629 2002     1208      1     2.644
## 3361  60950699119 3.855 2002     1208      1     2.813
## 3380  61249427664 3.502 2002     1208      1     2.517
## 3382  61249520644 3.955 2002     1208      1     2.913
## 3439  70349348163 4.050 2002     1208      1     3.065
## 3440  70349378641 3.629 2002     1208      1     2.587
## 3460  70450000901 3.629 2002     1208      1     2.644
## 3504  61249583521 3.963 2002     1208      2     2.978
## 3509  61449442102 3.533 2002     1208      2     2.548
## 3515  62449109752 4.033 2002     1208      2     2.991
## 3516  62449174089 3.672 2002     1208      2     2.687
## 3517  62449296550 4.063 2002     1208      2     3.021
## 3657  60950279634 3.963 2002     1208      3     2.921
## 3793  85012444559 3.746 2002     1202      3     2.761
## 3843   0036526749 3.670 2002     1202      6     2.628
## 3845   1442335800 3.694 2002     1202      6     2.709
## 3884  34047093694 4.436 2003     1208      1     3.392
## 3887  46749133770 4.109 2003     1208      1     3.123
## 3912  60949464452 3.780 2003     1208      1     2.794
```

```
## 3918  60949658016 5.148 2003     1208      1     4.104
## 3921  60949700952 4.436 2003     1208      1     3.392
## 3960  60950414028 4.651 2003     1208      1     3.665
## 3963  60950438763 4.034 2003     1208      1     3.048
## 3977  60950500372 4.249 2003     1208      1     3.263
## 3993  60950588015 4.314 2003     1208      1     3.328
## 3999  60950625661 5.403 2003     1208      1     4.417
## 4036  61249262249 4.034 2003     1208      1     3.048
## 4094  70449761302 3.580 2003     1208      1     2.594
## 4141   8544244308 4.547 2003     1208      1     3.561
## 4142  26944502435 4.044 2003     1208      2     3.000
## 4143  33646681740 6.083 2003     1208      2     5.039
## 4146  34547600256 4.341 2003     1208      2     3.297
## 4197  60950066571 4.869 2003     1208      2     3.825
## 4230  61149570569 3.528 2003     1208      2     2.542
## 4489  45449108735 5.113 2003     1208      1     4.069
## 4583  61949186737 4.181 2003     1208      1     3.195
## 4600  34748861400 4.651 2003     1208      1     3.607
## 4621  41549119242 4.109 2003     1208      1     3.123
## 4691  36248970780 3.954 2003     1208      1     2.910
## 4697  60949457140 8.312 2003     1208      1     7.326
## 4710  60950639325 3.780 2003     1208      1     2.736
## 4713  60950694241 3.954 2003     1208      1     2.968
## 4736  61249635138 4.109 2003     1208      1     3.065
## 4742  61449197953 4.922 2003     1208      1     3.936
## 4747  61449337582 4.837 2003     1208      1     3.793
## 4782  64949161384 3.580 2003     1208      1     2.536
## 4822  70449843968 3.683 2003     1208      1     2.639
## 4863  29744456134 5.218 2003     1208      2     4.232
## 4872  34547980224 3.959 2003     1208      2     2.973
## 4873  48549096636 5.652 2003     1208      2     4.608
## 4888  61249716500 4.045 2003     1208      2     3.059
## 4990  70449856706 3.528 2003     1208      2     2.542
## 5092  61949103780 3.552 2003     1208      3     2.508
## 5274  28244501538 5.482 2004     1208      1     4.428
## 5280  34748922675 4.292 2004     1208      1     3.180
## 5307  60949771564 3.622 2004     1208      1     2.568
## 5445  61949116307 3.947 2004     1208      1     2.893
## 5463  67049083067 4.804 2004     1208      1     3.750
## 5468  67649743610 4.213 2004     1208      1     3.101
## 5514  34250312985 4.933 2004     1208      2     3.821
## 5562  60950391613 4.521 2004     1208      2     3.409
## 5581  60950719934 3.827 2004     1208      2     2.773
## 5935  60949318761 3.739 2004     1208      1     2.918
## 6108  33750435368 5.377 2004     1208      1     4.323
## 6112  33947671157 4.506 2004     1208      1     3.394
## 6126  60949382516 6.369 2004     1208      1     5.315
## 6152  61049449507 5.223 2004     1208      1     4.169
## 6178  61449455372 3.739 2004     1208      1     2.685
## 6196  62549140156 3.622 2004     1208      1     2.568
```

```
## 6270  71249085963 3.622 2004     1208      1     2.568
## 6285  60950338479 5.580 2004     1208      2     4.468
## 6299  61849165077 3.827 2004     1208      2     2.715
## 6350  70449786556 4.699 2004     1208      2     3.645
## 6352  70449824679 4.933 2004     1208      2     3.879
## 6699  34248708340 3.945 2005     1208      1     2.997
## 6705  34548494655 3.822 2005     1208      1     2.817
## 6729  60949655783 3.688 2005     1208      1     2.740
## 6765  60950409349 3.945 2005     1208      1     2.997
## 6790  60950596472 3.541 2005     1208      1     2.536
## 6800  60950708929 3.541 2005     1208      1     2.593
## 6850  61149580054 4.683 2005     1208      1     3.678
## 6901  65449143734 4.358 2005     1208      1     3.353
## 6902  65449165806 4.683 2005     1208      1     3.678
## 6909  67649784613 4.445 2005     1208      1     3.497
## 6947  70449899952 3.822 2005     1208      1     2.874
## 6955  70449970464 3.541 2005     1208      1     2.593
## 6966  70450058332 3.688 2005     1208      1     2.740
## 6985  33750587792 3.538 2005     1208      2     2.533
## 6989  34250008387 4.562 2005     1208      2     3.614
## 6994  38649134273 6.094 2005     1208      2     5.089
## 6999  60949254410 4.465 2005     1208      2     3.517
## 7041  60950274865 3.944 2005     1208      2     2.996
## 7043  60950350045 5.477 2005     1208      2     4.529
## 7070  60950562957 3.861 2005     1208      2     2.913
## 7078  60950642207 3.861 2005     1208      2     2.913
## 7177  33845250551 4.010 2005     1202      4     3.062
## 7380  62749191854 3.744 2005     1202      3     2.739
## 7383  63849109891 3.549 2005     1202      3     2.544
## 7553  60950504855 3.650 2005     1208      3     2.645
## 7604  61949241615 3.822 2005     1208      1     2.817
## 7630  61049327976 3.822 2005     1208      1     2.817
## 7650  21544433512 4.003 2005     1208      2     3.289
## 7723  33846337496 6.348 2005     1208      1     5.343
## 7724  46249120022 4.059 2005     1208      1     3.054
## 7725  46249120887 3.945 2005     1208      1     2.997
## 7734  60949594699 3.822 2005     1208      1     2.817
## 7742  60950233081 4.059 2005     1208      1     3.111
## 7743  60950413511 3.541 2005     1208      1     2.593
## 7755  60950658414 3.541 2005     1208      1     2.536
## 7804  61449502250 3.688 2005     1208      1     2.683
## 7832  67649792074 3.822 2005     1208      1     2.874
## 7837  67650088057 3.541 2005     1208      1     2.593
## 7844  67650178186 4.059 2005     1208      1     3.111
## 7845  68549117379 3.822 2005     1208      1     2.874
## 7868  70450064684 3.822 2005     1208      1     2.874
## 7876  70450093197 3.541 2005     1208      1     2.593
## 7898  44449106909 4.250 2005     1208      2     3.245
## 7902  60950423747 5.118 2005     1208      2     4.113
## 7914  61249595569 3.788 2005     1208      2     2.840
```

```
## 7915  61449306806 3.654 2005     1208      2     2.940
## 7919  61449561847 3.582 2005     1208      2     2.577
## 7937  62449266342 3.723 2005     1208      2     2.775
## 8040  33846333428 3.605 2005     1208      3     2.657
## 8313  50949119108 4.710 2006     1208      1     3.759
## 8336  60949594614 3.653 2006     1208      1     2.645
## 8347  60950223477 3.786 2006     1208      1     2.778
## 8356  60950442490 3.908 2006     1208      1     2.957
## 8368  60950493268 4.224 2006     1208      1     3.273
## 8369  60950493735 3.653 2006     1208      1     2.645
## 8373  60950504606 3.786 2006     1208      1     2.778
## 8394  60950726337 4.317 2006     1208      1     3.309
## 8444  61449369584 3.508 2006     1208      1     2.557
## 8484  67650109778 3.786 2006     1208      1     2.778
## 8486  67650128316 3.653 2006     1208      1     2.702
## 8487  67650135684 3.786 2006     1208      1     2.835
## 8490  67650164706 4.317 2006     1208      1     3.309
## 8529  70450105045 4.224 2006     1208      1     3.273
## 9018  60949478073 3.653 2006     1208      1     2.702
## 9237  61949128551 4.404 2006     1208      1     3.453
## 9296  84870103615 4.317 2006     1208      1     3.366
## 9311  42449161018 5.130 2006     1208      1     4.122
## 9313  53549130671 7.825 2006     1208      1     6.817
## 9314  55349128031 5.280 2006     1208      1     4.272
## 9325  60950413512 4.777 2006     1208      1     3.769
## 9326  60950445899 3.908 2006     1208      1     2.900
## 9327  60950622690 4.126 2006     1208      1     3.175
## 9342  61249099688 4.021 2006     1208      1     3.013
## 9344  61249305083 4.317 2006     1208      1     3.366
## 9351  61449427822 3.653 2006     1208      1     2.645
## 9352  61449427823 4.126 2006     1208      1     3.118
## 9361  61949197756 4.021 2006     1208      1     3.070
## 9365  62349122821 3.653 2006     1208      1     2.645
## 9367  62449189059 4.777 2006     1208      1     3.826
## 9371  62749148771 4.777 2006     1208      1     3.826
## 9381  64949139647 4.224 2006     1208      1     3.216
## 9426  67650079640 3.908 2006     1208      1     2.900
## 9440  70449824307 3.786 2006     1208      1     2.835
## 9465  70450032962 3.786 2006     1208      1     2.544
## 9468  70450045620 4.126 2006     1208      1     3.118
## 9469  70450058396 3.786 2006     1208      1     2.835
## 9479  70749096336 4.021 2006     1208      1     3.013
## 9491  33846152354 3.582 2006     1208      2     2.631
## 9497  34249084654 5.314 2006     1208      2     4.306
## 9503  44449151884 6.523 2006     1208      2     5.515
## 9504  60949270586 3.627 2006     1208      2     2.676
## 9524  61449282187 3.627 2006     1208      2     2.676
## 9541  61449530625 6.523 2006     1208      2     5.572
## 9574  70349296952 3.627 2006     1208      2     2.676
## 9577  70449775017 4.234 2006     1208      2     3.283
```

```
## 9589  70450078889 3.627 2006     1208      2     2.619
## 9803  57749157632 6.254 2007     1208      1     5.277
## 9867  60950549890 3.610 2007     1208      1     2.633
## 9899  61149182825 4.555 2007     1208      1     3.635
## 9949  62449306099 4.555 2007     1208      1     3.635
## 9967  66049152167 4.214 2007     1208      1     3.237
## 9978  67650155463 3.783 2007     1208      1     2.806
## 9981  67650162724 3.610 2007     1208      1     2.633
## 9982  67650168953 3.939 2007     1208      1     2.962
## 10006 70449930640 3.610 2007     1208      1     2.690
## 10040 84868448556 3.610 2007     1208      1     2.690
## 10075 43249135021 4.018 2007     1208      2     3.041
## 10078 43249146684 4.726 2007     1208      2     3.749
## 10107 60949438530 4.726 2007     1208      2     3.806
## 10155 60950643835 4.839 2007     1208      2     3.919
## 10158 60950662433 3.834 2007     1208      2     2.914
## 10415 61949141429 3.441 2007     1208      3     2.521
## 10670 61049446035 5.641 2007     1208      1     4.721
## 10676 70450032239 3.610 2007     1208      1     2.690
## 10809 70450064287 5.151 2007     1208      1     4.231
## 10866 60949326226 3.610 2007     1208      1     2.633
## 10867 60949370742 5.976 2007     1208      1     5.056
## 10880 60950131470 3.939 2007     1208      1     2.962
## 10892 60950536451 4.449 2007     1208      1     3.472
## 10925 62449170699 5.841 2007     1208      1     4.864
## 10968 67649683064 4.214 2007     1208      1     3.294
## 10984 67650096553 4.082 2007     1208      1     3.162
## 10985 67650143129 5.474 2007     1208      1     4.554
## 10986 67650925409 3.939 2007     1208      1     3.019
## 10988 70249093068 3.610 2007     1208      1     2.690
## 10992 70449859265 3.610 2007     1208      1     2.690
## 10994 70449877375 3.610 2007     1208      1     2.690
## 11007 70450015708 6.179 2007     1208      1     5.259
## 11014 70450064757 3.783 2007     1208      1     2.806
## 11015 70450077380 4.654 2007     1208      1     3.734
## 11112 62449186372 3.857 2007     1208      2     2.880
## 11165 70450078906 4.018 2007     1208      2     3.098
## 11217 61249239835 4.188 2007     1208      3     3.268
## 11228 61249660537 3.741 2007     1208      3     2.821
## 11409 60949485276 3.367 2008     1208      1     2.509
## 11434 60950624831 3.802 2008     1208      1     2.944
## 11478 60950719981 5.997 2008     1208      1     5.081
## 11494 60950608143 3.692 2008     1203      3     2.543
## 11532 60950299886 3.598 2008     1208      1     2.740
## 11540 57749125280 4.798 2008     1208      1     3.940
## 11548 60949270646 4.686 2008     1208      1     3.828
## 11573 61049554358 3.802 2008     1208      1     2.944
## 11605 61949182548 3.598 2008     1208      1     2.682
## 11611 61949371494 3.367 2008     1208      1     2.509
## 11618 62449098267 5.426 2008     1208      1     4.568
```

```
## 11624 62449159076 3.598 2008     1208      1     2.740
## 11626 62449175332 5.942 2008     1208      1     5.026
## 11633 62449269964 3.984 2008     1208      1     3.126
## 11641 62949147900 4.300 2008     1208      1     3.384
## 11644 62949210835 3.367 2008     1208      1     2.509
## 11672 67650099391 3.984 2008     1208      1     3.068
## 11676 67650112946 3.984 2008     1208      1     3.068
## 11681 67650132635 4.149 2008     1208      1     3.233
## 11684 67650151239 3.984 2008     1208      1     3.068
## 11687 67650155476 3.367 2008     1208      1     2.509
## 11688 67650162729 4.798 2008     1208      1     3.882
## 11691 67650178260 5.001 2008     1208      1     4.085
## 11695 70449752752 4.300 2008     1208      1     3.442
## 11696 70449780196 3.367 2008     1208      1     2.509
## 11723 70450068192 3.367 2008     1208      1     2.509
## 11792 61149209901 3.427 2008     1208      2     2.569
## 11796 61149472770 4.222 2008     1208      2     3.306
## 12102 65749261392 3.585 2008     1202      3     2.669
## 12226 70450003031 5.184 2008     1208      1     4.268
## 12362 70450082589 3.516 2008     1205      3     2.600
## 12399 45349087542 3.582 2008     1208      2     2.724
## 12435 61249636505 6.426 2008     1208      1     5.568
## 12468 60949164169 3.720 2008     1208      3     2.862
## 12498 70449866030 3.802 2008     1208      1     2.886
## 12500 70450060571 4.686 2008     1208      1     3.770
## 12551 61049119505 3.367 2008     1208      1     2.509
## 12629 57749127836 3.984 2008     1208      1     3.126
## 12658 61249639675 3.598 2008     1208      1     2.682
## 12661 61449513631 3.984 2008     1208      1     3.126
## 12664 61949197772 4.149 2008     1208      1     3.291
## 12671 62449262219 4.300 2008     1208      1     3.442
## 12675 62749110813 6.152 2008     1208      1     5.294
## 12677 62749122811 5.268 2008     1208      1     4.410
## 12679 62749136269 3.984 2008     1208      1     3.126
## 12682 62749151909 4.798 2008     1208      1     3.940
## 12683 62749155048 4.149 2008     1208      1     3.233
## 12684 62749156172 5.942 2008     1208      1     5.084
## 12691 64949093382 3.598 2008     1208      1     2.682
## 12696 64949107038 3.802 2008     1208      1     2.886
## 12700 64949153413 4.300 2008     1208      1     3.442
## 12727 67650070793 4.149 2008     1208      1     3.291
## 12735 67650090943 5.570 2008     1208      1     4.712
## 12739 70449931160 5.268 2008     1208      1     4.352
## 12746 78650710041 3.802 2008     1208      1     2.886
## 12795 61249468411 4.222 2008     1208      2     3.364
## 12797 61249548069 3.427 2008     1208      2     2.569
## 12861 40449115608 3.564 2008     1202      4     2.939
## 13002 85014946308 3.363 2008     1203      4     2.505
## 13097 77950310142 4.054 2009     1208      1     3.287
## 13214 84884186648 4.054 2009     1208      1     3.344
```

```
## 13249 61949462044 3.777 2009     1208      3     3.067
## 13400 73949133087 4.425 2009     1208      1     3.715
## 13418 70349840389 4.249 2009     1208      1     3.539
## 13474 68549123389 3.591 2009     1208      1     2.881
## 13490 84867054792 3.306 2009     1208      1     2.830
## 13559 73649148390 3.360 2009     1202      4     2.650
## 13613 67649716605 3.591 2009     1208      1     2.824
## 13663 67649950267 4.054 2009     1208      1     3.344
## 13679 77950325600 4.997 2009     1208      1     4.287
## 13680 77950788847 3.306 2009     1208      1     2.596
## 13683 77950809870 4.054 2009     1208      1     3.054
## 13684 77950815002 6.760 2009     1208      1     5.993
## 13685 77950822756 5.116 2009     1208      1     4.349
## 13686 77950831009 4.585 2009     1208      1     3.875
## 13688 77950841010 6.560 2009     1208      1     5.793
## 13693 68949083793 3.825 2009     1208      2     3.115
## 13695 68949090683 4.540 2009     1208      2     3.830
## 13697 68949092620 3.727 2009     1208      2     3.017
## 13700 68949123188 3.423 2009     1208      2     2.713
## 13702 68949135917 3.636 2009     1208      2     2.926
## 13893 79958822732 4.054 2009     1208      1     3.287
## 13895 80054949947 3.837 2009     1208      1     3.070
## 13939 77950229849 3.223 2009     1208      2     2.513
## 13940 77950253669 3.223 2009     1208      2     2.513
## 14020 65849111025 3.591 2009     1208      1     2.824
## 14021 65849141252 3.306 2009     1208      1     2.596
## 14039 77956351978 3.837 2009     1208      1     3.070
## 14092 77950181710 4.785 2009     1205      2     4.018
## 14275 77950904394 3.302 2010     1208      1     2.629
## 14278 78649401790 3.586 2010     1208      1     2.856
## 14282 78649430807 3.586 2010     1208      1     2.856
## 14293 84855406055 3.832 2010     1208      1     3.102
## 14542 77957559973 4.049 2010     1208      1     3.376
## 14543 77957584700 3.302 2010     1208      1     2.572
## 14572 77956632352 3.506 2010     1208      2     2.776
## 14575 78649656075 4.243 2010     1208      1     3.513
## 14591 77956331789 4.376 2010     1208      2     3.703
## 14712 79959202296 3.634 2010     1208      2     2.961
## 14810 77950211324 3.302 2010     1208      1     2.629
## 14880 79956372974 3.668 2010     1202      4     2.995
## 14898 77949582311 4.243 2010     1208      1     3.513
## 14902 77949615907 3.586 2010     1208      1     2.913
## 14927 77950140620 3.832 2010     1208      1     3.159
## 14928 77950253361 3.302 2010     1208      1     2.629
## 14934 77950886825 3.586 2010     1208      1     2.856
## 14935 77950888966 3.586 2010     1208      1     2.856
## 15039 76349106441 3.586 2010     1208      1     2.856
## 15050 77954443733 3.302 2010     1208      1     2.572
## 15094 77950287584 3.586 2010     1208      1     2.913
## 15120 79955418712 4.243 2010     1208      1     3.513
```

```
## 15143 73249128238 4.062 2010     1208      2     3.622
## 15226 77956500865 4.776 2010     1208      3     4.103
## 15228 78649472664 5.064 2010     1208      3     4.100
## 15279 82455228764 3.393 2010     1208      3     2.720
## 15298 77951908318 3.254 2010     1202      6     2.581
## 15358 84855209572 3.491 2011     1208      1     2.756
## 15359 84855210092 3.491 2011     1208      1     2.698
## 15365 84855762177 3.887 2011     1208      1     3.152
## 15368 84855797090 3.491 2011     1208      1     2.756
## 15380 84859358430 3.491 2011     1208      1     2.756
## 15382 84860778742 5.724 2011     1208      1     4.989
## 15403 84870329240 3.491 2011     1208      1     2.756
## 15411 84872712403 3.491 2011     1208      1     2.698
## 15419 84873283995 3.491 2011     1208      1     2.698
## 15424 84873302367 3.887 2011     1208      1     3.152
## 15447 84884198571 3.887 2011     1208      1     3.152
## 15465 84887967106 3.491 2011     1208      1     2.756
## 15474 84855781328 3.573 2011     1208      2     2.838
## 15475 84855791481 5.406 2011     1208      2     4.613
## 15484 84857779496 3.573 2011     1208      2     2.547
## 15615 84857935340 3.474 2011     1202      6     2.739
## 15674 80155194058 3.906 2011     1202      5     3.113
## 15694 80054789940 4.221 2011     1208      1     3.428
## 15698 80054708194 3.887 2011     1208      1     3.152
## 15717 80053615731 4.511 2011     1208      1     3.776
## 15803 80052975515 3.243 2011     1205      2     2.508
## 15826 80052542479 3.887 2011     1208      1     3.152
## 15845 84055192224 6.801 2011     1208      1     6.008
## 15847 84856196552 3.887 2011     1208      1     3.152
## 15853 84862322836 3.887 2011     1208      1     3.094
## 15869 84883480879 4.995 2011     1208      1     4.202
## 15899 84861070464 5.091 2011     1208      2     4.356
## 16068 79959952724 4.221 2011     1208      1     3.428
## 16069 79959977747 3.491 2011     1208      1     2.698
## 16212 79955377318 6.894 2011     1208      1     6.101
## 16216 79955439141 4.221 2011     1208      1     3.486
## 16219 79957888983 3.887 2011     1208      1     3.152
## 16290 79958846842 3.887 2011     1208      1     3.152
## 16360 79955674091 3.677 2011     1203      6     2.884
## 16379 79955677215 6.146 2011     1208      1     5.353
## 16383 80054915853 3.887 2011     1208      1     3.094
## 16498 79955368617 3.375 2011     1203      3     2.582
## 16516 78650843435 5.564 2011     1208      1     4.829
## 16517 78650845690 4.221 2011     1208      1     3.486
## 16541 79960860760 4.221 2011     1208      1     3.428
## 16542 79960871745 4.221 2011     1208      1     3.486
## 16550 81355124072 3.887 2011     1208      1     3.094
## 16567 84055198510 3.887 2011     1208      1     3.094
## 16586 85011484631 3.887 2011     1208      1     3.094
## 16615 80555156002 3.648 2011     1208      2     2.855
```

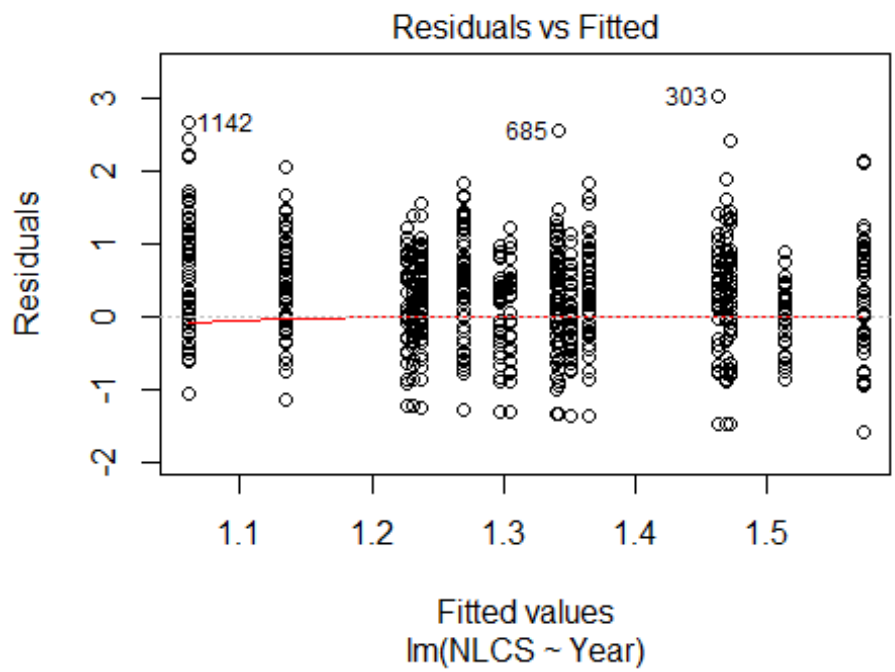
```
## 16630 84555179123 4.538 2011     1208      2     3.803
## 16726 82455232963 4.204 2011     1208      3     3.411
## 16727 82455244076 4.997 2011     1208      3     4.262
## 16835 84871206897 3.953 2012     1208      2     3.313
## 16837 84871229041 3.842 2012     1208      2     3.202
## 16880 84870331366 3.391 2012     1208      1     2.751
## 16882 84870350788 6.781 2012     1208      1     6.141
## 16892 84855231775 3.391 2012     1208      1     2.693
## 16896 84856905745 3.391 2012     1208      1     2.751
## 16898 84856945735 4.382 2012     1208      1     3.742
## 16903 84871438787 3.937 2012     1208      1     3.239
## 16909 84871985774 3.391 2012     1208      1     2.751
## 16914 84872706828 3.391 2012     1208      1     2.751
## 16926 84873281046 3.937 2012     1208      1     3.297
## 16949 84874079832 7.447 2012     1208      1     6.749
## 16950 84874095385 3.937 2012     1208      1     3.239
## 16957 84876994217 5.086 2012     1208      1     4.388
## 16958 84876998823 3.391 2012     1208      1     2.693
## 16959 84877000419 3.937 2012     1208      1     3.239
## 16960 84878432148 4.382 2012     1208      1     3.742
## 17050 84871269332 3.156 2012     1208      2     2.516
## 17170 84871367509 3.245 2012     1202      3     2.547
## 17171 84871394290 5.615 2012     1202      3     4.917
## 17215 84874341423 3.664 2012     1202      3     3.024
## 17318 84869003069 5.086 2012     1208      1     4.446
## 17320 84869033091 3.937 2012     1208      1     3.239
## 17326 84868366032 3.394 2012     1208      2     2.754
## 17327 84868368393 3.394 2012     1208      2     2.754
## 17375 84874056751 3.264 2012     1202      4     2.566
## 17431 84871011871 3.937 2012     1208      1     3.297
## 17462 84867923419 3.313 2012     1208      2     2.906
## 17464 84868631017 3.559 2012     1208      2     2.861
## 17465 84868671128 3.962 2012     1208      2     3.322
## 17569 84865697715 3.939 2012     1208      2     3.299
## 17570 84865700895 4.859 2012     1208      2     4.161
## 17571 84865204088 3.937 2012     1208      1     3.297
## 17574 84866596965 4.759 2012     1208      1     4.119
## 17578 84866635016 3.937 2012     1208      1     3.297
## 17594 84867391849 3.391 2012     1208      1     2.693
## 17604 84870451817 3.937 2012     1208      1     3.297
## 17607 84870509187 4.382 2012     1208      1     3.742
## 17616 84871215069 3.391 2012     1208      1     2.693
## 17617 84871229501 3.937 2012     1208      1     3.297
## 17622 84871342985 5.374 2012     1208      1     4.676
## 17624 84871749299 3.391 2012     1208      1     2.693
## 17631 84873941703 3.937 2012     1208      1     3.297
## 17632 84873952317 3.391 2012     1208      1     2.751
## 17633 84873956808 3.937 2012     1208      1     3.239
## 17668 84885611499 5.086 2012     1208      1     4.446
## 17727 84865713540 3.451 2012     1202      3     2.811
```

```
## 17750 84866640267 5.127 2012     1208      3     4.429
## 17762 84866982605 3.164 2012     1203      3     2.524
## 17828 84864711707 3.258 2012     1202      5     2.618
## 17849 84865371557 5.086 2012     1208      1     4.388
## 17850 84865386536 6.781 2012     1208      1     6.141
## 17858 84865347957 3.391 2012     1208      1     2.693
## 17863 84865160135 5.086 2012     1208      1     4.446
## 17864 84865162714 5.086 2012     1208      1     4.388
## 17874 84864926904 4.302 2012     1208      2     3.662
## 17897 84883323231 3.240 2012     1208      2     2.600
## 17928 84870409111 3.654 2012     1203      3     3.247
## 17953 84864349568 5.374 2012     1208      1     4.676
## 17968 84874377033 4.759 2012     1208      1     4.119
## 17987 84864035512 4.620 2012     1208      2     3.980
## 18033 84864201531 3.937 2012     1208      1     3.239
## 18034 84864215638 6.455 2012     1208      1     5.757
## 18035 84864219524 5.086 2012     1208      1     4.388
## 18036 84864248050 5.374 2012     1208      1     4.676
## 18037 84864268571 6.455 2012     1208      1     5.815
## 18043 84865177201 3.937 2012     1208      1     3.297
## 18054 84875659820 5.632 2012     1208      1     4.992
## 18091 84866718537 4.477 2012     1208      2     3.837
## 18215 84859330957 5.086 2012     1208      1     4.446
## 18233 84861746194 3.412 2012     1208      2     2.772
## 18242 84883352820 3.619 2012     1208      2     2.979
## 18247 84883394654 3.596 2012     1208      2     2.898
## 18259 84872852785 3.240 2012     1203      3     2.600
## 18266 84872871724 3.721 2012     1203      3     3.023
## 18281 84859745370 4.759 2012     1208      1     4.061
## 18284 84859760309 3.391 2012     1208      1     2.751
## 18296 84876226058 4.759 2012     1208      1     4.061
## 18414 84865691275 3.391 2012     1208      1     2.751
## 18427 84884668586 3.391 2012     1208      1     2.751
## 18448 84861350433 4.598 2012     1208      2     3.900
## 18449 84861359249 3.559 2012     1208      2     2.861
## 18452 84861395909 3.559 2012     1208      2     2.861
## 18480 84860734606 3.999 2012     1202      3     3.301
## 18489 84862115131 3.380 2012     1202      4     2.682
## 18490 84862116557 3.380 2012     1202      4     2.740
## 18555 84856799214 3.937 2012     1208      1     3.239
## 18557 84856642915 4.382 2012     1208      1     3.684
## 18558 84856653530 4.382 2012     1208      1     3.684
## 18565 84856448236 3.227 2012     1202      3     2.587
## 18566 84856468758 4.331 2012     1202      3     3.633
## 18567 84856478984 3.344 2012     1202      3     2.646
## 18576 84864376206 4.361 2012     1208      2     3.721
## 18632 84856135031 3.937 2012     1208      1     3.239
## 18636 84856185796 4.759 2012     1208      1     4.119
## 18637 84856189103 4.759 2012     1208      1     4.119
## 18638 84856197118 6.781 2012     1208      1     6.141
```

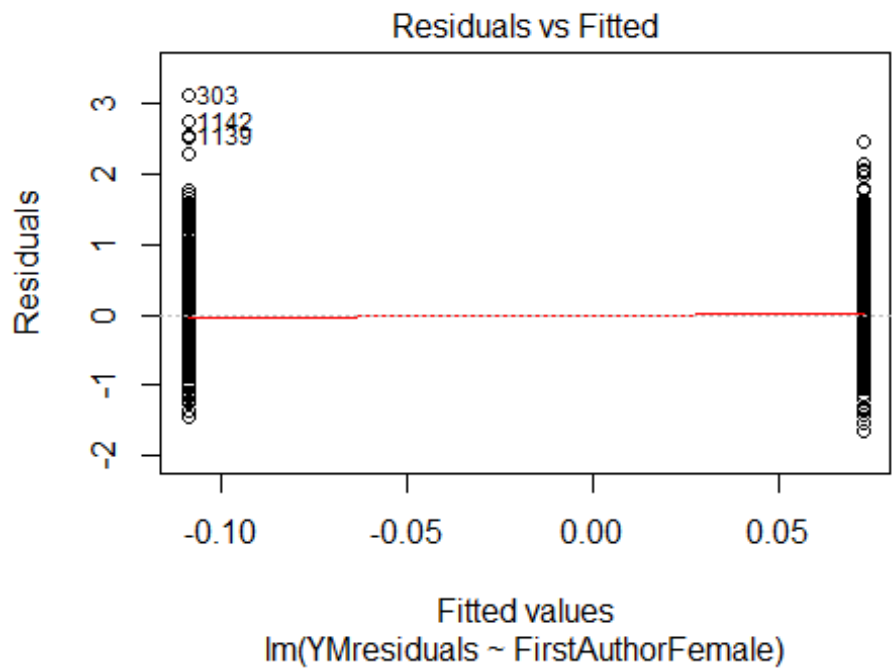
```
## 18648 84863619762 3.391 2012     1208      1     2.751
## 18664 84867394653 3.937 2012     1208      1     3.297
## 18737 84865710516 3.559 2012     1208      2     2.861
## 18740 84867763226 3.284 2012     1205      2     2.644
## 18745 84870312908 3.962 2012     1208      2     3.322
## 18747 84870334955 6.510 2012     1208      2     5.812
## 18751 84870345747 6.391 2012     1208      2     5.751
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.1148 -0.7695 -0.0362  0.7462  7.7750
##
## Coefficients:
##                   Estimate Std. Error t value Pr(>|t|)
## (Intercept)         0.7695     0.0495   15.55  < 2e-16 ***
## LastAuthorFemale1   0.0646     0.0161    4.02  5.9e-05 ***
## Year1997            0.0139     0.0660    0.21  0.83306
## Year1998            0.0510     0.0653    0.78  0.43493
## Year1999           -0.0251     0.0617   -0.41  0.68419
## Year2000            0.0479     0.0636    0.75  0.45097
## Year2001            0.0150     0.0616    0.24  0.80724
## Year2002            0.2125     0.0589    3.61  0.00031 ***
## Year2003            0.2137     0.0581    3.68  0.00024 ***
## Year2004            0.2808     0.0581    4.83  1.4e-06 ***
## Year2005            0.1749     0.0577    3.03  0.00244 **
## Year2006            0.1794     0.0577    3.11  0.00187 **
## Year2007            0.1470     0.0567    2.59  0.00949 **
## Year2008            0.0864     0.0561    1.54  0.12391
## Year2009           -0.0615     0.0569   -1.08  0.27965
## Year2010           -0.0995     0.0569   -1.75  0.08000 .
## Year2011           -0.0374     0.0563   -0.66  0.50688
## Year2012           -0.1318     0.0543   -2.43  0.01527 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.919
## Multiple R-squared:  0.0193, Adjusted R-squared:  0.0183
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
##  63 observations
c(2421,3496,3615,4107,4269,4621,5346,5362,5507,6122,6167,6744,8136,8307,8344,
8580,9339,9506,9559,9611,9629,10043,10161,10169,10858,11065,11067,11072,11116
,11120,11941,11942,11945,13274,13395,13479,13798,13846,14110,14249,14369,1468
4,14748,14755,14952,15080,15341,15378,15452,15539,15540,15551,15552,15630,156
```

```
96,15697,15698,15699,15715,15863,16233,16332,16334)
##   are outliers with |weight| = 0 ( < 6.1e-06);
##  1200 weights are ~= 1. The remaining 15231 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.000   0.892   0.933   0.891   0.971   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          6.06e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Sample size for the above analysis:  16494"
## [1] ""
## [1] ""
## [1] "####################################"
## [1] "Analysis of AJSC 1209"
## [1] "####################################"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##   42   32   46   44   45   53   53   51   68   57   75   70   66   97  101
## 2011 2012
##  108  117
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##   37   30   43   37   39   42   43   46   59   52   69   65   56   91   85
## 2011 2012
##   96  103
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##   34   29   43   35   34   38   43   42   58   48   68   62   53   89   85
## 2011 2012
##   93  100
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
##  Bartlett test of homogeneity of variances
##
```

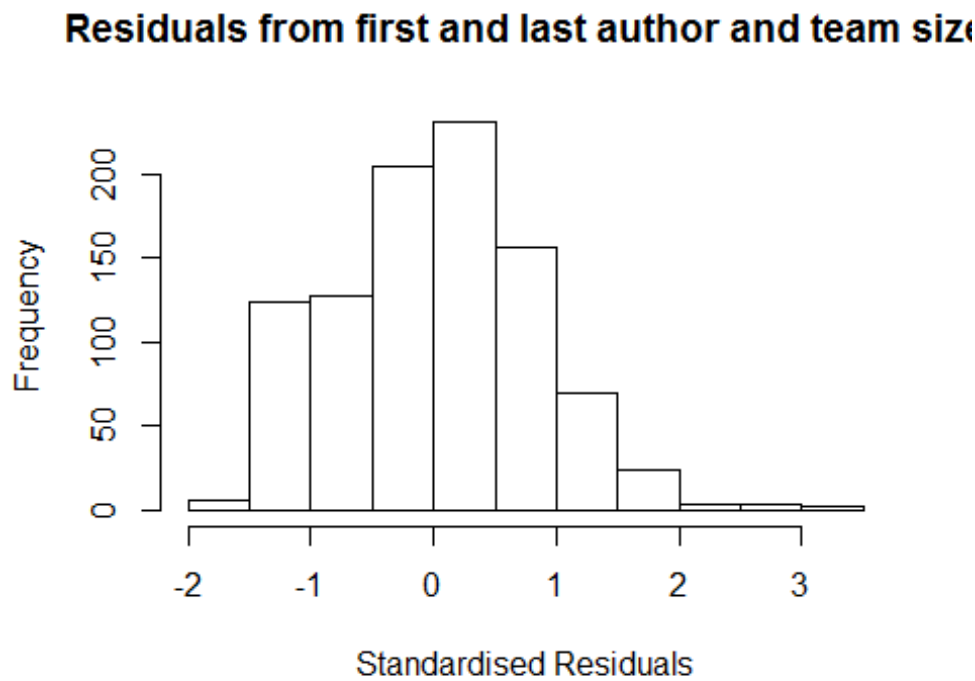
```
## data:  NLCS by Year
## Bartlett's K-squared = 83, df = 16, p-value = 4e-11
```




```
##
##  Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 0.12, df = 1, p-value = 0.7
```



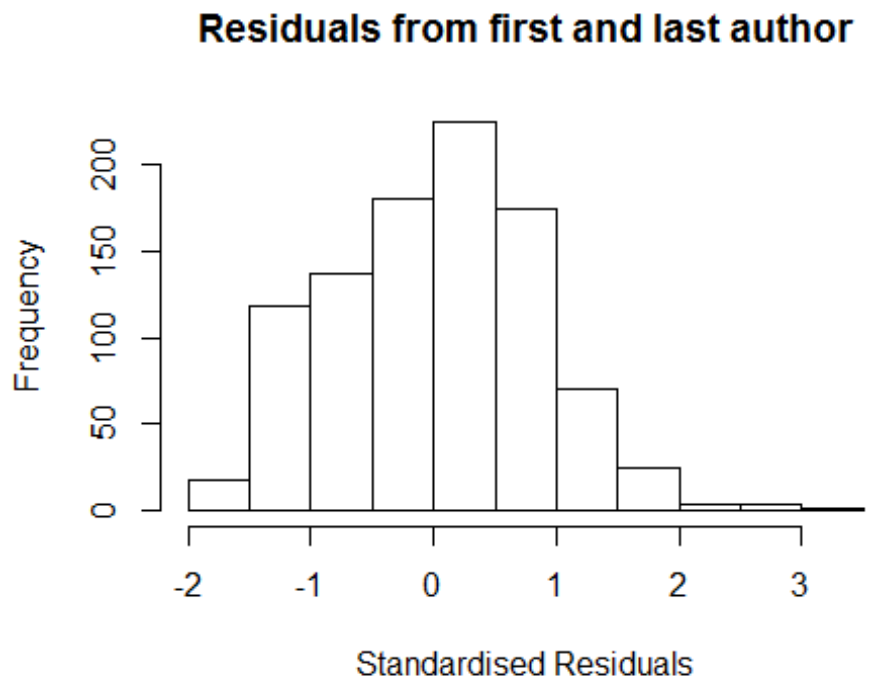

```
## [1] "Female first author team size 2018 geometric mean: 1.43614615050828"
## [1] "Male first author team size 2018 geometric mean: 1.42300765335865"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 1300, p-value = 0.8
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 1.36901753698476"
## [1] "Male last author team size 2018 geometric mean: 1.52421416193898"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 1200, p-value = 0.2
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##                     GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.865  1           1.366
## LastAuthorFemale  1.930  1           1.389
## UniqueAuthors     1.479  4           1.050
## Year              1.605 16           1.015
```




```
## [1] "List of  6 outliers with residuals above  2.5"
##         ScopusId  NLCS Year OneField Fields residuals
## 303   0036596824 4.474 2002     1202      3     3.247
## 351   0042885451 3.882 2003     1202      3     2.767
## 685  46849103881 3.875 2008     1202      3     2.674
## 1088 84876439550 3.251 2012     1201      4     2.537
## 1139 84857525840 3.491 2012     1202      3     2.777
## 1142 84857597635 3.715 2012     1202      3     3.001
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
UniqueAuthors +
##     Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##     k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.6799 -0.5674  0.0308  0.5584  3.2474
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.42163    0.09567   14.86  < 2e-16 ***
## FirstAuthorFemale1 -0.12560    0.07294   -1.72   0.0854 .
## LastAuthorFemale1  -0.07170    0.07429   -0.97   0.3347
## UniqueAuthors2      0.36645    0.06808    5.38  9.3e-08 ***
## UniqueAuthors3      0.49118    0.08364    5.87  6.0e-09 ***
```

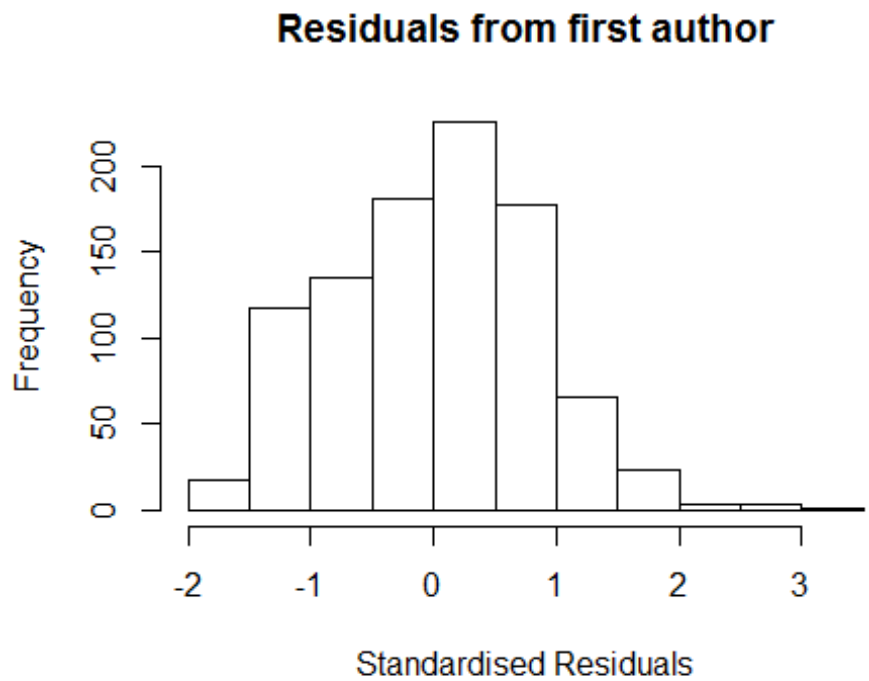
```
## UniqueAuthors4      0.54230    0.13419    4.04  5.8e-05 ***
## UniqueAuthors5      0.29645    0.19617    1.51   0.1311
## Year1997           -0.22231    0.15079   -1.47   0.1407
## Year1998           -0.14918    0.11589   -1.29   0.1983
## Year1999           -0.16640    0.13063   -1.27   0.2030
## Year2000           -0.22212    0.15030   -1.48   0.1398
## Year2001           -0.24164    0.12943   -1.87   0.0622 .
## Year2002            0.00222    0.17249    0.01   0.9897
## Year2003           -0.10911    0.18559   -0.59   0.5567
## Year2004            0.10155    0.15153    0.67   0.5029
## Year2005            0.03637    0.15849    0.23   0.8185
## Year2006           -0.03651    0.14591   -0.25   0.8025
## Year2007           -0.15770    0.15518   -1.02   0.3098
## Year2008           -0.22091    0.12538   -1.76   0.0784 .
## Year2009           -0.32672    0.13292   -2.46   0.0141 *
## Year2010           -0.15501    0.11274   -1.37   0.1695
## Year2011           -0.17203    0.15301   -1.12   0.2612
## Year2012           -0.51053    0.14357   -3.56   0.0004 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.796
## Multiple R-squared:  0.0964, Adjusted R-squared:  0.075
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
##  77 weights are ~= 1. The remaining 877 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0587  0.8480  0.9460  0.9040  0.9860  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.05e-04          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 2.016  1           1.420
```

```
## LastAuthorFemale  2.008  1           1.417
## Year              1.181 16           1.005
```



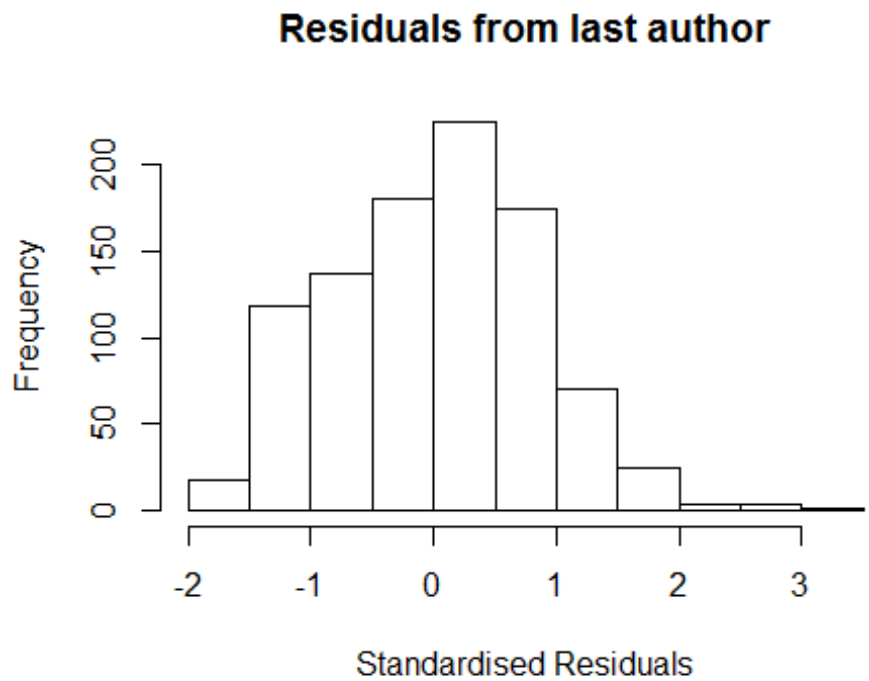

```
## [1] "List of  4 outliers with residuals above  2.5"
##         ScopusId  NLCS Year OneField Fields residuals
## 303   0036596824 4.474 2002     1202      3     3.175
## 351   0042885451 3.882 2003     1202      3     2.608
## 1139 84857525840 3.491 2012     1202      3     2.674
## 1142 84857597635 3.715 2012     1202      3     2.898
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##    Min     1Q Median     3Q    Max
## -1.701 -0.622  0.053  0.572  3.175
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)           1.5911     0.0840   18.93  < 2e-16 ***
## FirstAuthorFemale1   -0.1278     0.0799   -1.60  0.10996
## LastAuthorFemale1    -0.1213     0.0801   -1.51  0.13061
## Year1997             -0.2326     0.1553   -1.50  0.13455
## Year1998             -0.1799     0.1144   -1.57  0.11627
## Year1999             -0.1895     0.1319   -1.44  0.15106
```

```
## Year2000            -0.3194     0.1487   -2.15  0.03195 *
## Year2001            -0.2798     0.1317   -2.13  0.03383 *
## Year2002            -0.0428     0.1778   -0.24  0.80985
## Year2003            -0.0679     0.1968   -0.34  0.73021
## Year2004             0.1095     0.1488    0.74  0.46201
## Year2005            -0.0438     0.1587   -0.28  0.78252
## Year2006            -0.1150     0.1443   -0.80  0.42579
## Year2007            -0.1703     0.1556   -1.09  0.27401
## Year2008            -0.1920     0.1200   -1.60  0.11013
## Year2009            -0.3621     0.1285   -2.82  0.00492 **
## Year2010            -0.2280     0.1077   -2.12  0.03450 *
## Year2011            -0.2501     0.1476   -1.69  0.09044 .
## Year2012            -0.5252     0.1421   -3.70  0.00023 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.822
## Multiple R-squared:  0.053,  Adjusted R-squared:  0.0348
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
##  80 weights are ~= 1. The remaining 874 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.103   0.853   0.942   0.906   0.983   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.05e-04          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.078  1           1.038
## Year              1.078 16           1.002
```




```
## [1] "List of  4 outliers with residuals above  2.5"
##         ScopusId  NLCS Year OneField Fields residuals
## 303   0036596824 4.474 2002     1202      3     3.175
## 351   0042885451 3.882 2003     1202      3     2.608
## 1139 84857525840 3.491 2012     1202      3     2.674
## 1142 84857597635 3.715 2012     1202      3     2.898
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.6919 -0.6247  0.0548  0.5889  3.1601
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)          1.5779     0.0822   19.20  < 2e-16 ***
## FirstAuthorFemale1  -0.2219     0.0591   -3.75  0.00019 ***
## Year1997            -0.2128     0.1542   -1.38  0.16773
## Year1998            -0.1788     0.1112   -1.61  0.10822
## Year1999            -0.1837     0.1324   -1.39  0.16577
## Year2000            -0.3182     0.1482   -2.15  0.03201 *
## Year2001            -0.2810     0.1301   -2.16  0.03097 *
## Year2002            -0.0421     0.1775   -0.24  0.81242
## Year2003            -0.0705     0.1947   -0.36  0.71720
```

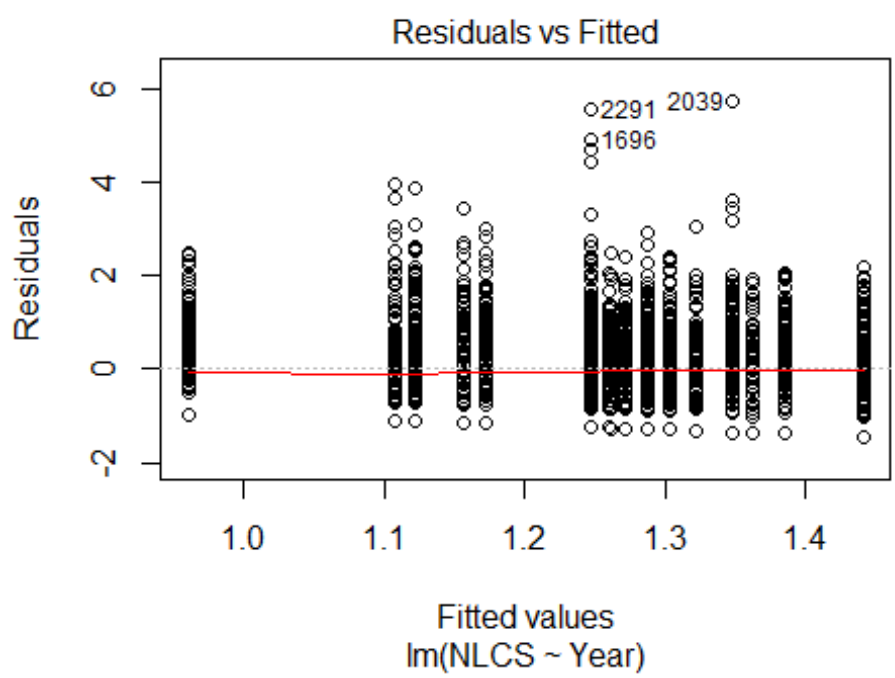
```
## Year2004             0.1140     0.1476    0.77  0.44007
## Year2005            -0.0379     0.1595   -0.24  0.81212
## Year2006            -0.1114     0.1442   -0.77  0.44024
## Year2007            -0.1635     0.1559   -1.05  0.29467
## Year2008            -0.1763     0.1196   -1.47  0.14091
## Year2009            -0.3674     0.1276   -2.88  0.00407 **
## Year2010            -0.2210     0.1067   -2.07  0.03861 *
## Year2011            -0.2485     0.1472   -1.69  0.09166 .
## Year2012            -0.5276     0.1419   -3.72  0.00021 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.822
## Multiple R-squared:  0.0512, Adjusted R-squared:  0.034
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
##  81 weights are ~= 1. The remaining 873 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.106   0.857   0.942   0.905   0.984   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.05e-04          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##                   GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.071  1           1.035
## Year             1.071 16           1.002
```



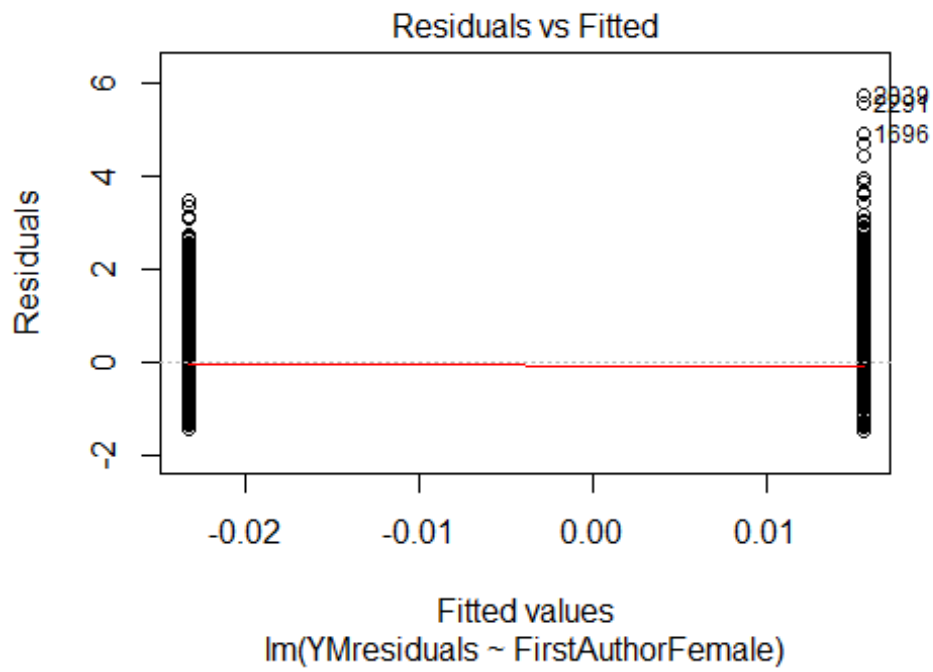

```
## [1] "List of  4 outliers with residuals above  2.5"
##         ScopusId  NLCS Year OneField Fields residuals
## 303   0036596824 4.474 2002     1202      3     3.175
## 351   0042885451 3.882 2003     1202      3     2.608
## 1139 84857525840 3.491 2012     1202      3     2.674
## 1142 84857597635 3.715 2012     1202      3     2.898
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.6824 -0.6251  0.0549  0.5768  3.1577
##
## Coefficients:
##                   Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.5843     0.0842   18.81  < 2e-16 ***
## LastAuthorFemale1  -0.2200     0.0594   -3.71  0.00022 ***
## Year1997           -0.2477     0.1539   -1.61  0.10777
## Year1998           -0.1755     0.1173   -1.50  0.13507
## Year1999           -0.1955     0.1308   -1.49  0.13526
## Year2000           -0.3174     0.1490   -2.13  0.03340 *
## Year2001           -0.2791     0.1333   -2.09  0.03659 *
## Year2002           -0.0480     0.1779   -0.27  0.78735
## Year2003           -0.0616     0.1981   -0.31  0.75589
```

```
## Year2004            0.0981     0.1489    0.66  0.51013
## Year2005           -0.0563     0.1579   -0.36  0.72132
## Year2006           -0.1218     0.1433   -0.85  0.39565
## Year2007           -0.1797     0.1534   -1.17  0.24193
## Year2008           -0.2152     0.1190   -1.81  0.07076 .
## Year2009           -0.3642     0.1289   -2.82  0.00483 **
## Year2010           -0.2413     0.1074   -2.25  0.02487 *
## Year2011           -0.2528     0.1479   -1.71  0.08761 .
## Year2012           -0.5283     0.1428   -3.70  0.00023 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.825
## Multiple R-squared:  0.0505, Adjusted R-squared:  0.0332
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
##  79 weights are ~= 1. The remaining 875 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.111   0.856   0.944   0.906   0.984   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.05e-04          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Sample size for the above analysis:  954"
## [1] ""
## [1] ""
## [1] "####################################"
## [1] "Analysis of AJSC 1210"
## [1] "####################################"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  121  134  115  102   94  124  188  184  259  235  256  243  289  249  217
## 2011 2012
##  343  376
##
```

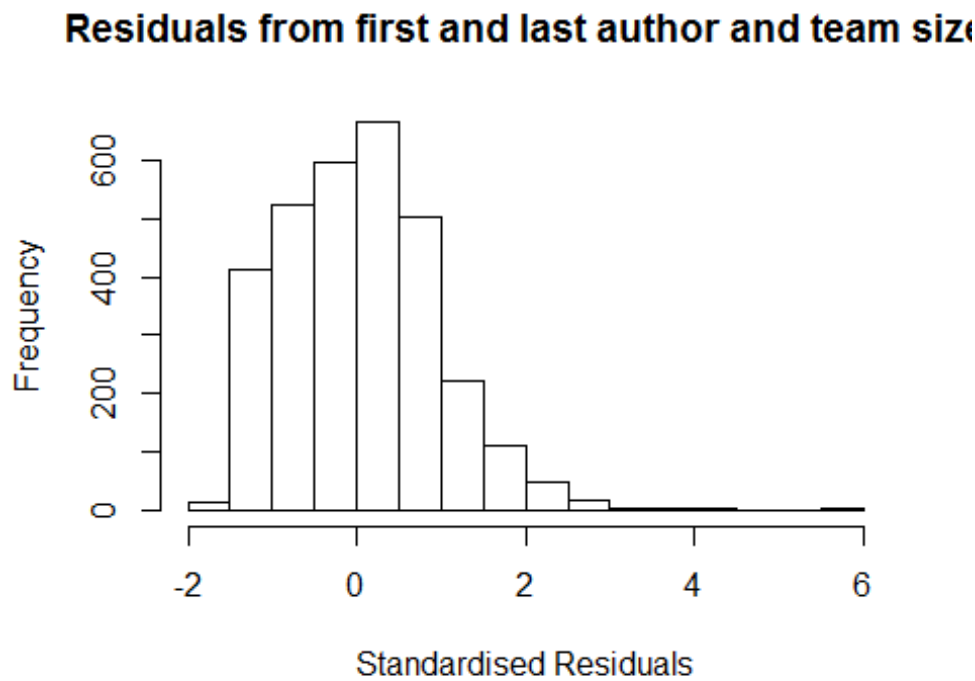
```
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  110  128   97   89   88   99  169  175  232  220  237  221  261  229  189
## 2011 2012
##  311  331
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  109  127   94   87   87   98  166  174  226  210  229  218  255  223  186
## 2011 2012
##  303  326
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
##  Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 57, df = 16, p-value = 2e-06
```




```
##
##  Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 16, df = 1, p-value = 5e-05
```



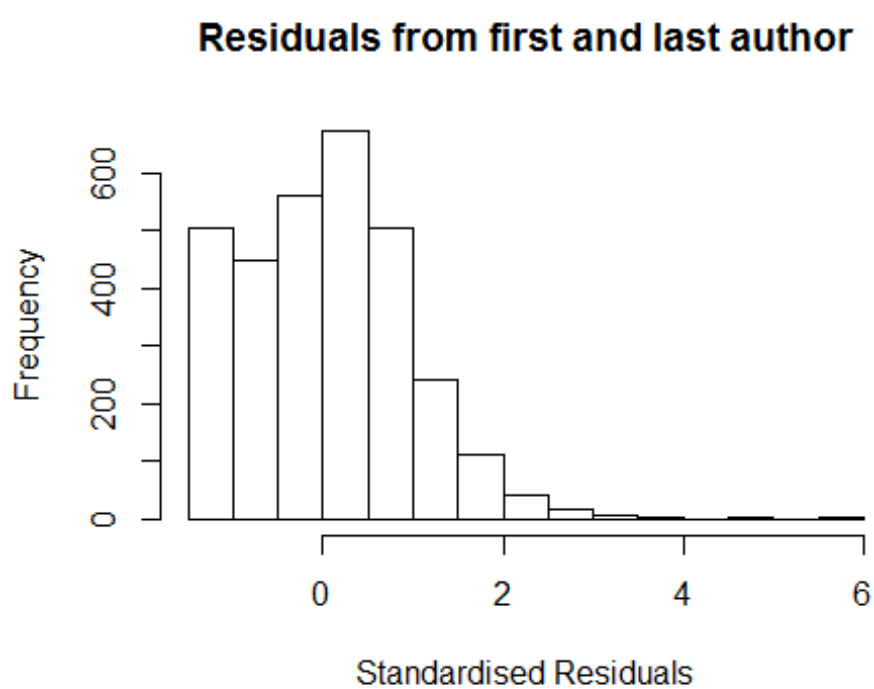

```
## [1] "Female first author team size 2018 geometric mean: 1.21993482434654"
## [1] "Male first author team size 2018 geometric mean: 1.17951096550236"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 12000, p-value = 0.2
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 1.18578070482994"
## [1] "Male last author team size 2018 geometric mean: 1.20872920173386"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 11000, p-value = 0.7
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 3.837  1           1.959
## LastAuthorFemale  3.818  1           1.954
## UniqueAuthors     1.127  4           1.015
## Year              1.141 16           1.004
```




```
## [1] "List of  29 outliers with residuals above  2.5"
##         ScopusId  NLCS Year OneField Fields residuals
## 77    0001903821 3.758 1996     1210      1     2.789
## 78    0004755588 3.620 1996     1210      1     2.636
## 79    0011299279 4.607 1996     1210      1     3.638
## 81    0242454754 3.878 1996     1210      1     2.909
## 337   0001194993 4.390 1998     1210      1     2.767
## 1247 63149119675 3.725 2004     1210      1     2.507
## 1487 33847791651 4.223 2005     1210      1     2.658
## 1503 33646729315 3.969 2005     1210      1     2.815
## 1696 60950491226 6.170 2006     1210      2     5.032
## 2029 34047192756 4.513 2007     1210      1     2.870
## 2031 34047235257 4.790 2007     1210      1     3.147
## 2039 36049016758 7.080 2007     1210      2     5.884
## 2130 60950499484 3.660 2008     1210      2     2.614
## 2289 57049157739 5.668 2008     1210      2     4.211
## 2290 57049173452 5.953 2008     1210      2     4.496
## 2291 57049179599 6.823 2008     1210      2     5.761
## 2300 42349109603 4.568 2008     1210      1     3.035
## 2323 38949195334 4.014 2008     1210      1     2.505
## 2438 76849084888 4.006 2009     1210      1     2.507
## 2579 75149146477 4.172 2009     1210      2     3.125
## 2779 77952935263 3.635 2010     1210      1     2.788
## 2797 77951477155 3.967 2010     1210      1     2.709
## 2894 77956500865 4.776 2010     1208      3     3.513
## 2895 78649472664 5.064 2010     1208      3     3.663
## 2899 82455228764 3.393 2010     1208      3     2.531
```

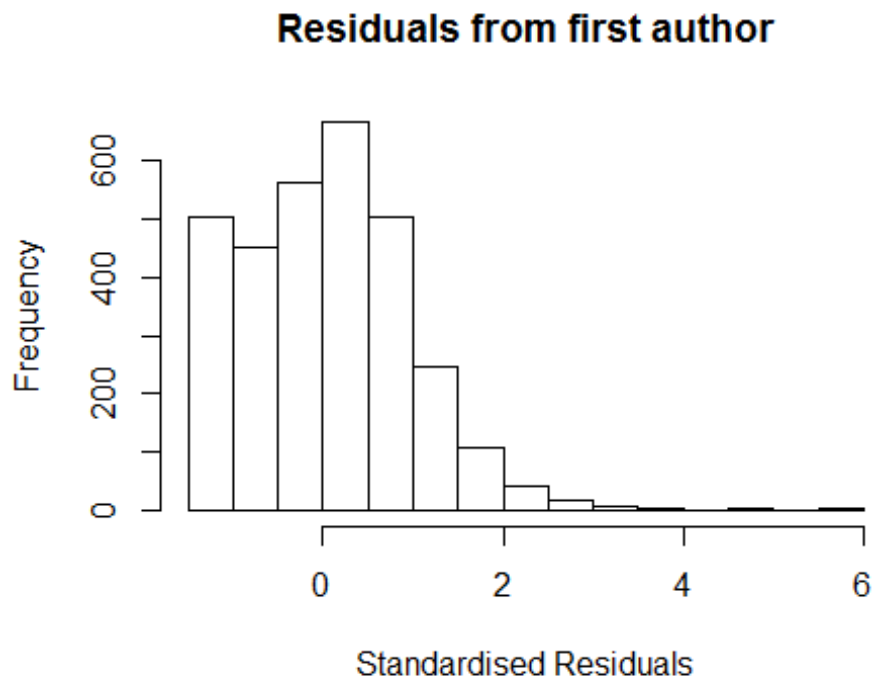
```
## 2936 84858822292 3.704 2011     1210      2     2.733
## 3248 82455232963 4.204 2011     1208      3     2.854
## 3249 82455244076 4.997 2011     1208      3     3.631
## 3389 84869152027 3.467 2012     1210      3     2.643
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
UniqueAuthors +
##     Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##     k.max = 1000))
##  \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.78266 -0.69141  0.00711  0.61050  5.88412
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)          0.9843     0.1239    7.94  2.8e-15 ***
## FirstAuthorFemale1  -0.0673     0.0636   -1.06   0.2897
## LastAuthorFemale1    0.0516     0.0635    0.81   0.4164
## UniqueAuthors2       0.3954     0.0463    8.54  < 2e-16 ***
## UniqueAuthors3       0.4006     0.0942    4.25  2.2e-05 ***
## UniqueAuthors4       0.7397     0.0885    8.36  < 2e-16 ***
## UniqueAuthors5       0.4867     0.1647    2.95   0.0032 **
## Year1997             0.3161     0.1492    2.12   0.0342 *
## Year1998             0.1914     0.1547    1.24   0.2162
## Year1999             0.1971     0.1553    1.27   0.2043
## Year2000             0.2761     0.1519    1.82   0.0692 .
## Year2001             0.1803     0.1466    1.23   0.2189
## Year2002             0.2254     0.1409    1.60   0.1098
## Year2003             0.3978     0.1420    2.80   0.0051 **
## Year2004             0.2332     0.1365    1.71   0.0877 .
## Year2005             0.1856     0.1358    1.37   0.1718
## Year2006             0.1541     0.1361    1.13   0.2575
## Year2007             0.2116     0.1389    1.52   0.1277
## Year2008             0.0775     0.1355    0.57   0.5672
## Year2009             0.0628     0.1394    0.45   0.6522
## Year2010            -0.1218     0.1416   -0.86   0.3897
## Year2011            -0.0136     0.1331   -0.10   0.9188
## Year2012            -0.1604     0.1320   -1.22   0.2242
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.906
## Multiple R-squared:  0.0615, Adjusted R-squared:  0.0548
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
##  4 observations c(1456,1746,1964,1965)
```

```
##   are outliers with |weight| = 0 ( < 3.2e-05);
##  269 weights are ~= 1. The remaining 2845 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0003  0.8790  0.9430  0.9120  0.9830  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          3.21e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 4.338  1           2.083
## LastAuthorFemale  4.328  1           2.081
## Year              1.045 16           1.001
```



```
## [1] "List of  32 outliers with residuals above  2.5"
##         ScopusId  NLCS Year OneField Fields residuals
## 77    0001903821 3.758 1996     1210      1     2.720
## 78    0004755588 3.620 1996     1210      1     2.584
## 79    0011299279 4.607 1996     1210      1     3.569
## 81    0242454754 3.878 1996     1210      1     2.840
## 337   0001194993 4.390 1998     1210      1     3.037
## 1487 33847791651 4.223 2005     1210      1     2.981
## 1503 33646729315 3.969 2005     1210      1     2.725
## 1696 60950491226 6.170 2006     1210      2     4.965
## 2029 34047192756 4.513 2007     1210      1     3.163
## 2031 34047235257 4.790 2007     1210      1     3.440
## 2039 36049016758 7.080 2007     1210      2     5.802
## 2130 60950499484 3.660 2008     1210      2     2.524
## 2289 57049157739 5.668 2008     1210      2     4.534
## 2290 57049173452 5.953 2008     1210      2     4.819
## 2291 57049179599 6.823 2008     1210      2     5.689
## 2300 42349109603 4.568 2008     1210      1     3.432
## 2321 38949105986 3.882 2008     1210      1     2.676
## 2323 38949195334 4.014 2008     1210      1     2.808
## 2438 76849084888 4.006 2009     1210      1     2.819
## 2565 63049091821 3.645 2009     1210      1     2.600
## 2579 75149146477 4.172 2009     1210      2     3.057
## 2735 77955800148 3.635 2010     1210      2     2.681
## 2779 77952935263 3.635 2010     1210      1     2.680
## 2797 77951477155 3.967 2010     1210      1     3.013
## 2894 77956500865 4.776 2010     1208      3     3.822
## 2895 78649472664 5.064 2010     1208      3     4.038
## 2936 84858822292 3.704 2011     1210      2     2.647
## 3123 79954543516 3.648 2011     1210      1     2.661
## 3219 79960456514 3.730 2011     1210      2     2.673
## 3248 82455232963 4.204 2011     1208      3     3.145
## 3249 82455244076 4.997 2011     1208      3     3.940
## 3389 84869152027 3.467 2012     1210      3     2.552
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.4404 -0.6740  0.0344  0.6184  5.8022
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)          1.0363     0.1245    8.33   <2e-16 ***
## FirstAuthorFemale1  -0.0706     0.0692   -1.02   0.3075
## LastAuthorFemale1    0.0721     0.0690    1.04   0.2964
## Year1997             0.3342     0.1504    2.22   0.0263 *
```

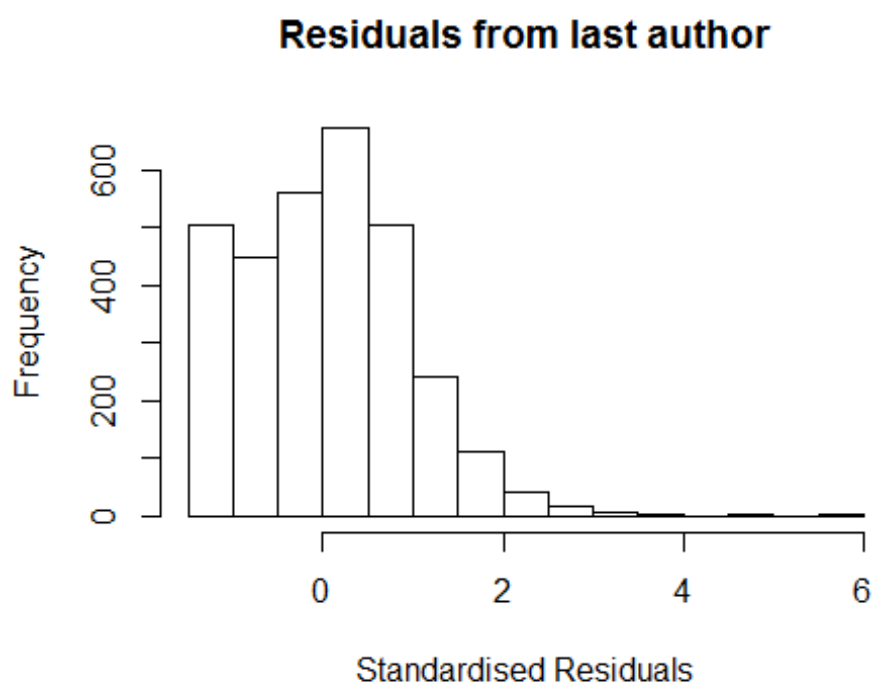
```
## Year1998             0.2444     0.1578    1.55   0.1215
## Year1999             0.2067     0.1579    1.31   0.1908
## Year2000             0.2955     0.1527    1.94   0.0530 .
## Year2001             0.1998     0.1491    1.34   0.1804
## Year2002             0.2199     0.1417    1.55   0.1207
## Year2003             0.4027     0.1436    2.80   0.0051 **
## Year2004             0.2261     0.1375    1.64   0.1002
## Year2005             0.2060     0.1369    1.50   0.1325
## Year2006             0.1691     0.1373    1.23   0.2182
## Year2007             0.2415     0.1398    1.73   0.0843 .
## Year2008             0.0978     0.1368    0.72   0.4746
## Year2009             0.0789     0.1412    0.56   0.5766
## Year2010            -0.0827     0.1444   -0.57   0.5670
## Year2011             0.0209     0.1344    0.16   0.8762
## Year2012            -0.1215     0.1332   -0.91   0.3619
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.929
## Multiple R-squared:  0.0259, Adjusted R-squared:  0.0203
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
##  5 observations c(1456,1746,1963,1964,1965)
##   are outliers with |weight| = 0 ( < 3.2e-05);
##  281 weights are ~= 1. The remaining 2832 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0192  0.8730  0.9460  0.9130  0.9840  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          3.21e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##                     GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.019  1           1.010
## Year              1.019 16           1.001
```




```
## [1] "List of  32 outliers with residuals above  2.5"
##         ScopusId  NLCS Year OneField Fields residuals
## 77    0001903821 3.758 1996     1210      1     2.720
## 78    0004755588 3.620 1996     1210      1     2.584
## 79    0011299279 4.607 1996     1210      1     3.569
## 81    0242454754 3.878 1996     1210      1     2.840
## 337   0001194993 4.390 1998     1210      1     3.037
## 1487 33847791651 4.223 2005     1210      1     2.981
## 1503 33646729315 3.969 2005     1210      1     2.725
## 1696 60950491226 6.170 2006     1210      2     4.965
## 2029 34047192756 4.513 2007     1210      1     3.163
## 2031 34047235257 4.790 2007     1210      1     3.440
## 2039 36049016758 7.080 2007     1210      2     5.802
## 2130 60950499484 3.660 2008     1210      2     2.524
## 2289 57049157739 5.668 2008     1210      2     4.534
## 2290 57049173452 5.953 2008     1210      2     4.819
## 2291 57049179599 6.823 2008     1210      2     5.689
## 2300 42349109603 4.568 2008     1210      1     3.432
## 2321 38949105986 3.882 2008     1210      1     2.676
## 2323 38949195334 4.014 2008     1210      1     2.808
## 2438 76849084888 4.006 2009     1210      1     2.819
## 2565 63049091821 3.645 2009     1210      1     2.600
## 2579 75149146477 4.172 2009     1210      2     3.057
## 2735 77955800148 3.635 2010     1210      2     2.681
## 2779 77952935263 3.635 2010     1210      1     2.680
## 2797 77951477155 3.967 2010     1210      1     3.013
## 2894 77956500865 4.776 2010     1208      3     3.822
```

```
## 2895 78649472664 5.064 2010     1208      3     4.038
## 2936 84858822292 3.704 2011     1210      2     2.647
## 3123 79954543516 3.648 2011     1210      1     2.661
## 3219 79960456514 3.730 2011     1210      2     2.673
## 3248 82455232963 4.204 2011     1208      3     3.145
## 3249 82455244076 4.997 2011     1208      3     3.940
## 3389 84869152027 3.467 2012     1210      3     2.552
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.4430 -0.6765  0.0278  0.6212  5.7987
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.04123    0.12426    8.38   <2e-16 ***
## FirstAuthorFemale1 -0.00776    0.03350   -0.23   0.8169
## Year1997            0.33248    0.15027    2.21   0.0270 *
## Year1998            0.24428    0.15796    1.55   0.1221
## Year1999            0.20282    0.15779    1.29   0.1987
## Year2000            0.29842    0.15243    1.96   0.0504 .
## Year2001            0.20000    0.14919    1.34   0.1801
## Year2002            0.21985    0.14167    1.55   0.1208
## Year2003            0.40172    0.14354    2.80   0.0052 **
## Year2004            0.22526    0.13744    1.64   0.1013
## Year2005            0.20556    0.13682    1.50   0.1331
## Year2006            0.16870    0.13726    1.23   0.2192
## Year2007            0.24010    0.13971    1.72   0.0858 .
## Year2008            0.09466    0.13650    0.69   0.4881
## Year2009            0.07904    0.14119    0.56   0.5757
## Year2010           -0.08355    0.14435   -0.58   0.5627
## Year2011            0.01906    0.13429    0.14   0.8871
## Year2012           -0.12319    0.13311   -0.93   0.3548
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.928
## Multiple R-squared:  0.0256, Adjusted R-squared:  0.0203
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
##  5 observations c(1456,1746,1963,1964,1965)
##   are outliers with |weight| = 0 ( < 3.2e-05);
##  277 weights are ~= 1. The remaining 2836 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0118  0.8720  0.9460  0.9130  0.9850  0.9990
```

```
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          3.21e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##                   GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.017  1           1.009
## Year             1.017 16           1.001
```

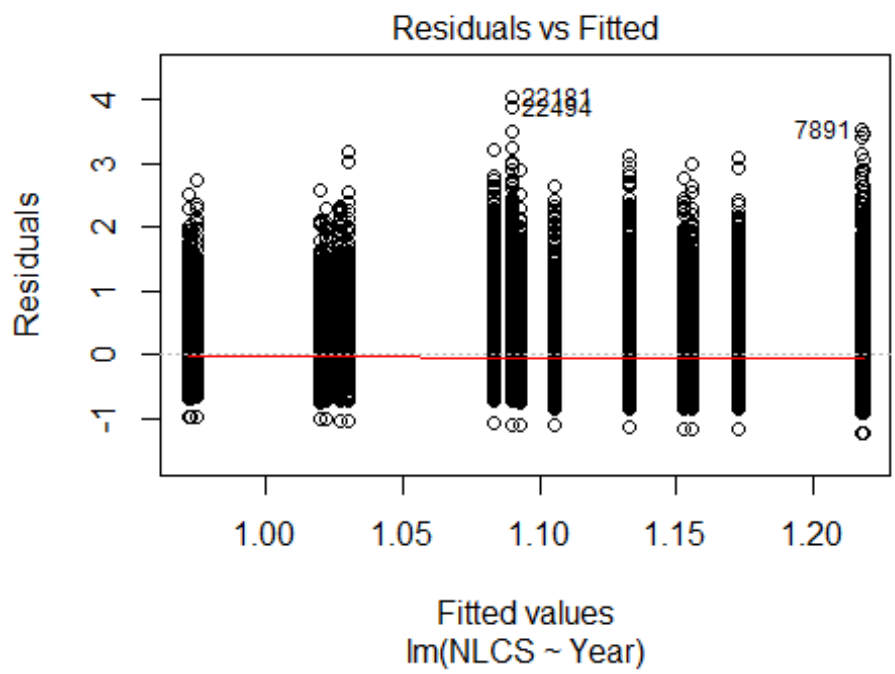



```
## [1] "List of  32 outliers with residuals above  2.5"
##         ScopusId  NLCS Year OneField Fields residuals
## 77    0001903821 3.758 1996     1210      1     2.720
## 78    0004755588 3.620 1996     1210      1     2.584
## 79    0011299279 4.607 1996     1210      1     3.569
## 81    0242454754 3.878 1996     1210      1     2.840
```

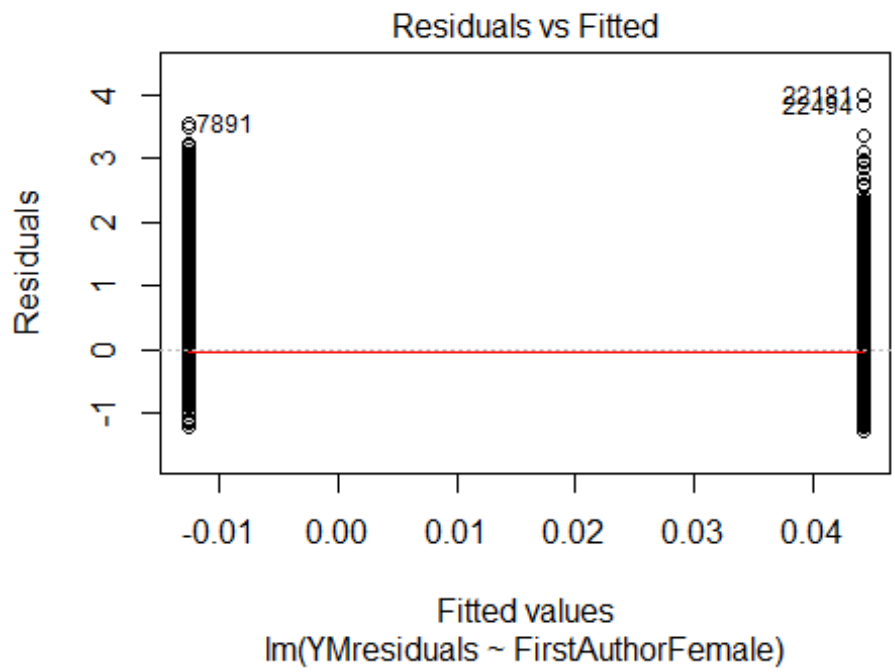
```
## 337   0001194993 4.390 1998     1210      1     3.037
## 1487 33847791651 4.223 2005     1210      1     2.981
## 1503 33646729315 3.969 2005     1210      1     2.725
## 1696 60950491226 6.170 2006     1210      2     4.965
## 2029 34047192756 4.513 2007     1210      1     3.163
## 2031 34047235257 4.790 2007     1210      1     3.440
## 2039 36049016758 7.080 2007     1210      2     5.802
## 2130 60950499484 3.660 2008     1210      2     2.524
## 2289 57049157739 5.668 2008     1210      2     4.534
## 2290 57049173452 5.953 2008     1210      2     4.819
## 2291 57049179599 6.823 2008     1210      2     5.689
## 2300 42349109603 4.568 2008     1210      1     3.432
## 2321 38949105986 3.882 2008     1210      1     2.676
## 2323 38949195334 4.014 2008     1210      1     2.808
## 2438 76849084888 4.006 2009     1210      1     2.819
## 2565 63049091821 3.645 2009     1210      1     2.600
## 2579 75149146477 4.172 2009     1210      2     3.057
## 2735 77955800148 3.635 2010     1210      2     2.681
## 2779 77952935263 3.635 2010     1210      1     2.680
## 2797 77951477155 3.967 2010     1210      1     3.013
## 2894 77956500865 4.776 2010     1208      3     3.822
## 2895 78649472664 5.064 2010     1208      3     4.038
## 2936 84858822292 3.704 2011     1210      2     2.647
## 3123 79954543516 3.648 2011     1210      1     2.661
## 3219 79960456514 3.730 2011     1210      2     2.673
## 3248 82455232963 4.204 2011     1208      3     3.145
## 3249 82455244076 4.997 2011     1208      3     3.940
## 3389 84869152027 3.467 2012     1210      3     2.552
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.4456 -0.6775  0.0352  0.6166  5.8056
##
## Coefficients:
##                   Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.0349     0.1242    8.33   <2e-16 ***
## LastAuthorFemale1   0.0107     0.0335    0.32   0.7488
## Year1997            0.3319     0.1500    2.21   0.0270 *
## Year1998            0.2438     0.1577    1.55   0.1223
## Year1999            0.2027     0.1576    1.29   0.1986
## Year2000            0.2967     0.1524    1.95   0.0516 .
## Year2001            0.1997     0.1490    1.34   0.1804
## Year2002            0.2189     0.1414    1.55   0.1218
## Year2003            0.4000     0.1433    2.79   0.0053 **
## Year2004            0.2239     0.1372    1.63   0.1027
```

```
## Year2005            0.2046     0.1366    1.50   0.1342
## Year2006            0.1683     0.1371    1.23   0.2197
## Year2007            0.2395     0.1395    1.72   0.0861 .
## Year2008            0.0935     0.1362    0.69   0.4923
## Year2009            0.0768     0.1410    0.54   0.5862
## Year2010           -0.0848     0.1441   -0.59   0.5561
## Year2011            0.0171     0.1340    0.13   0.8983
## Year2012           -0.1250     0.1328   -0.94   0.3470
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.928
## Multiple R-squared:  0.0256, Adjusted R-squared:  0.0203
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
##  5 observations c(1456,1746,1963,1964,1965)
##   are outliers with |weight| = 0 ( < 3.2e-05);
##  273 weights are ~= 1. The remaining 2840 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0121  0.8730  0.9460  0.9130  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          3.21e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                    psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Sample size for the above analysis:  3118"
## [1] ""
## [1] ""
## [1] "##################################"
## [1] "Analysis of AJSC 1211"
## [1] "##################################"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  970  968 1026 1014 1079 1195 1230 1125 1163 1239 1304 1526 1508 1453 1603
## 2011 2012
## 2092 2284
```

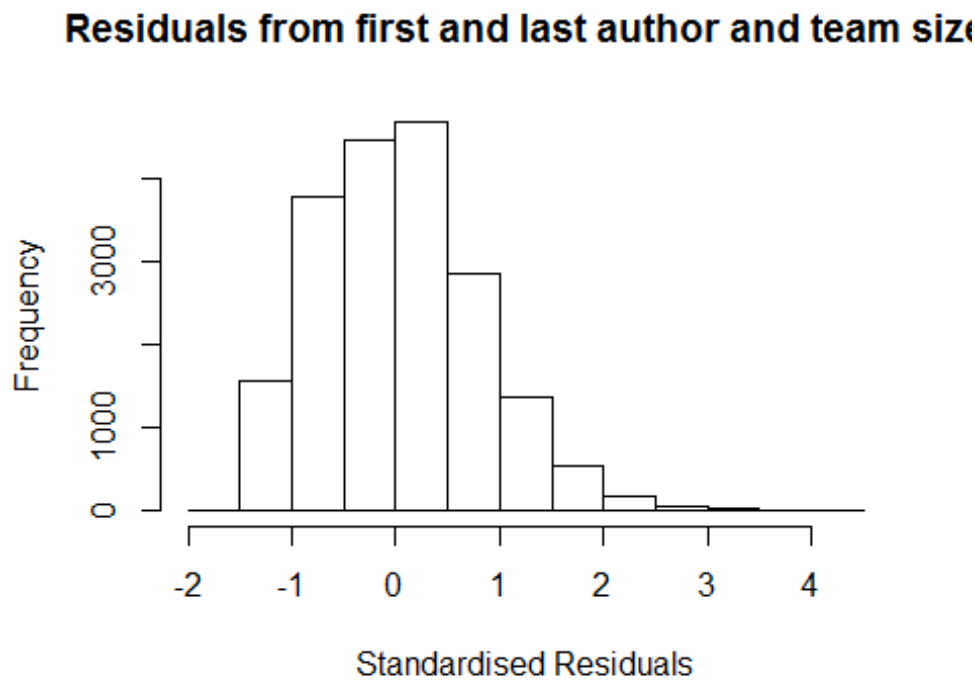
```
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  832  827  889  879  916 1001 1054  974 1020 1081 1139 1319 1313 1293 1408
## 2011 2012
## 1854 2024
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  820  817  879  866  899  990 1046  958 1005 1061 1119 1303 1288 1271 1385
## 2011 2012
## 1815 1982
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
##  Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 360, df = 16, p-value <2e-16
```




```
##
##  Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 0.02, df = 1, p-value = 0.9
```




```
## [1] "Female first author team size 2018 geometric mean: 1.31533856305557"
## [1] "Male first author team size 2018 geometric mean: 1.13494479549824"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 450000, p-value = 1e-10
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 1.32617582238127"
## [1] "Male last author team size 2018 geometric mean: 1.13179548037773"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 450000, p-value = 2e-10
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 2.723  1           1.650
## LastAuthorFemale  2.679  1           1.637
## UniqueAuthors     1.121  4           1.014
## Year              1.029 16           1.001
```

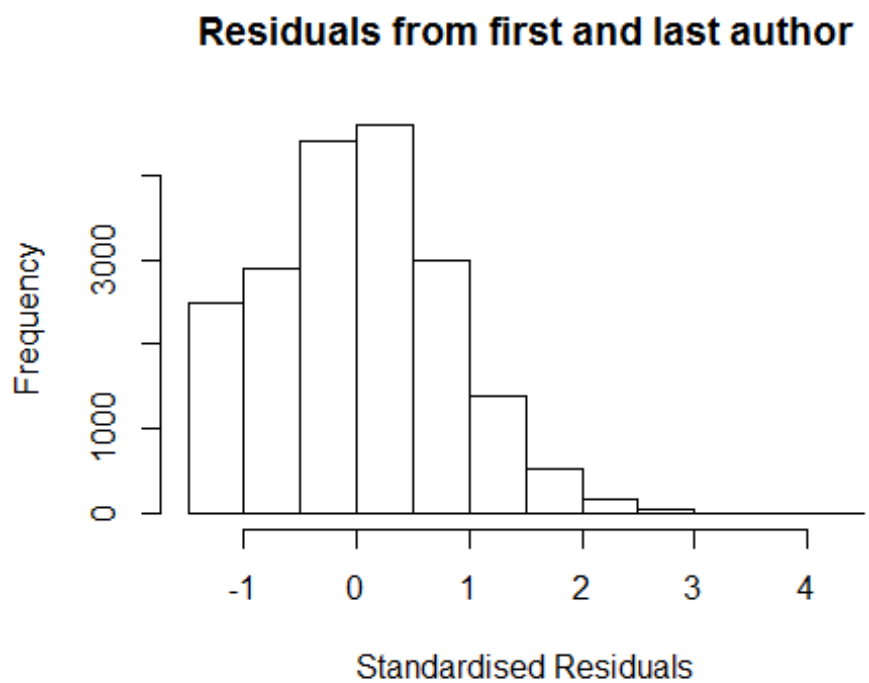



```
## [1] "List of  65 outliers with residuals above  2.5"
##          ScopusId  NLCS Year OneField Fields residuals
## 18    33845343748 4.229 1996     1211      1     3.274
## 574    0042684693 4.071 1996     1211      2     2.898
## 1282   0000478967 3.723 1997     1211      1     2.809
## 3106   0001217243 3.488 1999     1211      1     2.569
## 4471   0039916411 3.596 2000     1211      1     2.626
## 6568   0141557891 3.664 2002     1211      1     2.534
## 6646   3543089307 3.761 2002     1211      1     2.631
## 6659  21044438693 3.851 2002     1211      2     2.721
## 6991  33748322829 4.277 2002     1211      1     3.147
## 7078   1442335800 3.694 2002     1202      6     2.564
## 7106  17044385870 3.779 2002     1211      1     2.649
## 7384  33746151829 4.122 2002     1211      1     2.774
## 7415   0036526749 3.670 2002     1202      6     2.533
## 7617  70449876423 3.678 2002     1211      1     2.548
## 7669  61149169393 4.690 2002     1202      2     3.560
## 7891   0141888472 4.749 2003     1211      1     3.634
## 8121  33644775245 4.126 2003     1211      1     3.003
## 8249  33746156540 3.854 2003     1211      1     2.739
## 8520  33749449099 3.681 2003     1211      1     2.558
## 8654  11244330107 4.637 2003     1211      2     3.090
## 8739  32944461398 4.053 2003     1211      1     2.938
## 8839  18444365755 4.381 2003     1211      2     3.040
## 8924  33746453700 4.000 2003     1211      1     2.885
## 8971  33645120687 3.602 2004     1211      1     2.507
## 9555  33646900966 4.247 2004     1211      1     3.152
```

```
## 9597  36448965830 4.091 2004     1211      1     2.996
## 10166 34247240971 4.269 2005     1211      1     3.236
## 10185 34249698573 3.833 2005     1211      1     2.800
## 10201 37549039248 3.774 2005     1211      1     2.741
## 10383 32944462886 3.596 2005     1211      2     2.563
## 10627 33646497937 3.765 2005     1211      2     2.514
## 10629 33646527844 3.946 2005     1211      2     2.695
## 10685 77950024425 3.762 2005     1211      1     2.729
## 10821 36849082914 3.801 2005     1211      1     2.768
## 10974 33645117514 3.993 2005     1211      1     2.960
## 12260 37249044124 3.764 2006     1211      1     2.728
## 12838 57749198329 3.591 2007     1211      1     2.502
## 12841 58649090384 3.923 2007     1211      1     2.624
## 13304 58549101578 3.933 2007     1211      1     2.852
## 13719 34249694572 3.928 2007     1211      1     2.847
## 15224 60949164169 3.720 2008     1208      3     2.633
## 15305 42649133176 3.800 2008     1211      1     2.713
## 15657 61249393642 4.134 2008     1211      1     3.039
## 17020 70350428184 3.979 2009     1211      1     2.961
## 17021 77950534094 3.622 2009     1211      1     2.604
## 19142 84861758113 3.656 2011     1211      1     2.666
## 19669 79960983777 3.829 2011     1211      1     2.839
## 19717 80052611806 3.892 2011     1211      1     2.902
## 20383 79955056576 3.732 2011     1211      1     2.734
## 20834 79951935220 3.573 2011     1211      1     2.583
## 20937 80054858862 3.702 2011     1211      1     2.712
## 21253 84868593651 3.841 2012     1211      1     2.854
## 21532 84868280635 3.541 2012     1202      3     2.562
## 21801 84867218153 4.122 2012     1211      1     3.135
## 21927 84867202611 3.525 2012     1202      2     2.546
## 22027 84865019853 3.991 2012     1211      1     3.012
## 22041 84865421521 3.493 2012     1211      1     2.514
## 22181 84866640267 5.127 2012     1208      3     4.140
## 22274 84865419857 3.785 2012     1211      1     2.806
## 22494 84861636960 4.965 2012     1211      1     3.978
## 22554 84860861091 3.804 2012     1211      2     2.607
## 22606 84869471464 4.071 2012     1211      2     3.092
## 23000 84862637543 3.697 2012     1211      1     2.710
## 23275 84857943235 3.757 2012     1211      1     2.778
## 23342 84877747913 4.581 2012     1211      1     3.602
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
UniqueAuthors +
##     Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##     k.max = 1000))
##  \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
```

```
## -1.50591 -0.52800 -0.00617  0.52019  4.13959
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)          0.95496    0.02429   39.32  < 2e-16 ***
## FirstAuthorFemale1   0.00625    0.02212    0.28  0.77757
## LastAuthorFemale1    0.00172    0.02223    0.08  0.93822
## UniqueAuthors2       0.21803    0.02140   10.19  < 2e-16 ***
## UniqueAuthors3       0.29807    0.02994    9.95  < 2e-16 ***
## UniqueAuthors4       0.42338    0.03656   11.58  < 2e-16 ***
## UniqueAuthors5       0.47891    0.03010   15.91  < 2e-16 ***
## Year1997            -0.04143    0.03532   -1.17  0.24092
## Year1998             0.00422    0.03380    0.13  0.90051
## Year1999            -0.03558    0.03335   -1.07  0.28595
## Year2000             0.01523    0.03317    0.46  0.64616
## Year2001            -0.02984    0.03258   -0.92  0.35978
## Year2002             0.17455    0.03705    4.71  2.5e-06 ***
## Year2003             0.16032    0.03805    4.21  2.5e-05 ***
## Year2004             0.14027    0.03631    3.86  0.00011 ***
## Year2005             0.07847    0.03644    2.15  0.03131 *
## Year2006             0.08104    0.03314    2.45  0.01448 *
## Year2007             0.12586    0.03204    3.93  8.6e-05 ***
## Year2008             0.13238    0.03232    4.10  4.2e-05 ***
## Year2009             0.06321    0.03224    1.96  0.04995 *
## Year2010             0.00035    0.03112    0.01  0.99102
## Year2011             0.03535    0.03048    1.16  0.24614
## Year2012             0.02448    0.03202    0.76  0.44445
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.797
## Multiple R-squared:  0.0253, Adjusted R-squared:  0.0242
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  2 observations c(18346,18606) are outliers with |weight| = 0 ( < 5.1e-
06);
##  1667 weights are ~= 1. The remaining 17835 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0027  0.8670  0.9530  0.9130  0.9860  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          5.13e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
```

```
##           0        1000           0
##                   psi           subsampling                    cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 2.880  1           1.697
## LastAuthorFemale  2.877  1           1.696
## Year              1.007 16           1.000
```

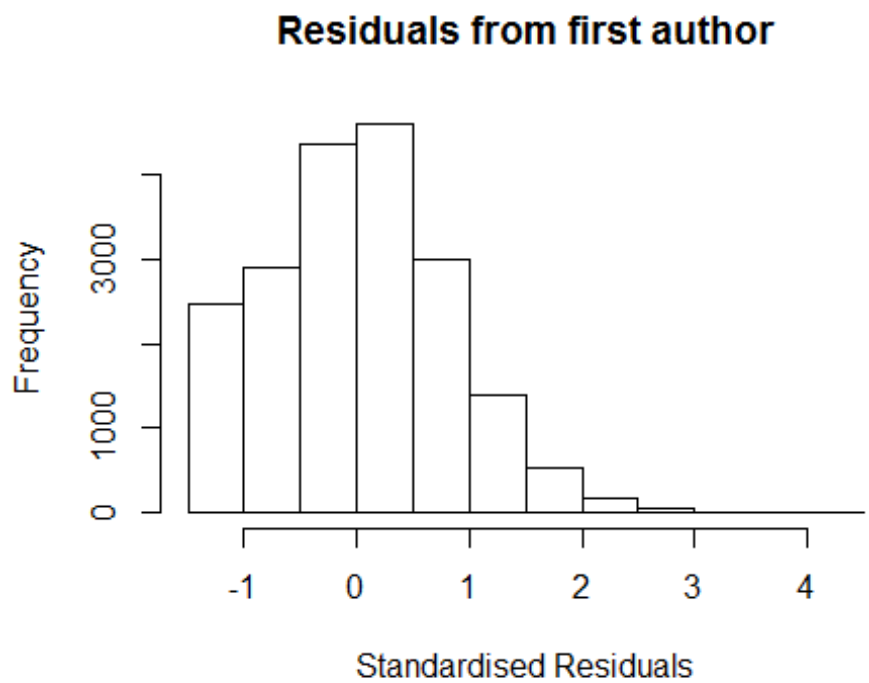



```
## [1] "List of  64 outliers with residuals above  2.5"
##          ScopusId  NLCS Year OneField Fields residuals
## 18    33845343748 4.229 1996     1211      1     3.245
## 463    0003143117 3.579 1996     1211      2     2.595
## 574    0042684693 4.071 1996     1211      2     3.087
## 1282   0000478967 3.723 1997     1211      1     2.788
## 3106   0001217243 3.488 1999     1211      1     2.555
## 4471   0039916411 3.596 2000     1211      1     2.602
## 6568   0141557891 3.664 2002     1211      1     2.505
## 6646   3543089307 3.761 2002     1211      1     2.602
## 6659  21044438693 3.851 2002     1211      2     2.692
## 6991  33748322829 4.277 2002     1211      1     3.118
## 7078   1442335800 3.694 2002     1202      6     2.535
## 7106  17044385870 3.779 2002     1211      1     2.620
```

```
## 7384  33746151829 4.122 2002     1211      1     2.963
## 7617  70449876423 3.678 2002     1211      1     2.519
## 7669  61149169393 4.690 2002     1202      2     3.531
## 7891   0141888472 4.749 2003     1211      1     3.600
## 8121  33644775245 4.126 2003     1211      1     2.917
## 8249  33746156540 3.854 2003     1211      1     2.705
## 8654  11244330107 4.637 2003     1211      2     3.428
## 8696  19944362281 3.702 2003     1211      1     2.553
## 8739  32944461398 4.053 2003     1211      1     2.904
## 8839  18444365755 4.381 2003     1211      2     3.172
## 8924  33746453700 4.000 2003     1211      1     2.851
## 9555  33646900966 4.247 2004     1211      1     3.118
## 9597  36448965830 4.091 2004     1211      1     2.962
## 10166 34247240971 4.269 2005     1211      1     3.204
## 10185 34249698573 3.833 2005     1211      1     2.768
## 10201 37549039248 3.774 2005     1211      1     2.709
## 10383 32944462886 3.596 2005     1211      2     2.531
## 10627 33646497937 3.765 2005     1211      2     2.700
## 10629 33646527844 3.946 2005     1211      2     2.881
## 10685 77950024425 3.762 2005     1211      1     2.697
## 10821 36849082914 3.801 2005     1211      1     2.736
## 10970 17444424014 3.828 2005     1211      1     2.763
## 10974 33645117514 3.993 2005     1211      1     2.928
## 12260 37249044124 3.764 2006     1211      1     2.698
## 12841 58649090384 3.923 2007     1211      1     2.811
## 13304 58549101578 3.933 2007     1211      1     2.821
## 13719 34249694572 3.928 2007     1211      1     2.816
## 15224 60949164169 3.720 2008     1208      3     2.604
## 15305 42649133176 3.800 2008     1211      1     2.684
## 15657 61249393642 4.134 2008     1211      1     2.958
## 17020 70350428184 3.979 2009     1211      1     2.928
## 17021 77950534094 3.622 2009     1211      1     2.571
## 19142 84861758113 3.656 2011     1211      1     2.633
## 19161 80255137237 3.700 2011     1211      2     2.677
## 19669 79960983777 3.829 2011     1211      1     2.806
## 19717 80052611806 3.892 2011     1211      1     2.869
## 20383 79955056576 3.732 2011     1211      1     2.648
## 20834 79951935220 3.573 2011     1211      1     2.550
## 20937 80054858862 3.702 2011     1211      1     2.679
## 21253 84868593651 3.841 2012     1211      1     2.770
## 21532 84868280635 3.541 2012     1202      3     2.530
## 21801 84867218153 4.122 2012     1211      1     3.051
## 21927 84867202611 3.525 2012     1202      2     2.514
## 22027 84865019853 3.991 2012     1211      1     2.980
## 22181 84866640267 5.127 2012     1208      3     4.056
## 22274 84865419857 3.785 2012     1211      1     2.774
## 22494 84861636960 4.965 2012     1211      1     3.894
## 22554 84860861091 3.804 2012     1211      2     2.793
## 22606 84869471464 4.071 2012     1211      2     3.060
## 23000 84862637543 3.697 2012     1211      1     2.626
```

```
## 23275 84857943235 3.757 2012     1211      1     2.746
## 23342 84877747913 4.581 2012     1211      1     3.570
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##       Min        1Q    Median        3Q       Max
## -1.219201 -0.549416 -0.000289  0.538232  4.056105
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         0.98384    0.02508   39.22  < 2e-16 ***
## FirstAuthorFemale1  0.06498    0.02349    2.77   0.0057 **
## LastAuthorFemale1  -0.00486    0.02385   -0.20   0.8386
## Year1997           -0.04844    0.03626   -1.34   0.1817
## Year1998            0.00175    0.03456    0.05   0.9595
## Year1999           -0.05100    0.03421   -1.49   0.1360
## Year2000            0.01045    0.03391    0.31   0.7579
## Year2001           -0.03757    0.03342   -1.12   0.2610
## Year2002            0.17523    0.03773    4.64  3.4e-06 ***
## Year2003            0.16515    0.03882    4.25  2.1e-05 ***
## Year2004            0.14560    0.03695    3.94  8.2e-05 ***
## Year2005            0.08117    0.03701    2.19   0.0283 *
## Year2006            0.08202    0.03402    2.41   0.0159 *
## Year2007            0.12805    0.03279    3.90  9.5e-05 ***
## Year2008            0.13178    0.03298    4.00  6.5e-05 ***
## Year2009            0.06745    0.03316    2.03   0.0420 *
## Year2010            0.00406    0.03208    0.13   0.8992
## Year2011            0.03961    0.03144    1.26   0.2077
## Year2012            0.02693    0.03285    0.82   0.4123
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.811
## Multiple R-squared:  0.00856,    Adjusted R-squared:  0.00765
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  2 observations c(18346,18606) are outliers with |weight| = 0 ( < 5.1e-
06);
##  1604 weights are ~= 1. The remaining 17898 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0103  0.8630  0.9510  0.9150  0.9860  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
```

```
##          1.00e-07          1.00e-07          5.13e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.005  1           1.003
## Year              1.005 16           1.000
```

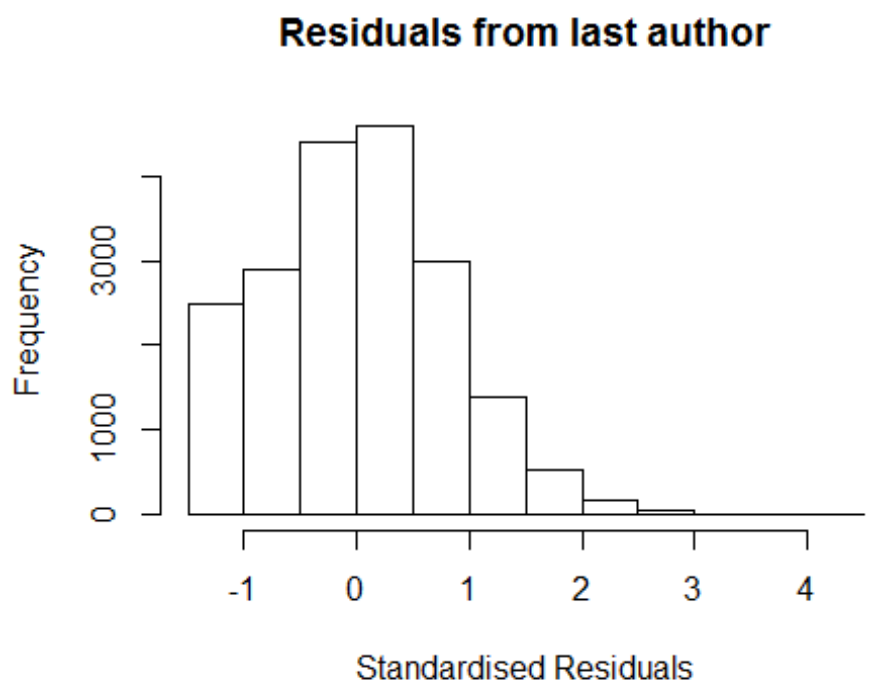



```
## [1] "List of  64 outliers with residuals above  2.5"
##          ScopusId  NLCS Year OneField Fields residuals
## 18    33845343748 4.229 1996     1211      1     3.245
## 463    0003143117 3.579 1996     1211      2     2.595
## 574    0042684693 4.071 1996     1211      2     3.087
## 1282   0000478967 3.723 1997     1211      1     2.788
## 3106   0001217243 3.488 1999     1211      1     2.555
## 4471   0039916411 3.596 2000     1211      1     2.602
## 6568   0141557891 3.664 2002     1211      1     2.505
## 6646   3543089307 3.761 2002     1211      1     2.602
```

```
## 6659  21044438693 3.851 2002     1211      2     2.692
## 6991  33748322829 4.277 2002     1211      1     3.118
## 7078   1442335800 3.694 2002     1202      6     2.535
## 7106  17044385870 3.779 2002     1211      1     2.620
## 7384  33746151829 4.122 2002     1211      1     2.963
## 7617  70449876423 3.678 2002     1211      1     2.519
## 7669  61149169393 4.690 2002     1202      2     3.531
## 7891   0141888472 4.749 2003     1211      1     3.600
## 8121  33644775245 4.126 2003     1211      1     2.917
## 8249  33746156540 3.854 2003     1211      1     2.705
## 8654  11244330107 4.637 2003     1211      2     3.428
## 8696  19944362281 3.702 2003     1211      1     2.553
## 8739  32944461398 4.053 2003     1211      1     2.904
## 8839  18444365755 4.381 2003     1211      2     3.172
## 8924  33746453700 4.000 2003     1211      1     2.851
## 9555  33646900966 4.247 2004     1211      1     3.118
## 9597  36448965830 4.091 2004     1211      1     2.962
## 10166 34247240971 4.269 2005     1211      1     3.204
## 10185 34249698573 3.833 2005     1211      1     2.768
## 10201 37549039248 3.774 2005     1211      1     2.709
## 10383 32944462886 3.596 2005     1211      2     2.531
## 10627 33646497937 3.765 2005     1211      2     2.700
## 10629 33646527844 3.946 2005     1211      2     2.881
## 10685 77950024425 3.762 2005     1211      1     2.697
## 10821 36849082914 3.801 2005     1211      1     2.736
## 10970 17444424014 3.828 2005     1211      1     2.763
## 10974 33645117514 3.993 2005     1211      1     2.928
## 12260 37249044124 3.764 2006     1211      1     2.698
## 12841 58649090384 3.923 2007     1211      1     2.811
## 13304 58549101578 3.933 2007     1211      1     2.821
## 13719 34249694572 3.928 2007     1211      1     2.816
## 15224 60949164169 3.720 2008     1208      3     2.604
## 15305 42649133176 3.800 2008     1211      1     2.684
## 15657 61249393642 4.134 2008     1211      1     2.958
## 17020 70350428184 3.979 2009     1211      1     2.928
## 17021 77950534094 3.622 2009     1211      1     2.571
## 19142 84861758113 3.656 2011     1211      1     2.633
## 19161 80255137237 3.700 2011     1211      2     2.677
## 19669 79960983777 3.829 2011     1211      1     2.806
## 19717 80052611806 3.892 2011     1211      1     2.869
## 20383 79955056576 3.732 2011     1211      1     2.648
## 20834 79951935220 3.573 2011     1211      1     2.550
## 20937 80054858862 3.702 2011     1211      1     2.679
## 21253 84868593651 3.841 2012     1211      1     2.770
## 21532 84868280635 3.541 2012     1202      3     2.530
## 21801 84867218153 4.122 2012     1211      1     3.051
## 21927 84867202611 3.525 2012     1202      2     2.514
## 22027 84865019853 3.991 2012     1211      1     2.980
## 22181 84866640267 5.127 2012     1208      3     4.056
## 22274 84865419857 3.785 2012     1211      1     2.774
```

```
## 22494 84861636960 4.965 2012     1211      1     3.894
## 22554 84860861091 3.804 2012     1211      2     2.793
## 22606 84869471464 4.071 2012     1211      2     3.060
## 23000 84862637543 3.697 2012     1211      1     2.626
## 23275 84857943235 3.757 2012     1211      1     2.746
## 23342 84877747913 4.581 2012     1211      1     3.570
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.21985 -0.54928 -0.00022  0.53853  4.05548
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         0.98373    0.02506   39.25  < 2e-16 ***
## FirstAuthorFemale1  0.06090    0.01392    4.37  1.2e-05 ***
## Year1997           -0.04846    0.03627   -1.34    0.181
## Year1998            0.00174    0.03457    0.05    0.960
## Year1999           -0.05101    0.03421   -1.49    0.136
## Year2000            0.01048    0.03391    0.31    0.757
## Year2001           -0.03762    0.03342   -1.13    0.260
## Year2002            0.17523    0.03773    4.64  3.4e-06 ***
## Year2003            0.16509    0.03882    4.25  2.1e-05 ***
## Year2004            0.14564    0.03694    3.94  8.1e-05 ***
## Year2005            0.08114    0.03701    2.19    0.028 *
## Year2006            0.08194    0.03402    2.41    0.016 *
## Year2007            0.12800    0.03280    3.90  9.5e-05 ***
## Year2008            0.13176    0.03299    3.99  6.5e-05 ***
## Year2009            0.06739    0.03317    2.03    0.042 *
## Year2010            0.00407    0.03208    0.13    0.899
## Year2011            0.03960    0.03144    1.26    0.208
## Year2012            0.02689    0.03285    0.82    0.413
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.811
## Multiple R-squared:  0.00856,    Adjusted R-squared:  0.0077
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  2 observations c(18346,18606) are outliers with |weight| = 0 ( < 5.1e-
06);
##  1603 weights are ~= 1. The remaining 17899 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0102  0.8630  0.9510  0.9150  0.9860  0.9990
## Algorithmic parameters:
```

```
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          5.13e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##                   GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.004  1           1.002
## Year             1.004 16           1.000
```

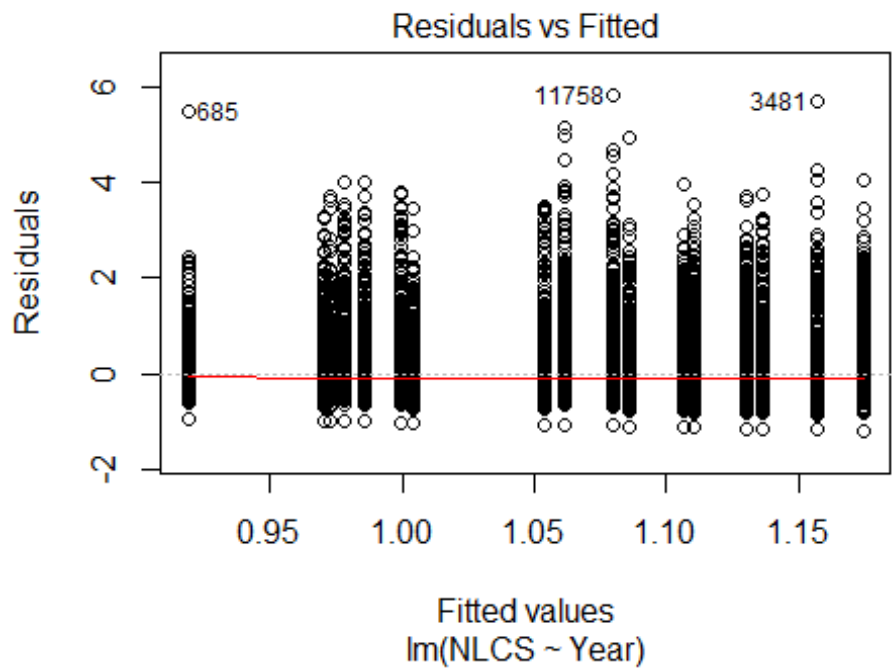



```
## [1] "List of  64 outliers with residuals above  2.5"
##          ScopusId  NLCS Year OneField Fields residuals
## 18    33845343748 4.229 1996     1211      1     3.245
## 463    0003143117 3.579 1996     1211      2     2.595
## 574    0042684693 4.071 1996     1211      2     3.087
## 1282   0000478967 3.723 1997     1211      1     2.788
## 3106   0001217243 3.488 1999     1211      1     2.555
```

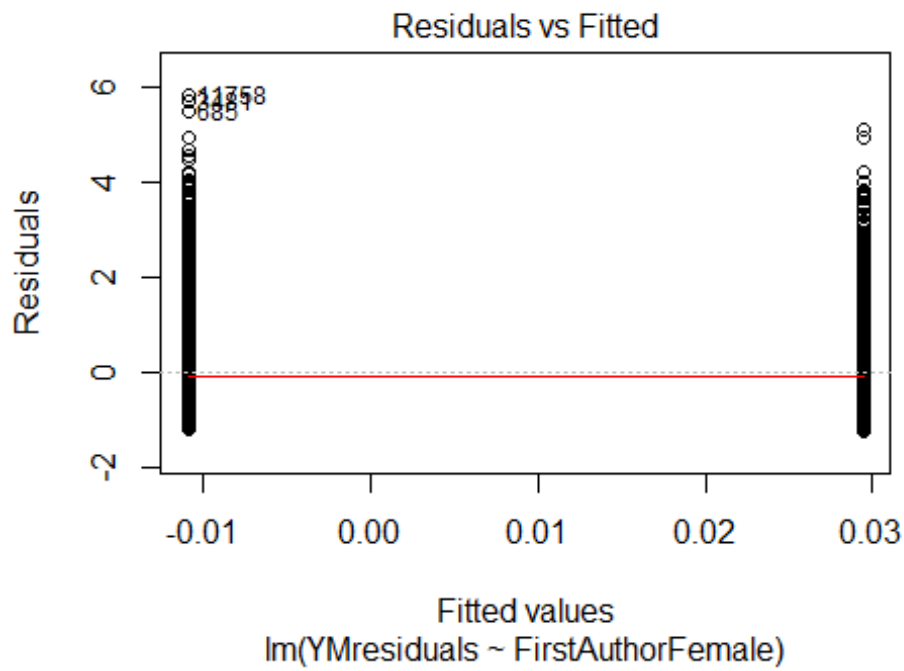
```
## 4471   0039916411 3.596 2000     1211      1     2.602
## 6568   0141557891 3.664 2002     1211      1     2.505
## 6646   3543089307 3.761 2002     1211      1     2.602
## 6659  21044438693 3.851 2002     1211      2     2.692
## 6991  33748322829 4.277 2002     1211      1     3.118
## 7078   1442335800 3.694 2002     1202      6     2.535
## 7106  17044385870 3.779 2002     1211      1     2.620
## 7384  33746151829 4.122 2002     1211      1     2.963
## 7617  70449876423 3.678 2002     1211      1     2.519
## 7669  61149169393 4.690 2002     1202      2     3.531
## 7891   0141888472 4.749 2003     1211      1     3.600
## 8121  33644775245 4.126 2003     1211      1     2.917
## 8249  33746156540 3.854 2003     1211      1     2.705
## 8654  11244330107 4.637 2003     1211      2     3.428
## 8696  19944362281 3.702 2003     1211      1     2.553
## 8739  32944461398 4.053 2003     1211      1     2.904
## 8839  18444365755 4.381 2003     1211      2     3.172
## 8924  33746453700 4.000 2003     1211      1     2.851
## 9555  33646900966 4.247 2004     1211      1     3.118
## 9597  36448965830 4.091 2004     1211      1     2.962
## 10166 34247240971 4.269 2005     1211      1     3.204
## 10185 34249698573 3.833 2005     1211      1     2.768
## 10201 37549039248 3.774 2005     1211      1     2.709
## 10383 32944462886 3.596 2005     1211      2     2.531
## 10627 33646497937 3.765 2005     1211      2     2.700
## 10629 33646527844 3.946 2005     1211      2     2.881
## 10685 77950024425 3.762 2005     1211      1     2.697
## 10821 36849082914 3.801 2005     1211      1     2.736
## 10970 17444424014 3.828 2005     1211      1     2.763
## 10974 33645117514 3.993 2005     1211      1     2.928
## 12260 37249044124 3.764 2006     1211      1     2.698
## 12841 58649090384 3.923 2007     1211      1     2.811
## 13304 58549101578 3.933 2007     1211      1     2.821
## 13719 34249694572 3.928 2007     1211      1     2.816
## 15224 60949164169 3.720 2008     1208      3     2.604
## 15305 42649133176 3.800 2008     1211      1     2.684
## 15657 61249393642 4.134 2008     1211      1     2.958
## 17020 70350428184 3.979 2009     1211      1     2.928
## 17021 77950534094 3.622 2009     1211      1     2.571
## 19142 84861758113 3.656 2011     1211      1     2.633
## 19161 80255137237 3.700 2011     1211      2     2.677
## 19669 79960983777 3.829 2011     1211      1     2.806
## 19717 80052611806 3.892 2011     1211      1     2.869
## 20383 79955056576 3.732 2011     1211      1     2.648
## 20834 79951935220 3.573 2011     1211      1     2.550
## 20937 80054858862 3.702 2011     1211      1     2.679
## 21253 84868593651 3.841 2012     1211      1     2.770
## 21532 84868280635 3.541 2012     1202      3     2.530
## 21801 84867218153 4.122 2012     1211      1     3.051
## 21927 84867202611 3.525 2012     1202      2     2.514
```

```
## 22027 84865019853 3.991 2012     1211      1     2.980
## 22181 84866640267 5.127 2012     1208      3     4.056
## 22274 84865419857 3.785 2012     1211      1     2.774
## 22494 84861636960 4.965 2012     1211      1     3.894
## 22554 84860861091 3.804 2012     1211      2     2.793
## 22606 84869471464 4.071 2012     1211      2     3.060
## 23000 84862637543 3.697 2012     1211      1     2.626
## 23275 84857943235 3.757 2012     1211      1     2.746
## 23342 84877747913 4.581 2012     1211      1     3.570
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.21295 -0.54995 -0.00271  0.53817  4.06262
##
## Coefficients:
##                   Estimate Std. Error t value Pr(>|t|)
## (Intercept)        0.98649    0.02513   39.25  < 2e-16 ***
## LastAuthorFemale1  0.05096    0.01408    3.62   0.0003 ***
## Year1997          -0.04896    0.03630   -1.35   0.1775
## Year1998           0.00129    0.03462    0.04   0.9703
## Year1999          -0.05158    0.03425   -1.51   0.1321
## Year2000           0.01043    0.03394    0.31   0.7587
## Year2001          -0.03816    0.03345   -1.14   0.2540
## Year2002           0.17550    0.03778    4.65  3.4e-06 ***
## Year2003           0.16460    0.03886    4.24  2.3e-05 ***
## Year2004           0.14618    0.03697    3.95  7.7e-05 ***
## Year2005           0.08121    0.03702    2.19   0.0282 *
## Year2006           0.08078    0.03407    2.37   0.0177 *
## Year2007           0.12767    0.03285    3.89   0.0001 ***
## Year2008           0.13140    0.03304    3.98  7.0e-05 ***
## Year2009           0.06689    0.03322    2.01   0.0441 *
## Year2010           0.00446    0.03213    0.14   0.8896
## Year2011           0.03952    0.03147    1.26   0.2092
## Year2012           0.02693    0.03287    0.82   0.4126
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.811
## Multiple R-squared:  0.00825,    Adjusted R-squared:  0.00738
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  2 observations c(18346,18606) are outliers with |weight| = 0 ( < 5.1e-
06);
##  1594 weights are ~= 1. The remaining 17908 ones are summarized as
```

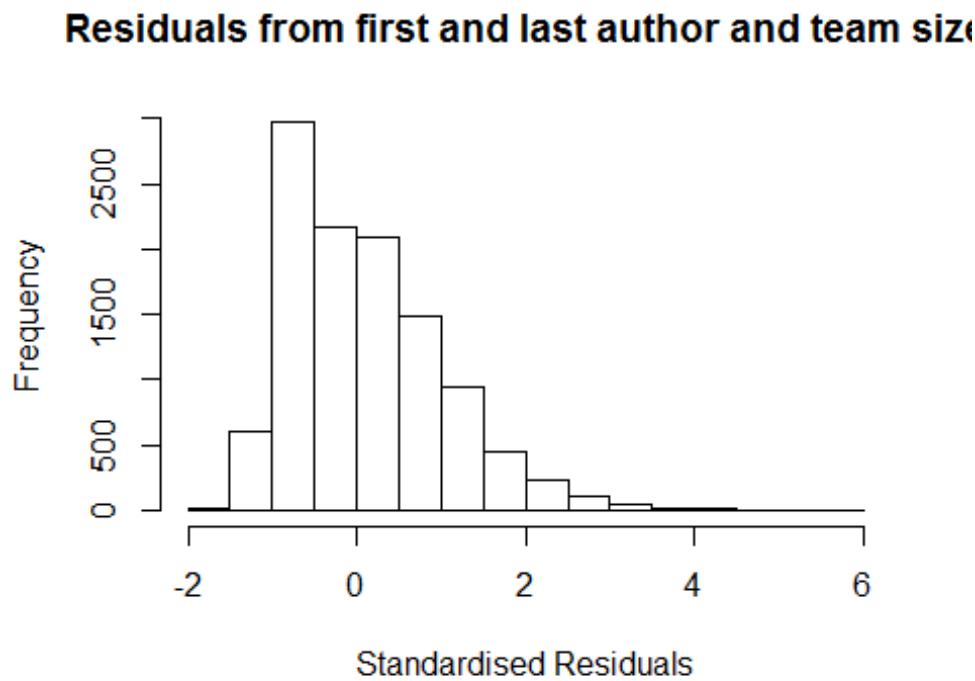
```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0109  0.8630  0.9520  0.9150  0.9860  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          5.13e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Sample size for the above analysis:  19504"
## [1] ""
## [1] ""
## [1] "####################################"
## [1] "Analysis of AJSC 1212"
## [1] "####################################"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  508  386  555  517  578  596  684  749  649  848  770  927  945  894  969
## 2011 2012
## 1219 1318
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  451  336  481  439  482  518  607  655  562  722  689  811  829  786  832
## 2011 2012
## 1046 1121
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  446  332  475  425  473  512  600  639  553  714  677  806  816  766  800
## 2011 2012
## 1027 1086
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
##  Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 95, df = 16, p-value = 3e-13
```




```
##
##  Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 0.024, df = 1, p-value = 0.9
```




```
## [1] "Female first author team size 2018 geometric mean: 1.25800092919429"
## [1] "Male first author team size 2018 geometric mean: 1.16888310649418"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 190000, p-value = 0.01
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 1.27620179649908"
## [1] "Male last author team size 2018 geometric mean: 1.16019526353907"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 190000, p-value = 0.002
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 4.454  1           2.110
## LastAuthorFemale  4.488  1           2.118
## UniqueAuthors     1.094  4           1.011
## Year              1.049 16           1.001
```




```
## [1] "List of  192 outliers with residuals above  2.5"
##          ScopusId  NLCS Year OneField Fields residuals
## 141    0000339785 4.220 1996     1212      1     2.943
## 144    0001811907 4.077 1996     1212      1     2.877
## 145    0002009675 3.936 1996     1212      1     3.133
## 146    0002288148 4.141 1996     1212      1     3.338
## 147    0002291743 4.518 1996     1212      1     3.726
## 151    0012467113 4.997 1996     1212      1     4.205
## 152    0013066072 3.885 1996     1212      1     3.082
## 153    0039860262 3.616 1996     1212      1     2.824
## 155    0242607951 3.961 1996     1212      1     3.158
## 160   21344473380 4.412 1996     1212      1     3.609
## 164   33745249484 3.390 1996     1212      1     2.587
## 282    0001565086 3.590 1996     1212      2     2.787
## 309    0042684693 4.071 1996     1211      2     2.871
## 680   84994881613 3.341 1997     1212      1     2.530
## 685    0001431225 6.397 1997     1212      1     5.066
## 1106   0000315626 3.948 1998     1212      1     2.573
## 1107   0000473275 4.202 1998     1212      1     2.746
## 1112   0002098619 3.653 1998     1212      1     2.686
## 1150  33748132596 6.001 1998     1212      1     4.510
## 1153  33748194686 4.108 1998     1212      1     2.652
## 1670   0001718525 4.602 1999     1212      1     3.243
## 1679   0033238699 3.803 1999     1212      1     2.967
## 1682   0041359166 4.417 1999     1212      1     3.173
## 1701  60950059093 3.550 1999     1212      1     2.714
## 1803  33749819408 3.636 1999     1212      2     2.790
```

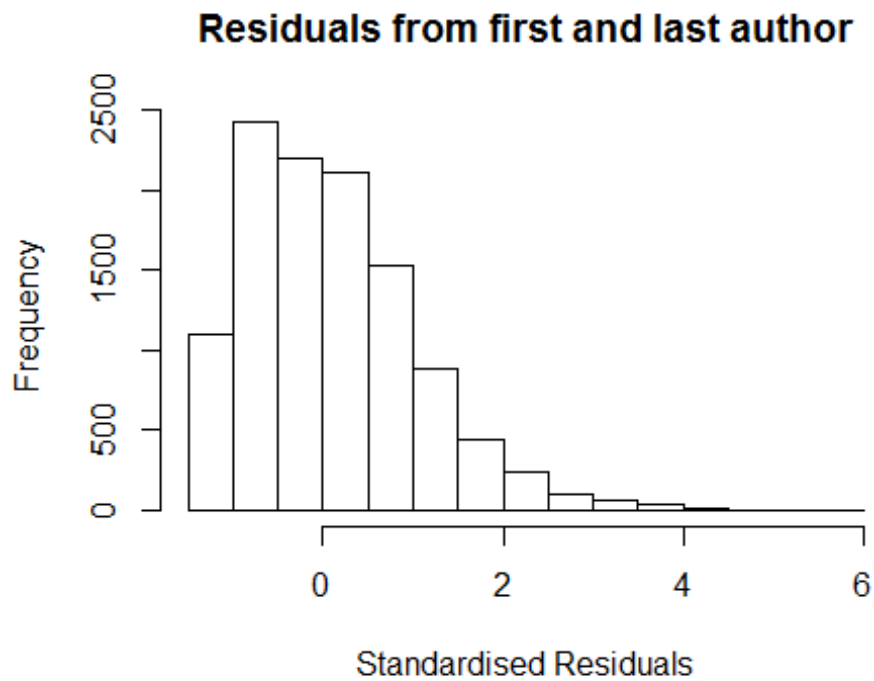
```
## 2114  33748347287 3.974 2000     1212      1     3.043
## 2181   0007337668 3.974 2000     1212      1     2.733
## 2191   0038437363 4.452 2000     1212      1     3.123
## 2194   0242564570 3.458 2000     1212      1     2.527
## 2196   1642316438 4.462 2000     1212      1     3.221
## 2610  61449201817 3.518 2001     1212      1     2.651
## 2784  33748072453 3.503 2001     1212      1     2.636
## 2827   0001415459 3.882 2001     1212      1     2.706
## 2829   0011182447 3.891 2001     1212      1     3.024
## 2830   0038694847 4.192 2001     1212      1     2.725
## 3442   0036747023 5.208 2002     1212      1     3.709
## 3443   0036748060 5.399 2002     1212      1     4.385
## 3444   0036748067 4.725 2002     1212      1     3.314
## 3445   0036749032 3.770 2002     1212      1     2.746
## 3481   0036623355 6.838 2002     1212      1     5.814
## 3517   0036521906 4.507 2002     1212      1     3.096
## 3545  33645260213 3.625 2002     1212      1     2.611
## 3564  60949697784 3.957 2002     1212      1     2.933
## 3751  61249583521 3.963 2002     1208      2     2.939
## 3756  61449442102 3.533 2002     1208      2     2.509
## 3761  62449174089 3.672 2002     1208      2     2.648
## 3763  62449296550 4.063 2002     1208      2     3.049
## 3769  67649155884 3.567 2002     1202      2     2.553
## 3878   0036526749 3.670 2002     1202      6     2.656
## 3880   1442335800 3.694 2002     1202      6     2.670
## 3914  60950296097 4.037 2003     1212      1     2.987
## 3964  70449986080 3.687 2003     1212      1     2.637
## 4112   0038489174 3.779 2003     1212      1     2.740
## 4235  61249404750 3.718 2003     1212      1     2.668
## 4356  11244330107 4.637 2003     1211      2     2.997
## 4361  18444365755 4.381 2003     1211      2     2.944
## 4364  29744456134 5.218 2003     1208      2     4.168
## 4383  34547980224 3.959 2003     1208      2     2.909
## 4412  61249716500 4.045 2003     1208      2     2.995
## 4657  37949018607 3.540 2004     1212      1     2.540
## 4672  60949844649 3.775 2004     1212      1     2.786
## 4705  61949157081 3.732 2004     1212      1     2.732
## 5033  84992828474 5.075 2004     1212      1     4.075
## 5076  85040393605 4.005 2004     1212      1     3.005
## 5316  36749058235 3.686 2005     1212      1     2.717
## 5337  60949946845 3.720 2005     1212      1     2.762
## 5538  62749191854 3.744 2005     1202      3     2.786
## 5592  61149560064 3.651 2005     1212      1     2.682
## 5620  33646527844 3.946 2005     1211      2     2.580
## 5672  27844473360 3.904 2005     1212      1     2.538
## 5673  27844485164 3.959 2005     1212      1     2.593
## 5674  27844500221 4.768 2005     1212      1     3.799
## 5675  27844592291 3.959 2005     1212      1     2.593
## 5765  27844515677 3.959 2005     1212      1     3.001
## 5766  27844593205 4.199 2005     1212      1     2.629
```

```
## 5813  21644482697 4.818 2005     1212      1     3.336
## 5838  60950725967 3.497 2005     1212      1     2.528
## 5880  84992776219 3.720 2005     1212      1     2.751
## 5949  61249595569 3.788 2005     1208      2     2.819
## 5952  61449561847 3.582 2005     1208      2     2.624
## 5955  62449266342 3.723 2005     1208      2     2.754
## 5973  70449890230 3.464 2005     1212      2     2.506
## 6177  33751047195 4.104 2006     1212      1     2.623
## 6187  34347309336 3.704 2006     1212      1     2.698
## 6223  60950249825 3.675 2006     1212      1     2.669
## 6324  41949093695 4.329 2006     1212      2     3.333
## 6332  43249163482 3.642 2006     1212      2     2.636
## 6453  33747151905 3.863 2006     1212      1     2.857
## 6456  33747171138 4.064 2006     1212      1     3.058
## 6458  33747184520 4.298 2006     1212      1     3.292
## 6522  33646756864 4.376 2006     1212      1     2.895
## 6523  33646757510 4.886 2006     1212      1     3.405
## 6737  84996238317 3.515 2006     1212      1     2.519
## 6959  36949019141 4.370 2007     1212      1     2.896
## 6961  36949024013 4.168 2007     1212      1     2.775
## 6976  57749162980 3.486 2007     1212      1     2.501
## 7296  34548661972 4.013 2007     1212      1     2.620
## 7297  34548675370 3.736 2007     1212      1     2.740
## 7388  34249740121 4.357 2007     1212      1     2.964
## 7472  33947233870 4.662 2007     1212      1     3.269
## 7721  62449186372 3.857 2007     1208      2     2.872
## 7926  57749143176 3.986 2008     1212      1     3.087
## 8004  77955455590 4.118 2008     1212      1     3.219
## 8030  61049425894 3.837 2008     1212      2     2.949
## 8236  56049110254 4.473 2008     1212      1     3.096
## 8238  56049115364 4.257 2008     1212      1     2.680
## 8241  56049125485 3.986 2008     1212      1     3.098
## 8242  56049126209 3.899 2008     1212      1     2.603
## 8296  51249095120 4.093 2008     1212      1     2.727
## 8299  51249104866 4.118 2008     1212      1     2.822
## 8378  44349086981 4.014 2008     1212      1     3.115
## 8382  44349132774 4.418 2008     1212      1     3.129
## 8446  39049105256 3.837 2008     1212      1     2.629
## 8457  40949157971 4.508 2008     1212      1     3.300
## 8862  71649089787 3.994 2009     1212      1     2.775
## 8868  71649103465 4.736 2009     1212      1     3.517
## 9055  70349676067 3.812 2009     1212      1     2.990
## 9108  70049086788 4.125 2009     1212      1     2.625
## 9113  70049102705 4.243 2009     1212      1     3.421
## 9114  70049107140 4.499 2009     1212      1     3.677
## 9215  77649288637 3.450 2009     1212      2     2.628
## 9237  66949141813 4.445 2009     1212      1     3.226
## 9238  66849089428 4.804 2009     1212      1     3.507
## 9240  66849099520 4.125 2009     1212      1     3.303
## 9242  66849113683 4.048 2009     1212      1     2.751
```

```
## 9243  66849131260 3.671 2009     1212      1     2.540
## 9355  62349113564 3.966 2009     1212      1     3.144
## 9357  62349125186 4.075 2009     1212      1     2.652
## 9358  62349129158 4.286 2009     1212      1     3.155
## 9522  77954193832 3.633 2009     1212      1     2.811
## 9775  78049519127 3.637 2010     1212      1     2.844
## 9784  78649886543 3.440 2010     1212      1     2.647
## 9785  78649898747 3.916 2010     1212      1     2.512
## 9946  84864916678 4.301 2010     1212      1     3.100
## 10180 77954430414 4.196 2010     1212      1     3.392
## 10184 77954439455 3.492 2010     1212      1     2.688
## 10186 77954449929 3.880 2010     1212      1     3.076
## 10357 79956372974 3.668 2010     1202      4     2.864
## 10392 77950931872 4.699 2010     1212      1     3.895
## 10393 77950932977 4.168 2010     1212      1     2.975
## 10398 77950952104 4.992 2010     1212      1     4.188
## 10399 77950953030 4.350 2010     1212      1     3.157
## 10810 82955178180 3.559 2011     1212      1     2.685
## 10811 82955227609 3.938 2011     1212      1     2.743
## 10816 82955234080 5.243 2011     1212      1     4.358
## 11021 84881605265 3.644 2011     1212      1     2.759
## 11043 84894292374 4.121 2011     1212      1     2.839
## 11088 80155194058 3.906 2011     1202      5     3.032
## 11183 80052195465 3.870 2011     1212      1     2.507
## 11184 80052203482 3.870 2011     1212      1     2.518
## 11186 80052224481 4.944 2011     1212      1     4.059
## 11187 80052230221 3.938 2011     1212      1     3.053
## 11189 80052246166 4.002 2011     1212      1     3.117
## 11190 80052254269 4.781 2011     1212      1     3.341
## 11441 79957671344 4.063 2011     1212      1     3.178
## 11443 79957677461 4.063 2011     1212      1     3.189
## 11444 79957679982 4.002 2011     1212      1     3.128
## 11456 79958069144 3.644 2011     1212      1     2.759
## 11458 79958076416 3.724 2011     1212      1     2.850
## 11669 84870277898 5.639 2011     1212      1     4.146
## 11672 79952213862 4.177 2011     1212      1     3.292
## 11673 79952214878 4.746 2011     1212      1     3.861
## 11676 79952224166 4.944 2011     1212      1     4.059
## 11677 79952229517 4.231 2011     1212      1     2.871
## 11678 79952231734 5.267 2011     1212      1     4.382
## 11758 84856228825 6.890 2011     1212      1     5.615
## 11896 84555179123 4.538 2011     1208      2     3.653
## 12058 84872371974 6.016 2012     1212      1     4.481
## 12065 84872420393 4.778 2012     1212      1     3.443
## 12300 84868366032 3.394 2012     1208      2     2.537
## 12301 84868368393 3.394 2012     1208      2     2.537
## 12314 84868144975 4.380 2012     1212      1     3.523
## 12356 84894355733 3.397 2012     1212      1     2.540
## 12453 84865697715 3.939 2012     1208      2     3.082
## 12454 84865700895 4.859 2012     1208      2     4.013
```

```
## 12476 84865764944 3.965 2012     1212      1     3.119
## 12753 84862130809 4.976 2012     1212      1     3.518
## 12755 84862144323 4.286 2012     1212      1     3.429
## 12768 84860861091 3.804 2012     1211      2     2.550
## 12819 84869471464 4.071 2012     1211      2     3.214
## 12871 84860600228 4.186 2012     1212      1     3.329
## 12945 84859452716 3.707 2012     1212      1     2.850
## 13027 84857938890 4.380 2012     1212      1     3.534
## 13031 84857967717 5.540 2012     1212      1     4.286
## 13033 84857979842 4.380 2012     1212      1     3.534
## 13042 84858207835 4.913 2012     1212      1     4.067
## 13046 84858234457 3.397 2012     1212      1     2.540
## 13049 84858256380 6.195 2012     1212      1     5.349
## 13056 84863396437 4.847 2012     1212      1     3.304
## 13148 84855208708 3.559 2012     1212      1     2.702
## 13185 84866235322 3.397 2012     1212      1     2.551
## 13186 84866240255 4.913 2012     1212      1     4.056
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
UniqueAuthors +
##     Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##     k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.6813 -0.8035 -0.0329  0.6474  5.8135
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)         0.802628   0.044392   18.08  < 2e-16 ***
## FirstAuthorFemale1  0.077353   0.041449    1.87  0.06204 .
## LastAuthorFemale1  -0.087980   0.041454   -2.12  0.03383 *
## UniqueAuthors2      0.397393   0.035460   11.21  < 2e-16 ***
## UniqueAuthors3      0.477926   0.056193    8.51  < 2e-16 ***
## UniqueAuthors4      0.600818   0.093498    6.43  1.4e-10 ***
## UniqueAuthors5      0.608294   0.074490    8.17  3.5e-16 ***
## Year1997            0.008005   0.064004    0.13  0.90047
## Year1998            0.175132   0.058866    2.98  0.00294 **
## Year1999            0.043559   0.061905    0.70  0.48167
## Year2000            0.128576   0.057759    2.23  0.02603 *
## Year2001            0.063945   0.056853    1.12  0.26072
## Year2002            0.221865   0.061093    3.63  0.00028 ***
## Year2003            0.247367   0.061253    4.04  5.4e-05 ***
## Year2004            0.197234   0.061434    3.21  0.00133 **
## Year2005            0.166258   0.059271    2.81  0.00504 **
## Year2006            0.203527   0.056198    3.62  0.00029 ***
## Year2007            0.193053   0.053976    3.58  0.00035 ***
## Year2008            0.096232   0.053736    1.79  0.07335 .
```

```
## Year2009            0.019176   0.053406    0.36  0.71956
## Year2010            0.000875   0.053145    0.02  0.98687
## Year2011            0.082476   0.051881    1.59  0.11193
## Year2012            0.054247   0.052954    1.02  0.30566
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.898
## Multiple R-squared:  0.0349, Adjusted R-squared:  0.033
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
##  11 observations c(132,586,977,2879,2912,9076,9779,9848,10096,10873,10890)
##   are outliers with |weight| <= 1.7e-08 ( < 9e-06);
##  881 weights are ~= 1. The remaining 10255 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0001  0.8910  0.9280  0.9040  0.9830  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          8.97e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##                     GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 4.395  1           2.096
## LastAuthorFemale  4.390  1           2.095
## Year              1.008 16           1.000
```




```
## [1] "List of  206 outliers with residuals above  2.5"
##          ScopusId  NLCS Year OneField Fields residuals
## 141    0000339785 4.220 1996     1212      1     3.306
## 144    0001811907 4.077 1996     1212      1     3.241
## 145    0002009675 3.936 1996     1212      1     3.100
## 146    0002288148 4.141 1996     1212      1     3.305
## 147    0002291743 4.518 1996     1212      1     3.647
## 151    0012467113 4.997 1996     1212      1     4.126
## 152    0013066072 3.885 1996     1212      1     3.049
## 153    0039860262 3.616 1996     1212      1     2.745
## 155    0242607951 3.961 1996     1212      1     3.125
## 160   21344473380 4.412 1996     1212      1     3.576
## 164   33745249484 3.390 1996     1212      1     2.554
## 282    0001565086 3.590 1996     1212      2     2.754
## 287    0003143117 3.579 1996     1211      2     2.743
## 309    0042684693 4.071 1996     1211      2     3.235
## 685    0001431225 6.397 1997     1212      1     5.590
## 1106   0000315626 3.948 1998     1212      1     2.935
## 1107   0000473275 4.202 1998     1212      1     3.189
## 1112   0002098619 3.653 1998     1212      1     2.606
## 1124   0009788167 3.653 1998     1212      1     2.640
## 1134   0039408604 3.528 1998     1212      1     2.515
## 1150  33748132596 6.001 1998     1212      1     5.031
## 1153  33748194686 4.108 1998     1212      1     3.095
## 1670   0001718525 4.602 1999     1212      1     3.756
## 1679   0033238699 3.803 1999     1212      1     2.879
## 1682   0041359166 4.417 1999     1212      1     3.528
```

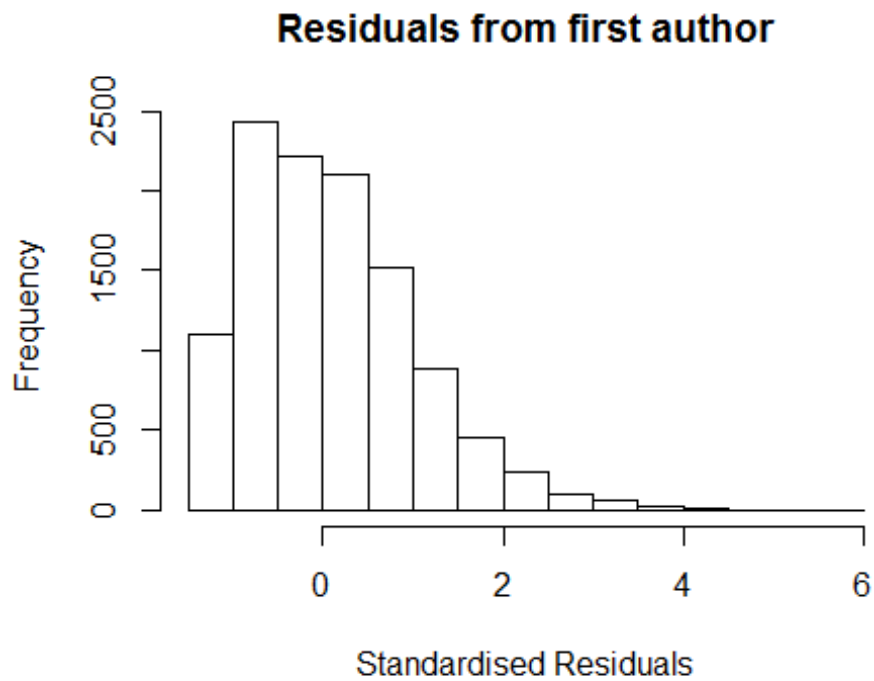
```
## 1701  60950059093 3.550 1999     1212      1     2.626
## 1803  33749819408 3.636 1999     1212      2     2.747
## 2114  33748347287 3.974 2000     1212      1     3.025
## 2181   0007337668 3.974 2000     1212      1     3.068
## 2191   0038437363 4.452 2000     1212      1     3.503
## 2194   0242564570 3.458 2000     1212      1     2.509
## 2196   1642316438 4.462 2000     1212      1     3.556
## 2610  61449201817 3.518 2001     1212      1     2.620
## 2784  33748072453 3.503 2001     1212      1     2.605
## 2827   0001415459 3.882 2001     1212      1     3.027
## 2829   0011182447 3.891 2001     1212      1     2.993
## 2830   0038694847 4.192 2001     1212      1     3.294
## 3442   0036747023 5.208 2002     1212      1     4.084
## 3443   0036748060 5.399 2002     1212      1     4.318
## 3444   0036748067 4.725 2002     1212      1     3.644
## 3445   0036749032 3.770 2002     1212      1     2.723
## 3481   0036623355 6.838 2002     1212      1     5.791
## 3516   0036521817 3.625 2002     1212      1     2.621
## 3517   0036521906 4.507 2002     1212      1     3.426
## 3545  33645260213 3.625 2002     1212      1     2.544
## 3564  60949697784 3.957 2002     1212      1     2.910
## 3751  61249583521 3.963 2002     1208      2     2.916
## 3761  62449174089 3.672 2002     1208      2     2.625
## 3763  62449296550 4.063 2002     1208      2     2.982
## 3878   0036526749 3.670 2002     1202      6     2.589
## 3880   1442335800 3.694 2002     1202      6     2.647
## 3914  60950296097 4.037 2003     1212      1     2.949
## 3964  70449986080 3.687 2003     1212      1     2.599
## 4112   0038489174 3.779 2003     1212      1     2.657
## 4235  61249404750 3.718 2003     1212      1     2.630
## 4356  11244330107 4.637 2003     1211      2     3.515
## 4361  18444365755 4.381 2003     1211      2     3.259
## 4364  29744456134 5.218 2003     1208      2     4.130
## 4383  34547980224 3.959 2003     1208      2     2.871
## 4412  61249716500 4.045 2003     1208      2     2.957
## 4657  37949018607 3.540 2004     1212      1     2.502
## 4672  60949844649 3.775 2004     1212      1     2.702
## 4705  61949157081 3.732 2004     1212      1     2.694
## 5033  84992828474 5.075 2004     1212      1     4.037
## 5076  85040393605 4.005 2004     1212      1     2.967
## 5316  36749058235 3.686 2005     1212      1     2.671
## 5337  60949946845 3.720 2005     1212      1     2.670
## 5538  62749191854 3.744 2005     1202      3     2.694
## 5592  61149560064 3.651 2005     1212      1     2.636
## 5618  33646497937 3.765 2005     1211      2     2.750
## 5620  33646527844 3.946 2005     1211      2     2.931
## 5669  27844441376 3.497 2005     1212      1     2.525
## 5672  27844473360 3.904 2005     1212      1     2.889
## 5673  27844485164 3.959 2005     1212      1     2.944
## 5674  27844500221 4.768 2005     1212      1     3.753
```

```
## 5675  27844592291 3.959 2005     1212      1     2.944
## 5765  27844515677 3.959 2005     1212      1     2.909
## 5766  27844593205 4.199 2005     1212      1     3.184
## 5813  21644482697 4.818 2005     1212      1     3.846
## 5817  27844577835 3.904 2005     1212      1     2.811
## 5880  84992776219 3.720 2005     1212      1     2.705
## 5949  61249595569 3.788 2005     1208      2     2.773
## 5950  61449306806 3.654 2005     1208      2     2.561
## 5952  61449561847 3.582 2005     1208      2     2.532
## 5955  62449266342 3.723 2005     1208      2     2.708
## 6177  33751047195 4.104 2006     1212      1     2.973
## 6187  34347309336 3.704 2006     1212      1     2.650
## 6223  60950249825 3.675 2006     1212      1     2.621
## 6324  41949093695 4.329 2006     1212      2     3.241
## 6332  43249163482 3.642 2006     1212      2     2.588
## 6453  33747151905 3.863 2006     1212      1     2.809
## 6456  33747171138 4.064 2006     1212      1     3.010
## 6458  33747184520 4.298 2006     1212      1     3.244
## 6522  33646756864 4.376 2006     1212      1     3.245
## 6523  33646757510 4.886 2006     1212      1     3.755
## 6957  36949003441 3.577 2007     1212      1     2.577
## 6959  36949019141 4.370 2007     1212      1     3.327
## 6961  36949024013 4.168 2007     1212      1     3.125
## 7296  34548661972 4.013 2007     1212      1     2.970
## 7297  34548675370 3.736 2007     1212      1     2.693
## 7388  34249740121 4.357 2007     1212      1     3.314
## 7472  33947233870 4.662 2007     1212      1     3.619
## 7478  33947270473 3.547 2007     1212      1     2.547
## 7721  62449186372 3.857 2007     1208      2     2.779
## 7926  57749143176 3.986 2008     1212      1     3.037
## 8004  77955455590 4.118 2008     1212      1     3.169
## 8030  61049425894 3.837 2008     1212      2     2.853
## 8236  56049110254 4.473 2008     1212      1     3.524
## 8238  56049115364 4.257 2008     1212      1     3.230
## 8239  56049118467 3.771 2008     1212      1     2.744
## 8241  56049125485 3.986 2008     1212      1     3.002
## 8242  56049126209 3.899 2008     1212      1     2.950
## 8294  51249088778 3.545 2008     1212      1     2.561
## 8296  51249095120 4.093 2008     1212      1     3.109
## 8299  51249104866 4.118 2008     1212      1     3.169
## 8378  44349086981 4.014 2008     1212      1     3.065
## 8379  44349090359 3.664 2008     1212      1     2.758
## 8382  44349132774 4.418 2008     1212      1     3.512
## 8446  39049105256 3.837 2008     1212      1     2.931
## 8457  40949157971 4.508 2008     1212      1     3.602
## 8862  71649089787 3.994 2009     1212      1     3.106
## 8863  71649091181 3.465 2009     1212      1     2.577
## 8868  71649103465 4.736 2009     1212      1     3.848
## 9055  70349676067 3.812 2009     1212      1     2.924
## 9108  70049086788 4.125 2009     1212      1     3.160
```

```
## 9113  70049102705 4.243 2009     1212      1     3.355
## 9114  70049107140 4.499 2009     1212      1     3.611
## 9131  74949109006 3.465 2009     1212      1     2.620
## 9215  77649288637 3.450 2009     1212      2     2.562
## 9237  66949141813 4.445 2009     1212      1     3.557
## 9238  66849089428 4.804 2009     1212      1     3.839
## 9240  66849099520 4.125 2009     1212      1     3.237
## 9242  66849113683 4.048 2009     1212      1     3.083
## 9243  66849131260 3.671 2009     1212      1     2.826
## 9352  62349085774 3.671 2009     1212      1     2.783
## 9355  62349113564 3.966 2009     1212      1     3.078
## 9357  62349125186 4.075 2009     1212      1     3.187
## 9358  62349129158 4.286 2009     1212      1     3.441
## 9389  61449262127 3.418 2009     1212      1     2.530
## 9522  77954193832 3.633 2009     1212      1     2.745
## 9775  78049519127 3.637 2010     1212      1     2.726
## 9784  78649886543 3.440 2010     1212      1     2.529
## 9785  78649898747 3.916 2010     1212      1     3.040
## 9946  84864916678 4.301 2010     1212      1     3.425
## 10035 77956406486 3.385 2010     1212      1     2.509
## 10180 77954430414 4.196 2010     1212      1     3.320
## 10183 77954438289 3.681 2010     1212      1     2.805
## 10184 77954439455 3.492 2010     1212      1     2.616
## 10186 77954449929 3.880 2010     1212      1     3.004
## 10187 77954451120 3.591 2010     1212      1     2.715
## 10357 79956372974 3.668 2010     1202      4     2.792
## 10382 77950251483 3.492 2010     1212      1     2.616
## 10392 77950931872 4.699 2010     1212      1     3.823
## 10393 77950932977 4.168 2010     1212      1     3.335
## 10398 77950952104 4.992 2010     1212      1     4.116
## 10399 77950953030 4.350 2010     1212      1     3.517
## 10810 82955178180 3.559 2011     1212      1     2.570
## 10811 82955227609 3.938 2011     1212      1     3.026
## 10815 82955234077 3.644 2011     1212      1     2.655
## 10816 82955234080 5.243 2011     1212      1     4.289
## 11021 84881605265 3.644 2011     1212      1     2.690
## 11043 84894292374 4.121 2011     1212      1     3.167
## 11088 80155194058 3.906 2011     1202      5     2.917
## 11183 80052195465 3.870 2011     1212      1     2.916
## 11184 80052203482 3.870 2011     1212      1     2.881
## 11186 80052224481 4.944 2011     1212      1     3.990
## 11187 80052230221 3.938 2011     1212      1     2.984
## 11189 80052246166 4.002 2011     1212      1     3.048
## 11190 80052254269 4.781 2011     1212      1     3.749
## 11428 84893744785 3.559 2011     1212      1     2.647
## 11441 79957671344 4.063 2011     1212      1     3.109
## 11443 79957677461 4.063 2011     1212      1     3.074
## 11444 79957679982 4.002 2011     1212      1     3.013
## 11456 79958069144 3.644 2011     1212      1     2.690
## 11458 79958076416 3.724 2011     1212      1     2.735
```

```
## 11669 84870277898 5.639 2011     1212      1     4.685
## 11672 79952213862 4.177 2011     1212      1     3.223
## 11673 79952214878 4.746 2011     1212      1     3.792
## 11676 79952224166 4.944 2011     1212      1     3.990
## 11677 79952229517 4.231 2011     1212      1     3.199
## 11678 79952231734 5.267 2011     1212      1     4.313
## 11758 84856228825 6.890 2011     1212      1     5.978
## 11896 84555179123 4.538 2011     1208      2     3.584
## 12058 84872371974 6.016 2012     1212      1     5.008
## 12065 84872420393 4.778 2012     1212      1     3.847
## 12314 84868144975 4.380 2012     1212      1     3.449
## 12453 84865697715 3.939 2012     1208      2     3.008
## 12454 84865700895 4.859 2012     1208      2     3.894
## 12469 84865721280 3.707 2012     1212      1     2.699
## 12476 84865764944 3.965 2012     1212      1     3.000
## 12743 84861553184 3.559 2012     1212      1     2.551
## 12753 84862130809 4.976 2012     1212      1     4.045
## 12754 84862133446 3.559 2012     1212      1     2.628
## 12755 84862144323 4.286 2012     1212      1     3.355
## 12768 84860861091 3.804 2012     1211      2     2.873
## 12819 84869471464 4.071 2012     1211      2     3.140
## 12871 84860600228 4.186 2012     1212      1     3.255
## 12945 84859452716 3.707 2012     1212      1     2.776
## 13027 84857938890 4.380 2012     1212      1     3.415
## 13031 84857967717 5.540 2012     1212      1     4.609
## 13033 84857979842 4.380 2012     1212      1     3.415
## 13042 84858207835 4.913 2012     1212      1     3.948
## 13049 84858256380 6.195 2012     1212      1     5.230
## 13056 84863396437 4.847 2012     1212      1     3.839
## 13148 84855208708 3.559 2012     1212      1     2.628
## 13186 84866240255 4.913 2012     1212      1     3.982
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.1652 -0.8758 -0.0435  0.6588  5.9784
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)          0.8362     0.0450   18.58  < 2e-16 ***
## FirstAuthorFemale1   0.0777     0.0418    1.86  0.06337 .
## LastAuthorFemale1   -0.0429     0.0416   -1.03  0.30275
## Year1997             0.0141     0.0648    0.22  0.82812
## Year1998             0.1765     0.0594    2.97  0.00299 **
## Year1999             0.0528     0.0621    0.85  0.39489
## Year2000             0.1130     0.0586    1.93  0.05387 .
```

```
## Year2001             0.0615     0.0579    1.06  0.28744
## Year2002             0.2105     0.0610    3.45  0.00057 ***
## Year2003             0.2513     0.0616    4.08  4.6e-05 ***
## Year2004             0.2019     0.0612    3.30  0.00098 ***
## Year2005             0.1787     0.0596    3.00  0.00272 **
## Year2006             0.2175     0.0569    3.82  0.00013 ***
## Year2007             0.2066     0.0545    3.79  0.00015 ***
## Year2008             0.1126     0.0546    2.06  0.03908 *
## Year2009             0.0515     0.0545    0.95  0.34467
## Year2010             0.0396     0.0539    0.73  0.46331
## Year2011             0.1183     0.0528    2.24  0.02522 *
## Year2012             0.0943     0.0538    1.75  0.07977 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.929
## Multiple R-squared:  0.00678,    Adjusted R-squared:  0.00517
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  8 observations c(586,977,2912,9770,9848,10096,10873,10890)
##   are outliers with |weight| = 0 ( < 9e-06);
##  922 weights are ~= 1. The remaining 10217 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0002  0.8890  0.9250  0.9060  0.9840  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          8.97e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##                     GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.005  1           1.003
## Year              1.005 16           1.000
```




```
## [1] "List of  206 outliers with residuals above  2.5"
##          ScopusId  NLCS Year OneField Fields residuals
## 141    0000339785 4.220 1996     1212      1     3.306
## 144    0001811907 4.077 1996     1212      1     3.241
## 145    0002009675 3.936 1996     1212      1     3.100
## 146    0002288148 4.141 1996     1212      1     3.305
## 147    0002291743 4.518 1996     1212      1     3.647
## 151    0012467113 4.997 1996     1212      1     4.126
## 152    0013066072 3.885 1996     1212      1     3.049
## 153    0039860262 3.616 1996     1212      1     2.745
## 155    0242607951 3.961 1996     1212      1     3.125
## 160   21344473380 4.412 1996     1212      1     3.576
## 164   33745249484 3.390 1996     1212      1     2.554
## 282    0001565086 3.590 1996     1212      2     2.754
## 287    0003143117 3.579 1996     1211      2     2.743
## 309    0042684693 4.071 1996     1211      2     3.235
## 685    0001431225 6.397 1997     1212      1     5.590
## 1106   0000315626 3.948 1998     1212      1     2.935
## 1107   0000473275 4.202 1998     1212      1     3.189
## 1112   0002098619 3.653 1998     1212      1     2.606
## 1124   0009788167 3.653 1998     1212      1     2.640
## 1134   0039408604 3.528 1998     1212      1     2.515
## 1150  33748132596 6.001 1998     1212      1     5.031
## 1153  33748194686 4.108 1998     1212      1     3.095
## 1670   0001718525 4.602 1999     1212      1     3.756
## 1679   0033238699 3.803 1999     1212      1     2.879
## 1682   0041359166 4.417 1999     1212      1     3.528
```

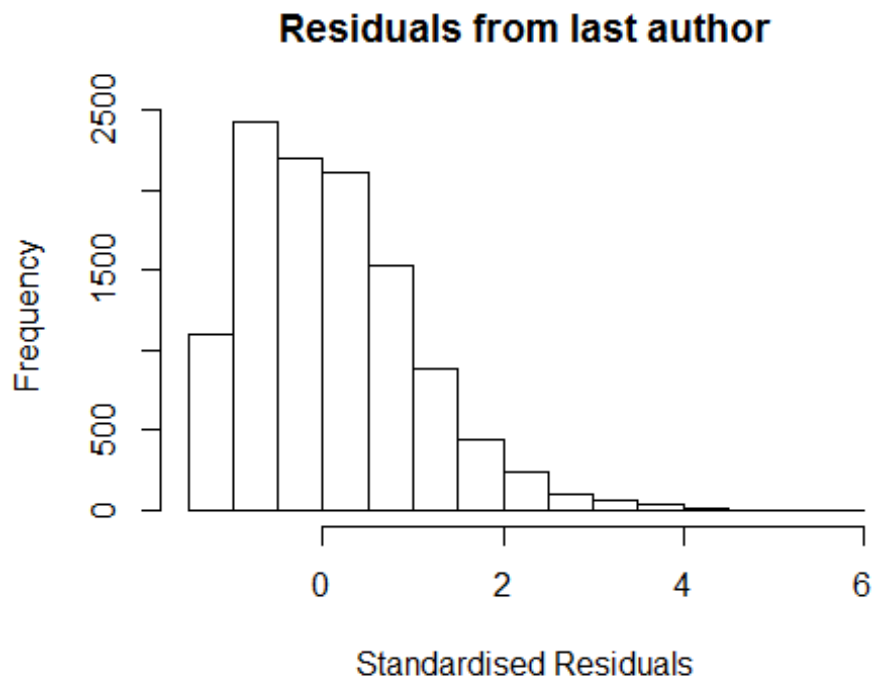
```
## 1701  60950059093 3.550 1999     1212      1     2.626
## 1803  33749819408 3.636 1999     1212      2     2.747
## 2114  33748347287 3.974 2000     1212      1     3.025
## 2181   0007337668 3.974 2000     1212      1     3.068
## 2191   0038437363 4.452 2000     1212      1     3.503
## 2194   0242564570 3.458 2000     1212      1     2.509
## 2196   1642316438 4.462 2000     1212      1     3.556
## 2610  61449201817 3.518 2001     1212      1     2.620
## 2784  33748072453 3.503 2001     1212      1     2.605
## 2827   0001415459 3.882 2001     1212      1     3.027
## 2829   0011182447 3.891 2001     1212      1     2.993
## 2830   0038694847 4.192 2001     1212      1     3.294
## 3442   0036747023 5.208 2002     1212      1     4.084
## 3443   0036748060 5.399 2002     1212      1     4.318
## 3444   0036748067 4.725 2002     1212      1     3.644
## 3445   0036749032 3.770 2002     1212      1     2.723
## 3481   0036623355 6.838 2002     1212      1     5.791
## 3516   0036521817 3.625 2002     1212      1     2.621
## 3517   0036521906 4.507 2002     1212      1     3.426
## 3545  33645260213 3.625 2002     1212      1     2.544
## 3564  60949697784 3.957 2002     1212      1     2.910
## 3751  61249583521 3.963 2002     1208      2     2.916
## 3761  62449174089 3.672 2002     1208      2     2.625
## 3763  62449296550 4.063 2002     1208      2     2.982
## 3878   0036526749 3.670 2002     1202      6     2.589
## 3880   1442335800 3.694 2002     1202      6     2.647
## 3914  60950296097 4.037 2003     1212      1     2.949
## 3964  70449986080 3.687 2003     1212      1     2.599
## 4112   0038489174 3.779 2003     1212      1     2.657
## 4235  61249404750 3.718 2003     1212      1     2.630
## 4356  11244330107 4.637 2003     1211      2     3.515
## 4361  18444365755 4.381 2003     1211      2     3.259
## 4364  29744456134 5.218 2003     1208      2     4.130
## 4383  34547980224 3.959 2003     1208      2     2.871
## 4412  61249716500 4.045 2003     1208      2     2.957
## 4657  37949018607 3.540 2004     1212      1     2.502
## 4672  60949844649 3.775 2004     1212      1     2.702
## 4705  61949157081 3.732 2004     1212      1     2.694
## 5033  84992828474 5.075 2004     1212      1     4.037
## 5076  85040393605 4.005 2004     1212      1     2.967
## 5316  36749058235 3.686 2005     1212      1     2.671
## 5337  60949946845 3.720 2005     1212      1     2.670
## 5538  62749191854 3.744 2005     1202      3     2.694
## 5592  61149560064 3.651 2005     1212      1     2.636
## 5618  33646497937 3.765 2005     1211      2     2.750
## 5620  33646527844 3.946 2005     1211      2     2.931
## 5669  27844441376 3.497 2005     1212      1     2.525
## 5672  27844473360 3.904 2005     1212      1     2.889
## 5673  27844485164 3.959 2005     1212      1     2.944
## 5674  27844500221 4.768 2005     1212      1     3.753
```

```
## 5675  27844592291 3.959 2005     1212      1     2.944
## 5765  27844515677 3.959 2005     1212      1     2.909
## 5766  27844593205 4.199 2005     1212      1     3.184
## 5813  21644482697 4.818 2005     1212      1     3.846
## 5817  27844577835 3.904 2005     1212      1     2.811
## 5880  84992776219 3.720 2005     1212      1     2.705
## 5949  61249595569 3.788 2005     1208      2     2.773
## 5950  61449306806 3.654 2005     1208      2     2.561
## 5952  61449561847 3.582 2005     1208      2     2.532
## 5955  62449266342 3.723 2005     1208      2     2.708
## 6177  33751047195 4.104 2006     1212      1     2.973
## 6187  34347309336 3.704 2006     1212      1     2.650
## 6223  60950249825 3.675 2006     1212      1     2.621
## 6324  41949093695 4.329 2006     1212      2     3.241
## 6332  43249163482 3.642 2006     1212      2     2.588
## 6453  33747151905 3.863 2006     1212      1     2.809
## 6456  33747171138 4.064 2006     1212      1     3.010
## 6458  33747184520 4.298 2006     1212      1     3.244
## 6522  33646756864 4.376 2006     1212      1     3.245
## 6523  33646757510 4.886 2006     1212      1     3.755
## 6957  36949003441 3.577 2007     1212      1     2.577
## 6959  36949019141 4.370 2007     1212      1     3.327
## 6961  36949024013 4.168 2007     1212      1     3.125
## 7296  34548661972 4.013 2007     1212      1     2.970
## 7297  34548675370 3.736 2007     1212      1     2.693
## 7388  34249740121 4.357 2007     1212      1     3.314
## 7472  33947233870 4.662 2007     1212      1     3.619
## 7478  33947270473 3.547 2007     1212      1     2.547
## 7721  62449186372 3.857 2007     1208      2     2.779
## 7926  57749143176 3.986 2008     1212      1     3.037
## 8004  77955455590 4.118 2008     1212      1     3.169
## 8030  61049425894 3.837 2008     1212      2     2.853
## 8236  56049110254 4.473 2008     1212      1     3.524
## 8238  56049115364 4.257 2008     1212      1     3.230
## 8239  56049118467 3.771 2008     1212      1     2.744
## 8241  56049125485 3.986 2008     1212      1     3.002
## 8242  56049126209 3.899 2008     1212      1     2.950
## 8294  51249088778 3.545 2008     1212      1     2.561
## 8296  51249095120 4.093 2008     1212      1     3.109
## 8299  51249104866 4.118 2008     1212      1     3.169
## 8378  44349086981 4.014 2008     1212      1     3.065
## 8379  44349090359 3.664 2008     1212      1     2.758
## 8382  44349132774 4.418 2008     1212      1     3.512
## 8446  39049105256 3.837 2008     1212      1     2.931
## 8457  40949157971 4.508 2008     1212      1     3.602
## 8862  71649089787 3.994 2009     1212      1     3.106
## 8863  71649091181 3.465 2009     1212      1     2.577
## 8868  71649103465 4.736 2009     1212      1     3.848
## 9055  70349676067 3.812 2009     1212      1     2.924
## 9108  70049086788 4.125 2009     1212      1     3.160
```

```
## 9113  70049102705 4.243 2009     1212      1     3.355
## 9114  70049107140 4.499 2009     1212      1     3.611
## 9131  74949109006 3.465 2009     1212      1     2.620
## 9215  77649288637 3.450 2009     1212      2     2.562
## 9237  66949141813 4.445 2009     1212      1     3.557
## 9238  66849089428 4.804 2009     1212      1     3.839
## 9240  66849099520 4.125 2009     1212      1     3.237
## 9242  66849113683 4.048 2009     1212      1     3.083
## 9243  66849131260 3.671 2009     1212      1     2.826
## 9352  62349085774 3.671 2009     1212      1     2.783
## 9355  62349113564 3.966 2009     1212      1     3.078
## 9357  62349125186 4.075 2009     1212      1     3.187
## 9358  62349129158 4.286 2009     1212      1     3.441
## 9389  61449262127 3.418 2009     1212      1     2.530
## 9522  77954193832 3.633 2009     1212      1     2.745
## 9775  78049519127 3.637 2010     1212      1     2.726
## 9784  78649886543 3.440 2010     1212      1     2.529
## 9785  78649898747 3.916 2010     1212      1     3.040
## 9946  84864916678 4.301 2010     1212      1     3.425
## 10035 77956406486 3.385 2010     1212      1     2.509
## 10180 77954430414 4.196 2010     1212      1     3.320
## 10183 77954438289 3.681 2010     1212      1     2.805
## 10184 77954439455 3.492 2010     1212      1     2.616
## 10186 77954449929 3.880 2010     1212      1     3.004
## 10187 77954451120 3.591 2010     1212      1     2.715
## 10357 79956372974 3.668 2010     1202      4     2.792
## 10382 77950251483 3.492 2010     1212      1     2.616
## 10392 77950931872 4.699 2010     1212      1     3.823
## 10393 77950932977 4.168 2010     1212      1     3.335
## 10398 77950952104 4.992 2010     1212      1     4.116
## 10399 77950953030 4.350 2010     1212      1     3.517
## 10810 82955178180 3.559 2011     1212      1     2.570
## 10811 82955227609 3.938 2011     1212      1     3.026
## 10815 82955234077 3.644 2011     1212      1     2.655
## 10816 82955234080 5.243 2011     1212      1     4.289
## 11021 84881605265 3.644 2011     1212      1     2.690
## 11043 84894292374 4.121 2011     1212      1     3.167
## 11088 80155194058 3.906 2011     1202      5     2.917
## 11183 80052195465 3.870 2011     1212      1     2.916
## 11184 80052203482 3.870 2011     1212      1     2.881
## 11186 80052224481 4.944 2011     1212      1     3.990
## 11187 80052230221 3.938 2011     1212      1     2.984
## 11189 80052246166 4.002 2011     1212      1     3.048
## 11190 80052254269 4.781 2011     1212      1     3.749
## 11428 84893744785 3.559 2011     1212      1     2.647
## 11441 79957671344 4.063 2011     1212      1     3.109
## 11443 79957677461 4.063 2011     1212      1     3.074
## 11444 79957679982 4.002 2011     1212      1     3.013
## 11456 79958069144 3.644 2011     1212      1     2.690
## 11458 79958076416 3.724 2011     1212      1     2.735
```

```
## 11669 84870277898 5.639 2011     1212      1     4.685
## 11672 79952213862 4.177 2011     1212      1     3.223
## 11673 79952214878 4.746 2011     1212      1     3.792
## 11676 79952224166 4.944 2011     1212      1     3.990
## 11677 79952229517 4.231 2011     1212      1     3.199
## 11678 79952231734 5.267 2011     1212      1     4.313
## 11758 84856228825 6.890 2011     1212      1     5.978
## 11896 84555179123 4.538 2011     1208      2     3.584
## 12058 84872371974 6.016 2012     1212      1     5.008
## 12065 84872420393 4.778 2012     1212      1     3.847
## 12314 84868144975 4.380 2012     1212      1     3.449
## 12453 84865697715 3.939 2012     1208      2     3.008
## 12454 84865700895 4.859 2012     1208      2     3.894
## 12469 84865721280 3.707 2012     1212      1     2.699
## 12476 84865764944 3.965 2012     1212      1     3.000
## 12743 84861553184 3.559 2012     1212      1     2.551
## 12753 84862130809 4.976 2012     1212      1     4.045
## 12754 84862133446 3.559 2012     1212      1     2.628
## 12755 84862144323 4.286 2012     1212      1     3.355
## 12768 84860861091 3.804 2012     1211      2     2.873
## 12819 84869471464 4.071 2012     1211      2     3.140
## 12871 84860600228 4.186 2012     1212      1     3.255
## 12945 84859452716 3.707 2012     1212      1     2.776
## 13027 84857938890 4.380 2012     1212      1     3.415
## 13031 84857967717 5.540 2012     1212      1     4.609
## 13033 84857979842 4.380 2012     1212      1     3.415
## 13042 84858207835 4.913 2012     1212      1     3.948
## 13049 84858256380 6.195 2012     1212      1     5.230
## 13056 84863396437 4.847 2012     1212      1     3.839
## 13148 84855208708 3.559 2012     1212      1     2.628
## 13186 84866240255 4.913 2012     1212      1     3.982
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.1263 -0.8742 -0.0415  0.6600  5.9375
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)          0.8343     0.0449   18.56  < 2e-16 ***
## FirstAuthorFemale1   0.0402     0.0200    2.01  0.04466 *
## Year1997             0.0143     0.0648    0.22  0.82515
## Year1998             0.1767     0.0595    2.97  0.00297 **
## Year1999             0.0536     0.0621    0.86  0.38801
## Year2000             0.1139     0.0586    1.94  0.05183 .
## Year2001             0.0617     0.0578    1.07  0.28652
```

```
## Year2002             0.2111     0.0610    3.46  0.00054 ***
## Year2003             0.2517     0.0616    4.08  4.4e-05 ***
## Year2004             0.2025     0.0612    3.31  0.00094 ***
## Year2005             0.1788     0.0596    3.00  0.00271 **
## Year2006             0.2181     0.0568    3.84  0.00012 ***
## Year2007             0.2064     0.0545    3.79  0.00015 ***
## Year2008             0.1125     0.0546    2.06  0.03926 *
## Year2009             0.0515     0.0544    0.95  0.34437
## Year2010             0.0399     0.0539    0.74  0.45947
## Year2011             0.1182     0.0528    2.24  0.02527 *
## Year2012             0.0944     0.0538    1.75  0.07942 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.929
## Multiple R-squared:  0.00666,    Adjusted R-squared:  0.00514
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  8 observations c(586,977,2912,9770,9848,10096,10873,10890)
##   are outliers with |weight| = 0 ( < 9e-06);
##  928 weights are ~= 1. The remaining 10211 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0003  0.8900  0.9250  0.9060  0.9840  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          8.97e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##                   GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.004  1           1.002
## Year             1.004 16           1.000
```




```
## [1] "List of  206 outliers with residuals above  2.5"
##          ScopusId  NLCS Year OneField Fields residuals
## 141    0000339785 4.220 1996     1212      1     3.306
## 144    0001811907 4.077 1996     1212      1     3.241
## 145    0002009675 3.936 1996     1212      1     3.100
## 146    0002288148 4.141 1996     1212      1     3.305
## 147    0002291743 4.518 1996     1212      1     3.647
## 151    0012467113 4.997 1996     1212      1     4.126
## 152    0013066072 3.885 1996     1212      1     3.049
## 153    0039860262 3.616 1996     1212      1     2.745
## 155    0242607951 3.961 1996     1212      1     3.125
## 160   21344473380 4.412 1996     1212      1     3.576
## 164   33745249484 3.390 1996     1212      1     2.554
## 282    0001565086 3.590 1996     1212      2     2.754
## 287    0003143117 3.579 1996     1211      2     2.743
## 309    0042684693 4.071 1996     1211      2     3.235
## 685    0001431225 6.397 1997     1212      1     5.590
## 1106   0000315626 3.948 1998     1212      1     2.935
## 1107   0000473275 4.202 1998     1212      1     3.189
## 1112   0002098619 3.653 1998     1212      1     2.606
## 1124   0009788167 3.653 1998     1212      1     2.640
## 1134   0039408604 3.528 1998     1212      1     2.515
## 1150  33748132596 6.001 1998     1212      1     5.031
## 1153  33748194686 4.108 1998     1212      1     3.095
## 1670   0001718525 4.602 1999     1212      1     3.756
## 1679   0033238699 3.803 1999     1212      1     2.879
## 1682   0041359166 4.417 1999     1212      1     3.528
```

```
## 1701  60950059093 3.550 1999     1212      1     2.626
## 1803  33749819408 3.636 1999     1212      2     2.747
## 2114  33748347287 3.974 2000     1212      1     3.025
## 2181   0007337668 3.974 2000     1212      1     3.068
## 2191   0038437363 4.452 2000     1212      1     3.503
## 2194   0242564570 3.458 2000     1212      1     2.509
## 2196   1642316438 4.462 2000     1212      1     3.556
## 2610  61449201817 3.518 2001     1212      1     2.620
## 2784  33748072453 3.503 2001     1212      1     2.605
## 2827   0001415459 3.882 2001     1212      1     3.027
## 2829   0011182447 3.891 2001     1212      1     2.993
## 2830   0038694847 4.192 2001     1212      1     3.294
## 3442   0036747023 5.208 2002     1212      1     4.084
## 3443   0036748060 5.399 2002     1212      1     4.318
## 3444   0036748067 4.725 2002     1212      1     3.644
## 3445   0036749032 3.770 2002     1212      1     2.723
## 3481   0036623355 6.838 2002     1212      1     5.791
## 3516   0036521817 3.625 2002     1212      1     2.621
## 3517   0036521906 4.507 2002     1212      1     3.426
## 3545  33645260213 3.625 2002     1212      1     2.544
## 3564  60949697784 3.957 2002     1212      1     2.910
## 3751  61249583521 3.963 2002     1208      2     2.916
## 3761  62449174089 3.672 2002     1208      2     2.625
## 3763  62449296550 4.063 2002     1208      2     2.982
## 3878   0036526749 3.670 2002     1202      6     2.589
## 3880   1442335800 3.694 2002     1202      6     2.647
## 3914  60950296097 4.037 2003     1212      1     2.949
## 3964  70449986080 3.687 2003     1212      1     2.599
## 4112   0038489174 3.779 2003     1212      1     2.657
## 4235  61249404750 3.718 2003     1212      1     2.630
## 4356  11244330107 4.637 2003     1211      2     3.515
## 4361  18444365755 4.381 2003     1211      2     3.259
## 4364  29744456134 5.218 2003     1208      2     4.130
## 4383  34547980224 3.959 2003     1208      2     2.871
## 4412  61249716500 4.045 2003     1208      2     2.957
## 4657  37949018607 3.540 2004     1212      1     2.502
## 4672  60949844649 3.775 2004     1212      1     2.702
## 4705  61949157081 3.732 2004     1212      1     2.694
## 5033  84992828474 5.075 2004     1212      1     4.037
## 5076  85040393605 4.005 2004     1212      1     2.967
## 5316  36749058235 3.686 2005     1212      1     2.671
## 5337  60949946845 3.720 2005     1212      1     2.670
## 5538  62749191854 3.744 2005     1202      3     2.694
## 5592  61149560064 3.651 2005     1212      1     2.636
## 5618  33646497937 3.765 2005     1211      2     2.750
## 5620  33646527844 3.946 2005     1211      2     2.931
## 5669  27844441376 3.497 2005     1212      1     2.525
## 5672  27844473360 3.904 2005     1212      1     2.889
## 5673  27844485164 3.959 2005     1212      1     2.944
## 5674  27844500221 4.768 2005     1212      1     3.753
```

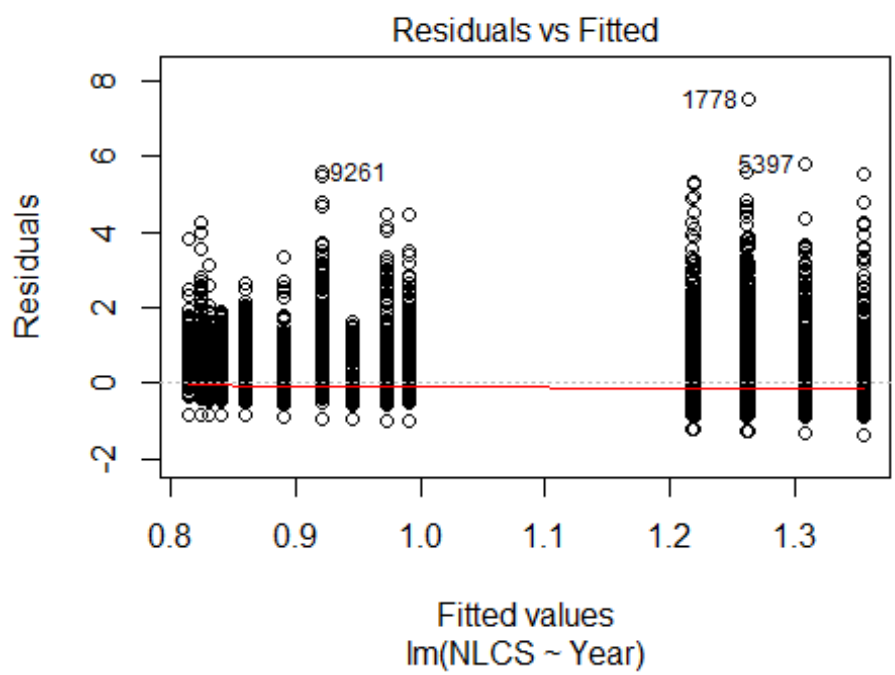
```
## 5675  27844592291 3.959 2005     1212      1     2.944
## 5765  27844515677 3.959 2005     1212      1     2.909
## 5766  27844593205 4.199 2005     1212      1     3.184
## 5813  21644482697 4.818 2005     1212      1     3.846
## 5817  27844577835 3.904 2005     1212      1     2.811
## 5880  84992776219 3.720 2005     1212      1     2.705
## 5949  61249595569 3.788 2005     1208      2     2.773
## 5950  61449306806 3.654 2005     1208      2     2.561
## 5952  61449561847 3.582 2005     1208      2     2.532
## 5955  62449266342 3.723 2005     1208      2     2.708
## 6177  33751047195 4.104 2006     1212      1     2.973
## 6187  34347309336 3.704 2006     1212      1     2.650
## 6223  60950249825 3.675 2006     1212      1     2.621
## 6324  41949093695 4.329 2006     1212      2     3.241
## 6332  43249163482 3.642 2006     1212      2     2.588
## 6453  33747151905 3.863 2006     1212      1     2.809
## 6456  33747171138 4.064 2006     1212      1     3.010
## 6458  33747184520 4.298 2006     1212      1     3.244
## 6522  33646756864 4.376 2006     1212      1     3.245
## 6523  33646757510 4.886 2006     1212      1     3.755
## 6957  36949003441 3.577 2007     1212      1     2.577
## 6959  36949019141 4.370 2007     1212      1     3.327
## 6961  36949024013 4.168 2007     1212      1     3.125
## 7296  34548661972 4.013 2007     1212      1     2.970
## 7297  34548675370 3.736 2007     1212      1     2.693
## 7388  34249740121 4.357 2007     1212      1     3.314
## 7472  33947233870 4.662 2007     1212      1     3.619
## 7478  33947270473 3.547 2007     1212      1     2.547
## 7721  62449186372 3.857 2007     1208      2     2.779
## 7926  57749143176 3.986 2008     1212      1     3.037
## 8004  77955455590 4.118 2008     1212      1     3.169
## 8030  61049425894 3.837 2008     1212      2     2.853
## 8236  56049110254 4.473 2008     1212      1     3.524
## 8238  56049115364 4.257 2008     1212      1     3.230
## 8239  56049118467 3.771 2008     1212      1     2.744
## 8241  56049125485 3.986 2008     1212      1     3.002
## 8242  56049126209 3.899 2008     1212      1     2.950
## 8294  51249088778 3.545 2008     1212      1     2.561
## 8296  51249095120 4.093 2008     1212      1     3.109
## 8299  51249104866 4.118 2008     1212      1     3.169
## 8378  44349086981 4.014 2008     1212      1     3.065
## 8379  44349090359 3.664 2008     1212      1     2.758
## 8382  44349132774 4.418 2008     1212      1     3.512
## 8446  39049105256 3.837 2008     1212      1     2.931
## 8457  40949157971 4.508 2008     1212      1     3.602
## 8862  71649089787 3.994 2009     1212      1     3.106
## 8863  71649091181 3.465 2009     1212      1     2.577
## 8868  71649103465 4.736 2009     1212      1     3.848
## 9055  70349676067 3.812 2009     1212      1     2.924
## 9108  70049086788 4.125 2009     1212      1     3.160
```

```
## 9113  70049102705 4.243 2009     1212      1     3.355
## 9114  70049107140 4.499 2009     1212      1     3.611
## 9131  74949109006 3.465 2009     1212      1     2.620
## 9215  77649288637 3.450 2009     1212      2     2.562
## 9237  66949141813 4.445 2009     1212      1     3.557
## 9238  66849089428 4.804 2009     1212      1     3.839
## 9240  66849099520 4.125 2009     1212      1     3.237
## 9242  66849113683 4.048 2009     1212      1     3.083
## 9243  66849131260 3.671 2009     1212      1     2.826
## 9352  62349085774 3.671 2009     1212      1     2.783
## 9355  62349113564 3.966 2009     1212      1     3.078
## 9357  62349125186 4.075 2009     1212      1     3.187
## 9358  62349129158 4.286 2009     1212      1     3.441
## 9389  61449262127 3.418 2009     1212      1     2.530
## 9522  77954193832 3.633 2009     1212      1     2.745
## 9775  78049519127 3.637 2010     1212      1     2.726
## 9784  78649886543 3.440 2010     1212      1     2.529
## 9785  78649898747 3.916 2010     1212      1     3.040
## 9946  84864916678 4.301 2010     1212      1     3.425
## 10035 77956406486 3.385 2010     1212      1     2.509
## 10180 77954430414 4.196 2010     1212      1     3.320
## 10183 77954438289 3.681 2010     1212      1     2.805
## 10184 77954439455 3.492 2010     1212      1     2.616
## 10186 77954449929 3.880 2010     1212      1     3.004
## 10187 77954451120 3.591 2010     1212      1     2.715
## 10357 79956372974 3.668 2010     1202      4     2.792
## 10382 77950251483 3.492 2010     1212      1     2.616
## 10392 77950931872 4.699 2010     1212      1     3.823
## 10393 77950932977 4.168 2010     1212      1     3.335
## 10398 77950952104 4.992 2010     1212      1     4.116
## 10399 77950953030 4.350 2010     1212      1     3.517
## 10810 82955178180 3.559 2011     1212      1     2.570
## 10811 82955227609 3.938 2011     1212      1     3.026
## 10815 82955234077 3.644 2011     1212      1     2.655
## 10816 82955234080 5.243 2011     1212      1     4.289
## 11021 84881605265 3.644 2011     1212      1     2.690
## 11043 84894292374 4.121 2011     1212      1     3.167
## 11088 80155194058 3.906 2011     1202      5     2.917
## 11183 80052195465 3.870 2011     1212      1     2.916
## 11184 80052203482 3.870 2011     1212      1     2.881
## 11186 80052224481 4.944 2011     1212      1     3.990
## 11187 80052230221 3.938 2011     1212      1     2.984
## 11189 80052246166 4.002 2011     1212      1     3.048
## 11190 80052254269 4.781 2011     1212      1     3.749
## 11428 84893744785 3.559 2011     1212      1     2.647
## 11441 79957671344 4.063 2011     1212      1     3.109
## 11443 79957677461 4.063 2011     1212      1     3.074
## 11444 79957679982 4.002 2011     1212      1     3.013
## 11456 79958069144 3.644 2011     1212      1     2.690
## 11458 79958076416 3.724 2011     1212      1     2.735
```

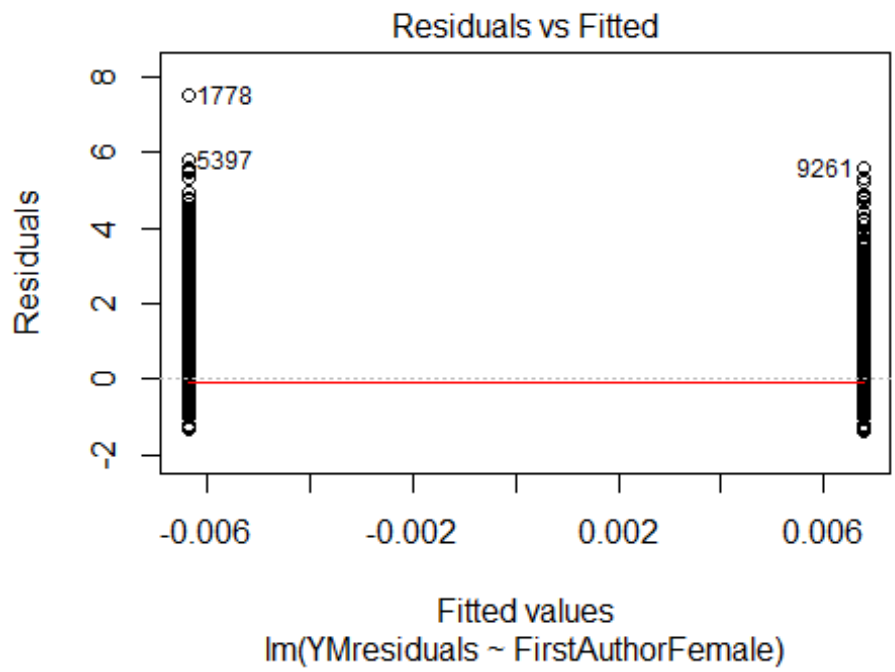
```
## 11669 84870277898 5.639 2011     1212      1     4.685
## 11672 79952213862 4.177 2011     1212      1     3.223
## 11673 79952214878 4.746 2011     1212      1     3.792
## 11676 79952224166 4.944 2011     1212      1     3.990
## 11677 79952229517 4.231 2011     1212      1     3.199
## 11678 79952231734 5.267 2011     1212      1     4.313
## 11758 84856228825 6.890 2011     1212      1     5.978
## 11896 84555179123 4.538 2011     1208      2     3.584
## 12058 84872371974 6.016 2012     1212      1     5.008
## 12065 84872420393 4.778 2012     1212      1     3.847
## 12314 84868144975 4.380 2012     1212      1     3.449
## 12453 84865697715 3.939 2012     1208      2     3.008
## 12454 84865700895 4.859 2012     1208      2     3.894
## 12469 84865721280 3.707 2012     1212      1     2.699
## 12476 84865764944 3.965 2012     1212      1     3.000
## 12743 84861553184 3.559 2012     1212      1     2.551
## 12753 84862130809 4.976 2012     1212      1     4.045
## 12754 84862133446 3.559 2012     1212      1     2.628
## 12755 84862144323 4.286 2012     1212      1     3.355
## 12768 84860861091 3.804 2012     1211      2     2.873
## 12819 84869471464 4.071 2012     1211      2     3.140
## 12871 84860600228 4.186 2012     1212      1     3.255
## 12945 84859452716 3.707 2012     1212      1     2.776
## 13027 84857938890 4.380 2012     1212      1     3.415
## 13031 84857967717 5.540 2012     1212      1     4.609
## 13033 84857979842 4.380 2012     1212      1     3.415
## 13042 84858207835 4.913 2012     1212      1     3.948
## 13049 84858256380 6.195 2012     1212      1     5.230
## 13056 84863396437 4.847 2012     1212      1     3.839
## 13148 84855208708 3.559 2012     1212      1     2.628
## 13186 84866240255 4.913 2012     1212      1     3.982
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.1147 -0.8788 -0.0462  0.6617  5.9085
##
## Coefficients:
##                   Estimate Std. Error t value Pr(>|t|)
## (Intercept)         0.8386     0.0450   18.66  < 2e-16 ***
## LastAuthorFemale1   0.0243     0.0199    1.22  0.22091
## Year1997            0.0139     0.0647    0.21  0.83059
## Year1998            0.1762     0.0594    2.96  0.00304 **
## Year1999            0.0536     0.0621    0.86  0.38868
## Year2000            0.1140     0.0585    1.95  0.05136 .
## Year2001            0.0614     0.0578    1.06  0.28863
```

```
## Year2002            0.2104     0.0610    3.45  0.00057 ***
## Year2003            0.2517     0.0616    4.09  4.4e-05 ***
## Year2004            0.2028     0.0612    3.31  0.00093 ***
## Year2005            0.1786     0.0596    3.00  0.00273 **
## Year2006            0.2184     0.0568    3.84  0.00012 ***
## Year2007            0.2059     0.0544    3.78  0.00016 ***
## Year2008            0.1123     0.0545    2.06  0.03942 *
## Year2009            0.0510     0.0544    0.94  0.34857
## Year2010            0.0402     0.0539    0.74  0.45640
## Year2011            0.1185     0.0528    2.25  0.02479 *
## Year2012            0.0943     0.0538    1.75  0.07970 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.929
## Multiple R-squared:  0.00642,    Adjusted R-squared:  0.0049
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  8 observations c(586,977,2912,9770,9848,10096,10873,10890)
##   are outliers with |weight| = 0 ( < 9e-06);
##  926 weights are ~= 1. The remaining 10213 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0001  0.8890  0.9250  0.9060  0.9840  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          8.97e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Sample size for the above analysis:  11147"
## [1] ""
## [1] ""
## [1] "####################################"
## [1] "Analysis of AJSC 1213"
## [1] "####################################"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
```

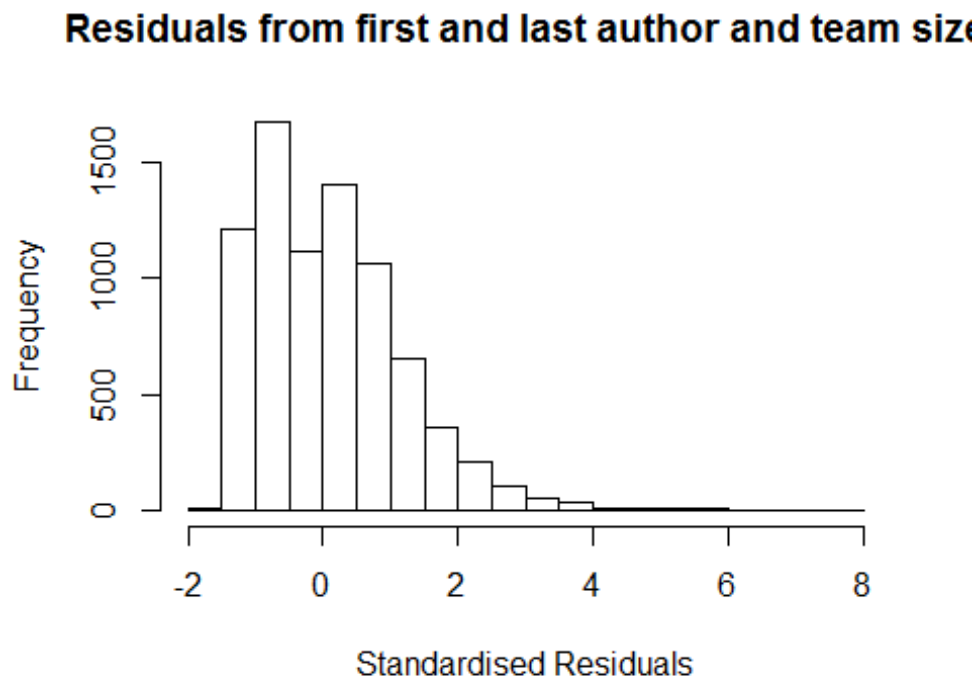
```
##  286  266  249  275  285  291  583  579  658  675  728  708  675  575  603
## 2011 2012
##  724 1069
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  247  229  214  241  260  256  519  509  578  591  639  611  598  501  498
## 2011 2012
##  600  927
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  244  223  211  239  258  253  513  507  573  583  625  601  584  492  485
## 2011 2012
##  586  916
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
##  Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 330, df = 16, p-value <2e-16
```




```
##
##  Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 0.077, df = 1, p-value = 0.8
```




```
## [1] "Female first author team size 2018 geometric mean: 1.12946028923721"
## [1] "Male first author team size 2018 geometric mean: 1.10523326889363"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 120000, p-value = 0.2
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 1.11189441857757"
## [1] "Male last author team size 2018 geometric mean: 1.12664316095034"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 110000, p-value = 0.8
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 5.445  1           2.333
## LastAuthorFemale  5.410  1           2.326
## UniqueAuthors     1.058  4           1.007
## Year              1.046 16           1.001
```




```
## [1] "List of  221 outliers with residuals above  2.5"
##         ScopusId  NLCS Year OneField Fields residuals
## 96    0007314215 3.292 1996     1213      1     2.556
## 102  62449259298 4.627 1996     1213      1     3.907
## 362  84876320180 3.334 1997     1210      3     2.596
## 391  43349105749 3.514 1997     1213      1     2.760
## 903  79959669049 3.949 1999     1213      1     3.188
## 1019  0013039289 3.420 1999     1208      3     2.659
## 1201 33748906142 4.237 2000     1213      1     3.416
## 1202 57949091899 3.432 2000     1213      1     2.611
## 1295 85011501181 3.568 2000     1213      2     2.762
## 1742 61049504518 4.879 2002     1213      1     3.761
## 1778  3142697241 8.757 2002     1208      2     7.655
## 1780 33646911853 5.146 2002     1208      2     4.028
## 1781 33748905673 5.092 2002     1208      2     3.990
## 1785 44149119239 3.798 2002     1208      2     2.680
## 1803 60949601992 4.417 2002     1208      2     3.299
## 1821 60950259244 4.417 2002     1208      2     3.299
## 2012 78650508823 3.843 2002     1213      1     2.725
## 2013 80051741609 4.229 2002     1213      1     3.111
## 2016 84901236898 3.690 2002     1213      1     2.588
## 2054 61049101473 4.366 2002     1213      2     3.248
## 2059 62449109752 4.033 2002     1208      2     2.915
## 2131 84998183101 4.984 2002     1213      2     3.866
## 2182  3142779116 4.956 2002     1213      3     3.854
## 2196 60950279634 3.963 2002     1208      3     2.845
## 2280  0036526749 3.670 2002     1202      6     2.552
```

```
## 2282  1442335800 3.694 2002     1202      6     2.592
## 2315 60950075969 4.072 2003     1213      1     2.931
## 2319 60950157550 4.599 2003     1213      1     3.474
## 2321 60950180274 3.911 2003     1213      1     2.770
## 2358 26944502435 4.044 2003     1208      2     2.903
## 2359 33646681740 6.083 2003     1208      2     4.942
## 2367 34547600256 4.341 2003     1208      2     3.200
## 2402 60950066571 4.869 2003     1208      2     3.728
## 2651 33749552300 5.088 2003     1213      1     3.686
## 2656 34249351441 3.911 2003     1213      1     2.770
## 2666 77956825002 5.467 2003     1213      1     3.719
## 2690 48549096636 5.652 2003     1208      2     4.511
## 2756 84997941907 3.951 2003     1213      2     2.810
## 2885 34249345547 4.059 2004     1213      1     2.861
## 2934 61449284695 6.897 2004     1213      1     5.699
## 2966 34250312985 4.933 2004     1208      2     3.720
## 2995 60950391613 4.521 2004     1208      2     3.308
## 3021 60950719934 3.827 2004     1208      2     2.629
## 3198 60949482998 4.653 2004     1213      1     3.440
## 3237 34249135961 5.281 2004     1213      1     3.804
## 3240 34547752001 3.874 2004     1213      1     2.661
## 3242 60249096596 3.874 2004     1213      1     2.661
## 3245 61049303528 6.110 2004     1213      1     4.912
## 3246 61149229629 4.522 2004     1213      1     3.324
## 3270 85008579728 4.888 2004     1213      1     3.690
## 3272 85008582168 4.227 2004     1213      1     3.014
## 3278 85012439959 3.874 2004     1213      1     2.661
## 3282 85012517282 4.059 2004     1213      1     2.846
## 3284 85012545761 4.059 2004     1213      1     2.846
## 3307 60950338479 5.580 2004     1208      2     4.367
## 3322 61849165077 3.827 2004     1208      2     2.614
## 3345 70449786556 4.699 2004     1208      2     3.501
## 3347 70449824679 4.933 2004     1208      2     3.735
## 3386 84998075563 5.553 2004     1213      2     4.355
## 3633 34250008387 4.562 2005     1208      2     3.491
## 3635 38649134273 6.094 2005     1208      2     5.007
## 3641 60949254410 4.465 2005     1208      2     3.394
## 3660 60950350045 5.477 2005     1208      2     4.406
## 3689 60950562957 3.861 2005     1208      2     2.790
## 3698 60950642207 3.861 2005     1208      2     2.790
## 3750 33845250551 4.010 2005     1202      4     2.939
## 3775 60950110696 3.592 2005     1202      3     2.505
## 3920 60950504855 3.650 2005     1208      3     2.563
## 4024 78650499949 4.052 2005     1213      1     2.965
## 4072 44449106909 4.250 2005     1208      2     3.163
## 4075 60950423747 5.118 2005     1208      2     4.031
## 4142 84998090320 4.780 2005     1213      2     3.709
## 4163 33846333428 3.605 2005     1208      3     2.534
## 4260 34249427022 5.304 2006     1213      1     4.261
## 4281 61049161278 3.745 2006     1213      1     2.718
```

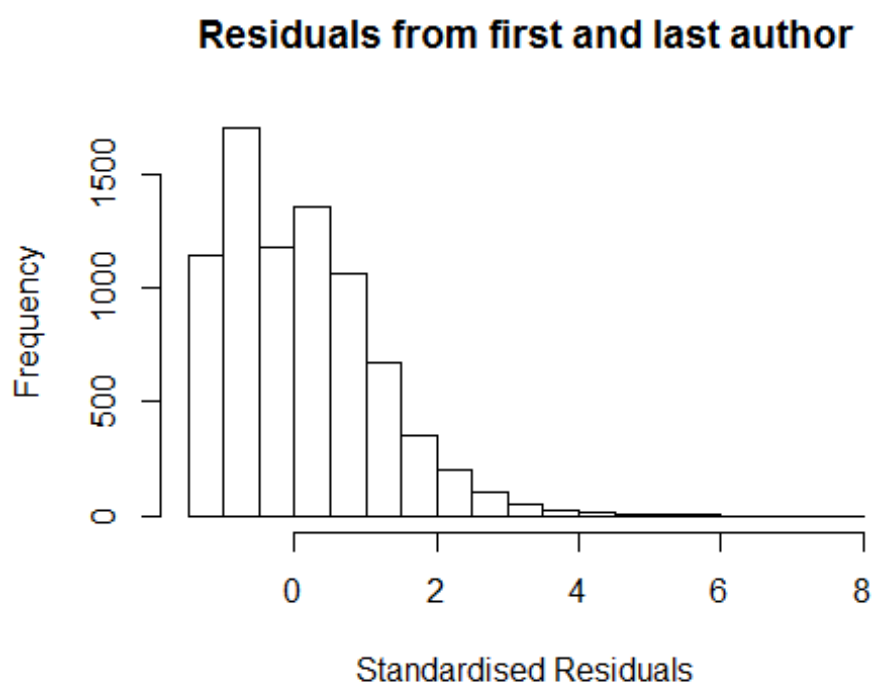
```
## 4284 61049322526 5.304 2006     1213      1     4.277
## 4299 63849133332 3.958 2006     1213      1     2.915
## 4320 41949093695 4.329 2006     1212      2     3.286
## 4321 43249163482 3.642 2006     1212      2     2.615
## 4353 60950491226 6.170 2006     1210      2     5.143
## 4610 60949389015 5.724 2006     1213      1     4.697
## 4700 34249142355 3.745 2006     1213      1     2.702
## 4702 34748865278 6.455 2006     1213      1     5.412
## 4711 64549121127 4.476 2006     1213      1     3.433
## 4720 84948118751 4.476 2006     1213      1     3.449
## 4722 84974819817 3.745 2006     1213      1     2.718
## 4724 84983368119 5.103 2006     1213      1     4.076
## 4732 85008525695 4.476 2006     1213      1     3.449
## 4757 34249084654 5.314 2006     1208      2     4.271
## 4765 41449090508 4.162 2006     1202      2     3.119
## 4782 44449151884 6.523 2006     1208      2     5.480
## 4783 60949270586 3.627 2006     1208      2     2.600
## 4808 61449282187 3.627 2006     1208      2     2.600
## 4811 61449530625 6.523 2006     1208      2     5.496
## 4829 70349296952 3.627 2006     1208      2     2.600
## 4832 70449775017 4.234 2006     1208      2     3.207
## 4842 70450078889 3.627 2006     1208      2     2.584
## 4864 84997860330 3.886 2006     1213      2     2.859
## 4870 84998153444 4.441 2006     1213      2     3.414
## 5006 43249136736 3.870 2007     1213      1     2.730
## 5007 43249138158 4.285 2007     1213      1     3.130
## 5036 61149150533 3.870 2007     1213      1     2.715
## 5059 43249135021 4.018 2007     1208      2     2.863
## 5069 43249146684 4.726 2007     1208      2     3.571
## 5097 60949438530 4.726 2007     1208      2     3.586
## 5127 60950643835 4.839 2007     1208      2     3.699
## 5130 60950662433 3.834 2007     1208      2     2.694
## 5359 43249145818 4.442 2007     1213      2     3.287
## 5381 84973631259 4.730 2007     1206      3     3.590
## 5396 84869154206 4.089 2007     1213      1     2.934
## 5397 36049016758 7.080 2007     1210      2     5.940
## 5431 63449100602 3.870 2007     1213      1     2.730
## 5436 77950873213 5.666 2007     1213      1     4.526
## 5440 84860473637 4.285 2007     1213      1     3.130
## 5451 84969630204 4.912 2007     1213      1     3.772
## 5453 85008521054 4.463 2007     1213      1     3.323
## 5468 85012475415 4.089 2007     1213      1     2.934
## 5481 43249133166 3.704 2007     1202      2     2.564
## 5538 70450078906 4.018 2007     1208      2     2.878
## 5567 84997943760 4.345 2007     1213      2     2.928
## 5644 61249239835 4.188 2007     1208      3     3.048
## 5652 61249660537 3.741 2007     1208      3     2.601
## 5775 60950096222 3.727 2008     1213      1     2.643
## 5792 61049095439 4.598 2008     1213      1     3.514
## 5796 61049355621 4.598 2008     1213      1     3.514
```

```
## 5806 61449337801 4.454 2008     1213      1     3.370
## 5817 63449139280 3.727 2008     1213      1     2.643
## 5851 60950666896 4.798 2008     1213      2     3.714
## 5863 61049425894 3.837 2008     1212      2     2.753
## 5875 61149472770 4.222 2008     1208      2     3.138
## 5991 65749261392 3.585 2008     1202      3     2.501
## 6056 60949119937 3.727 2008     1213      1     2.659
## 6057 60949253613 4.598 2008     1213      1     3.514
## 6101 57049157739 5.668 2008     1210      2     4.336
## 6102 57049173452 5.953 2008     1210      2     4.621
## 6103 57049179599 6.823 2008     1210      2     5.755
## 6169 64949099880 4.127 2008     1213      1     3.059
## 6174 77952917870 5.079 2008     1213      1     3.995
## 6181 84876153884 4.298 2008     1213      1     3.214
## 6226 61249468411 4.222 2008     1208      2     3.154
## 6235 70449355960 4.922 2008     1213      2     3.589
## 6237 77955354528 4.524 2008     1213      2     3.177
## 6262 84997884125 4.500 2008     1213      2     3.416
## 6263 84997901848 5.848 2008     1213      2     4.780
## 6269 84998153436 4.500 2008     1213      2     3.432
## 6485 61949462044 3.777 2009     1208      3     2.949
## 6597 70349828937 4.501 2009     1213      1     3.657
## 6617 73649148390 3.360 2009     1202      4     2.532
## 6638 68349157263 3.856 2009     1213      1     3.012
## 6649 77649288637 3.450 2009     1212      2     2.622
## 6654 67149097677 5.424 2009     1213      1     4.596
## 6683 65849172939 3.402 2009     1210      3     2.574
## 6736 61149318798 3.649 2009     1213      1     2.805
## 6744 67650102337 4.376 2009     1213      2     3.548
## 6753 75149146477 4.172 2009     1210      2     3.344
## 6807 79551671924 4.501 2009     1213      1     3.394
## 6808 79952872192 3.856 2009     1213      1     2.636
## 6812 84859341926 3.649 2009     1213      1     2.805
## 6815 84882067675 3.415 2009     1213      1     2.587
## 6817 84916206263 3.649 2009     1213      1     2.805
## 6818 84930366138 3.649 2009     1213      1     2.805
## 7013 77950787861 3.540 2010     1213      1     2.884
## 7015 77950848584 3.540 2010     1213      1     2.869
## 7142 77955800148 3.635 2010     1210      2     2.716
## 7144 77956331789 4.376 2010     1208      2     3.720
## 7147 78249251522 3.350 2010     1210      2     2.694
## 7233 79959202296 3.634 2010     1208      2     2.978
## 7579 77956500865 4.776 2010     1208      3     3.513
## 7582 78649472664 5.064 2010     1208      3     3.832
## 7600 82455228764 3.393 2010     1208      3     2.737
## 7616 77951908318 3.254 2010     1202      6     2.598
## 7654 84855781328 3.573 2011     1208      2     2.776
## 7655 84855791481 5.406 2011     1208      2     4.593
## 7669 84858822292 3.704 2011     1210      2     2.907
## 7784 80155194058 3.906 2011     1202      5     3.093
```

```
## 7861 80053198127 3.665 2011     1213      1     2.868
## 7883 84861070464 5.091 2011     1208      2     4.294
## 7956 79961064842 3.817 2011     1213      1     3.020
## 8027 79958705208 4.106 2011     1213      3     3.293
## 8147 79953030885 3.665 2011     1213      1     2.868
## 8149 79953054054 3.665 2011     1213      1     2.852
## 8181 79952719487 3.448 2011     1213      3     2.651
## 8232 82455226568 3.497 2011     1213      1     2.684
## 8347 82455232963 4.204 2011     1208      3     3.128
## 8348 82455244076 4.997 2011     1208      3     3.937
## 8419 84878611571 3.247 2012     1213      1     2.513
## 8499 84871367509 3.245 2012     1202      3     2.511
## 8500 84871394290 5.615 2012     1202      3     4.881
## 8544 84874341423 3.664 2012     1202      3     2.946
## 8593 84869392883 3.247 2012     1213      1     2.529
## 8620 84867392058 4.065 2012     1213      2     3.347
## 8681 84868631017 3.559 2012     1208      2     2.825
## 8682 84868671128 3.962 2012     1208      2     3.244
## 8729 84866600452 3.926 2012     1213      1     3.192
## 8739 84866634692 3.615 2012     1213      1     2.881
## 8751 84883428181 4.433 2012     1213      1     3.699
## 8790 84869152027 3.467 2012     1210      3     2.749
## 8824 84864711707 3.258 2012     1202      5     2.540
## 8833 84864926904 4.302 2012     1208      2     3.584
## 8891 84861970276 3.900 2012     1213      2     3.166
## 8945 84883516915 4.433 2012     1213      1     3.699
## 8996 84863094733 3.615 2012     1213      1     2.881
## 8999 84859949763 3.495 2012     1213      2     2.761
## 9001 84859998112 3.239 2012     1213      2     2.521
## 9048 84861569470 3.264 2012     1210      3     2.530
## 9070 84864138989 5.716 2012     1213      1     4.389
## 9088 84992122601 4.645 2012     1213      1     3.911
## 9108 84861350433 4.598 2012     1208      2     3.864
## 9109 84861359249 3.559 2012     1208      2     2.825
## 9112 84861395909 3.559 2012     1208      2     2.825
## 9145 84860734606 3.999 2012     1202      3     3.265
## 9153 84862115131 3.380 2012     1202      4     2.646
## 9154 84862116557 3.380 2012     1202      4     2.662
## 9182 84856817391 3.466 2012     1213      2     2.748
## 9201 84860317282 4.195 2012     1213      1     3.461
## 9208 84869098440 3.247 2012     1213      1     2.529
## 9246 84865710516 3.559 2012     1208      2     2.825
## 9259 84870312908 3.962 2012     1208      2     3.244
## 9261 84870334955 6.510 2012     1208      2     5.776
## 9265 84870345747 6.391 2012     1208      2     5.673
## 9363 84857285608 3.843 2012     1213      3     2.862
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
UniqueAuthors +
```

```
##     Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##     k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.8209 -0.7358 -0.0151  0.7255  7.6548
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         0.72013    0.05315   13.55  < 2e-16 ***
## FirstAuthorFemale1  0.00163    0.05342    0.03  0.97560
## LastAuthorFemale1   0.01403    0.05323    0.26  0.79204
## UniqueAuthors2      0.26340    0.04530    5.82  6.3e-09 ***
## UniqueAuthors3      0.60740    0.07352    8.26  < 2e-16 ***
## UniqueAuthors4      0.37638    0.12398    3.04  0.00241 **
## UniqueAuthors5      0.56233    0.16469    3.41  0.00064 ***
## Year1997            0.01789    0.07432    0.24  0.80974
## Year1998            0.16331    0.07193    2.27  0.02320 *
## Year1999            0.02486    0.07120    0.35  0.72696
## Year2000            0.08558    0.07351    1.16  0.24439
## Year2001            0.05743    0.06912    0.83  0.40601
## Year2002            0.38208    0.07249    5.27  1.4e-07 ***
## Year2003            0.40475    0.07099    5.70  1.2e-08 ***
## Year2004            0.47770    0.07149    6.68  2.5e-11 ***
## Year2005            0.35117    0.06972    5.04  4.8e-07 ***
## Year2006            0.30728    0.06915    4.44  9.0e-06 ***
## Year2007            0.41969    0.07079    5.93  3.2e-09 ***
## Year2008            0.34804    0.06989    4.98  6.5e-07 ***
## Year2009            0.10781    0.06881    1.57  0.11721
## Year2010           -0.06463    0.06477   -1.00  0.31839
## Year2011            0.07687    0.06423    1.20  0.23145
## Year2012           -0.00205    0.06194   -0.03  0.97357
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.968
## Multiple R-squared:  0.0429, Adjusted R-squared:  0.0402
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
##  19 observations
c(1497,1996,2491,2757,3089,3687,3896,3971,4037,4063,4561,5150,5151,5279,5619,
6409,7083,7711,7713)
##   are outliers with |weight| = 0 ( < 1.3e-05);
##  686 weights are ~= 1. The remaining 7188 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.000   0.882   0.940   0.898   0.977   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
```

```
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.27e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 6.063  1           2.462
## LastAuthorFemale  6.052  1           2.460
## Year              1.015 16           1.000
```




```
## [1] "List of  219 outliers with residuals above  2.5"
##         ScopusId  NLCS Year OneField Fields residuals
## 96    0007314215 3.292 1996     1213      1     2.520
## 102  62449259298 4.627 1996     1213      1     3.870
## 362  84876320180 3.334 1997     1210      3     2.542
## 391  43349105749 3.514 1997     1213      1     2.708
```

```
## 903  79959669049 3.949 1999     1213      1     3.157
## 1019  0013039289 3.420 1999     1208      3     2.628
## 1201 33748906142 4.237 2000     1213      1     3.382
## 1202 57949091899 3.432 2000     1213      1     2.577
## 1295 85011501181 3.568 2000     1213      2     2.728
## 1742 61049504518 4.879 2002     1213      1     3.729
## 1778  3142697241 8.757 2002     1208      2     7.622
## 1780 33646911853 5.146 2002     1208      2     3.996
## 1781 33748905673 5.092 2002     1208      2     3.957
## 1785 44149119239 3.798 2002     1208      2     2.648
## 1803 60949601992 4.417 2002     1208      2     3.267
## 1821 60950259244 4.417 2002     1208      2     3.267
## 2012 78650508823 3.843 2002     1213      1     2.693
## 2013 80051741609 4.229 2002     1213      1     3.079
## 2016 84901236898 3.690 2002     1213      1     2.555
## 2019 84937380085 3.843 2002     1213      1     2.708
## 2054 61049101473 4.366 2002     1213      2     3.216
## 2059 62449109752 4.033 2002     1208      2     2.883
## 2131 84998183101 4.984 2002     1213      2     3.834
## 2182  3142779116 4.956 2002     1213      3     3.821
## 2196 60950279634 3.963 2002     1208      3     2.813
## 2280  0036526749 3.670 2002     1202      6     2.520
## 2282  1442335800 3.694 2002     1202      6     2.559
## 2315 60950075969 4.072 2003     1213      1     2.908
## 2319 60950157550 4.599 2003     1213      1     3.450
## 2321 60950180274 3.911 2003     1213      1     2.747
## 2358 26944502435 4.044 2003     1208      2     2.880
## 2359 33646681740 6.083 2003     1208      2     4.919
## 2367 34547600256 4.341 2003     1208      2     3.177
## 2402 60950066571 4.869 2003     1208      2     3.705
## 2651 33749552300 5.088 2003     1213      1     3.909
## 2656 34249351441 3.911 2003     1213      1     2.747
## 2666 77956825002 5.467 2003     1213      1     4.303
## 2690 48549096636 5.652 2003     1208      2     4.488
## 2756 84997941907 3.951 2003     1213      2     2.787
## 2885 34249345547 4.059 2004     1213      1     2.831
## 2934 61449284695 6.897 2004     1213      1     5.669
## 2966 34250312985 4.933 2004     1208      2     3.691
## 2995 60950391613 4.521 2004     1208      2     3.279
## 3021 60950719934 3.827 2004     1208      2     2.599
## 3198 60949482998 4.653 2004     1213      1     3.411
## 3237 34249135961 5.281 2004     1213      1     4.039
## 3238 34249155307 3.874 2004     1213      1     2.632
## 3240 34547752001 3.874 2004     1213      1     2.632
## 3242 60249096596 3.874 2004     1213      1     2.632
## 3245 61049303528 6.110 2004     1213      1     4.882
## 3246 61149229629 4.522 2004     1213      1     3.294
## 3270 85008579728 4.888 2004     1213      1     3.660
## 3272 85008582168 4.227 2004     1213      1     2.985
## 3278 85012439959 3.874 2004     1213      1     2.632
```

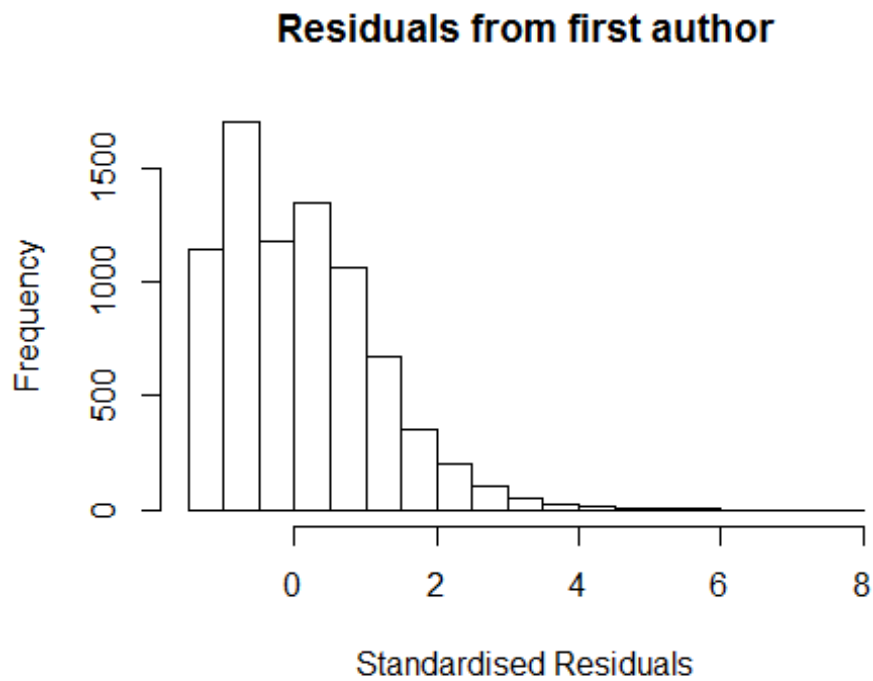
```
## 3282 85012517282 4.059 2004     1213      1     2.817
## 3284 85012545761 4.059 2004     1213      1     2.817
## 3307 60950338479 5.580 2004     1208      2     4.338
## 3322 61849165077 3.827 2004     1208      2     2.585
## 3345 70449786556 4.699 2004     1208      2     3.471
## 3347 70449824679 4.933 2004     1208      2     3.705
## 3386 84998075563 5.553 2004     1213      2     4.325
## 3633 34250008387 4.562 2005     1208      2     3.461
## 3635 38649134273 6.094 2005     1208      2     4.978
## 3641 60949254410 4.465 2005     1208      2     3.364
## 3660 60950350045 5.477 2005     1208      2     4.376
## 3689 60950562957 3.861 2005     1208      2     2.760
## 3698 60950642207 3.861 2005     1208      2     2.760
## 3750 33845250551 4.010 2005     1202      4     2.909
## 3920 60950504855 3.650 2005     1208      3     2.534
## 4024 78650499949 4.052 2005     1213      1     2.936
## 4072 44449106909 4.250 2005     1208      2     3.134
## 4075 60950423747 5.118 2005     1208      2     4.002
## 4142 84998090320 4.780 2005     1213      2     3.679
## 4163 33846333428 3.605 2005     1208      3     2.504
## 4260 34249427022 5.304 2006     1213      1     4.229
## 4281 61049161278 3.745 2006     1213      1     2.685
## 4284 61049322526 5.304 2006     1213      1     4.244
## 4299 63849133332 3.958 2006     1213      1     2.883
## 4320 41949093695 4.329 2006     1212      2     3.254
## 4321 43249163482 3.642 2006     1212      2     2.582
## 4353 60950491226 6.170 2006     1210      2     5.110
## 4610 60949389015 5.724 2006     1213      1     4.664
## 4700 34249142355 3.745 2006     1213      1     2.670
## 4702 34748865278 6.455 2006     1213      1     5.380
## 4711 64549121127 4.476 2006     1213      1     3.401
## 4720 84948118751 4.476 2006     1213      1     3.416
## 4722 84974819817 3.745 2006     1213      1     2.685
## 4724 84983368119 5.103 2006     1213      1     4.043
## 4732 85008525695 4.476 2006     1213      1     3.416
## 4757 34249084654 5.314 2006     1208      2     4.239
## 4765 41449090508 4.162 2006     1202      2     3.087
## 4782 44449151884 6.523 2006     1208      2     5.448
## 4783 60949270586 3.627 2006     1208      2     2.567
## 4808 61449282187 3.627 2006     1208      2     2.567
## 4811 61449530625 6.523 2006     1208      2     5.463
## 4829 70349296952 3.627 2006     1208      2     2.567
## 4832 70449775017 4.234 2006     1208      2     3.174
## 4842 70450078889 3.627 2006     1208      2     2.552
## 4864 84997860330 3.886 2006     1213      2     2.826
## 4870 84998153444 4.441 2006     1213      2     3.381
## 5006 43249136736 3.870 2007     1213      1     2.699
## 5007 43249138158 4.285 2007     1213      1     3.099
## 5036 61149150533 3.870 2007     1213      1     2.684
## 5059 43249135021 4.018 2007     1208      2     2.832
```

```
## 5069 43249146684 4.726 2007     1208      2     3.540
## 5097 60949438530 4.726 2007     1208      2     3.555
## 5127 60950643835 4.839 2007     1208      2     3.668
## 5130 60950662433 3.834 2007     1208      2     2.663
## 5359 43249145818 4.442 2007     1213      2     3.256
## 5381 84973631259 4.730 2007     1206      3     3.559
## 5396 84869154206 4.089 2007     1213      1     2.903
## 5397 36049016758 7.080 2007     1210      2     5.909
## 5421 51249143944 3.870 2007     1213      1     2.684
## 5431 63449100602 3.870 2007     1213      1     2.699
## 5436 77950873213 5.666 2007     1213      1     4.495
## 5440 84860473637 4.285 2007     1213      1     3.099
## 5451 84969630204 4.912 2007     1213      1     3.741
## 5453 85008521054 4.463 2007     1213      1     3.292
## 5468 85012475415 4.089 2007     1213      1     2.903
## 5481 43249133166 3.704 2007     1202      2     2.533
## 5538 70450078906 4.018 2007     1208      2     2.847
## 5567 84997943760 4.345 2007     1213      2     3.145
## 5644 61249239835 4.188 2007     1208      3     3.017
## 5652 61249660537 3.741 2007     1208      3     2.570
## 5775 60950096222 3.727 2008     1213      1     2.601
## 5792 61049095439 4.598 2008     1213      1     3.472
## 5796 61049355621 4.598 2008     1213      1     3.472
## 5806 61449337801 4.454 2008     1213      1     3.328
## 5817 63449139280 3.727 2008     1213      1     2.601
## 5851 60950666896 4.798 2008     1213      2     3.672
## 5863 61049425894 3.837 2008     1212      2     2.711
## 5875 61149472770 4.222 2008     1208      2     3.096
## 6056 60949119937 3.727 2008     1213      1     2.616
## 6057 60949253613 4.598 2008     1213      1     3.472
## 6101 57049157739 5.668 2008     1210      2     4.557
## 6102 57049173452 5.953 2008     1210      2     4.842
## 6103 57049179599 6.823 2008     1210      2     5.712
## 6169 64949099880 4.127 2008     1213      1     3.016
## 6174 77952917870 5.079 2008     1213      1     3.953
## 6181 84876153884 4.298 2008     1213      1     3.172
## 6226 61249468411 4.222 2008     1208      2     3.111
## 6235 70449355960 4.922 2008     1213      2     3.826
## 6237 77955354528 4.524 2008     1213      2     3.398
## 6262 84997884125 4.500 2008     1213      2     3.374
## 6263 84997901848 5.848 2008     1213      2     4.737
## 6269 84998153436 4.500 2008     1213      2     3.389
## 6485 61949462044 3.777 2009     1208      3     2.906
## 6597 70349828937 4.501 2009     1213      1     3.616
## 6638 68349157263 3.856 2009     1213      1     2.971
## 6649 77649288637 3.450 2009     1212      2     2.579
## 6654 67149097677 5.424 2009     1213      1     4.553
## 6683 65849172939 3.402 2009     1210      3     2.531
## 6736 61149318798 3.649 2009     1213      1     2.764
## 6744 67650102337 4.376 2009     1213      2     3.505
```

```
## 6753 75149146477 4.172 2009     1210      2     3.301
## 6807 79551671924 4.501 2009     1213      1     3.616
## 6808 79952872192 3.856 2009     1213      1     2.971
## 6810 79952895760 3.856 2009     1213      1     3.000
## 6812 84859341926 3.649 2009     1213      1     2.764
## 6815 84882067675 3.415 2009     1213      1     2.544
## 6817 84916206263 3.649 2009     1213      1     2.764
## 6818 84930366138 3.649 2009     1213      1     2.764
## 7013 77950787861 3.540 2010     1213      1     2.830
## 7015 77950848584 3.540 2010     1213      1     2.815
## 7142 77955800148 3.635 2010     1210      2     2.925
## 7144 77956331789 4.376 2010     1208      2     3.666
## 7147 78249251522 3.350 2010     1210      2     2.640
## 7233 79959202296 3.634 2010     1208      2     2.924
## 7417 82455232955 3.309 2010     1213      1     2.584
## 7578 77952198073 3.247 2010     1202      3     2.552
## 7579 77956500865 4.776 2010     1208      3     4.066
## 7582 78649472664 5.064 2010     1208      3     4.325
## 7600 82455228764 3.393 2010     1208      3     2.683
## 7616 77951908318 3.254 2010     1202      6     2.544
## 7654 84855781328 3.573 2011     1208      2     2.715
## 7655 84855791481 5.406 2011     1208      2     4.533
## 7664 84857779496 3.573 2011     1208      2     2.685
## 7669 84858822292 3.704 2011     1210      2     2.846
## 7725 84857935340 3.474 2011     1202      6     2.616
## 7784 80155194058 3.906 2011     1202      5     3.033
## 7861 80053198127 3.665 2011     1213      1     2.807
## 7883 84861070464 5.091 2011     1208      2     4.233
## 7956 79961064842 3.817 2011     1213      1     2.959
## 8027 79958705208 4.106 2011     1213      3     3.233
## 8147 79953030885 3.665 2011     1213      1     2.807
## 8149 79953054054 3.665 2011     1213      1     2.792
## 8181 79952719487 3.448 2011     1213      3     2.590
## 8232 82455226568 3.497 2011     1213      1     2.624
## 8347 82455232963 4.204 2011     1208      3     3.331
## 8348 82455244076 4.997 2011     1208      3     4.139
## 8500 84871394290 5.615 2012     1202      3     4.834
## 8544 84874341423 3.664 2012     1202      3     2.898
## 8620 84867392058 4.065 2012     1213      2     3.299
## 8681 84868631017 3.559 2012     1208      2     2.778
## 8682 84868671128 3.962 2012     1208      2     3.196
## 8729 84866600452 3.926 2012     1213      1     3.145
## 8739 84866634692 3.615 2012     1213      1     2.834
## 8751 84883428181 4.433 2012     1213      1     3.652
## 8790 84869152027 3.467 2012     1210      3     2.701
## 8833 84864926904 4.302 2012     1208      2     3.536
## 8891 84861970276 3.900 2012     1213      2     3.119
## 8945 84883516915 4.433 2012     1213      1     3.652
## 8996 84863094733 3.615 2012     1213      1     2.834
## 8999 84859949763 3.495 2012     1213      2     2.714
```

```
## 9070 84864138989 5.716 2012     1213      1     4.965
## 9088 84992122601 4.645 2012     1213      1     3.864
## 9108 84861350433 4.598 2012     1208      2     3.817
## 9109 84861359249 3.559 2012     1208      2     2.778
## 9112 84861395909 3.559 2012     1208      2     2.778
## 9145 84860734606 3.999 2012     1202      3     3.218
## 9153 84862115131 3.380 2012     1202      4     2.599
## 9154 84862116557 3.380 2012     1202      4     2.614
## 9182 84856817391 3.466 2012     1213      2     2.700
## 9201 84860317282 4.195 2012     1213      1     3.414
## 9246 84865710516 3.559 2012     1208      2     2.778
## 9259 84870312908 3.962 2012     1208      2     3.196
## 9261 84870334955 6.510 2012     1208      2     5.729
## 9265 84870345747 6.391 2012     1208      2     5.625
## 9363 84857285608 3.843 2012     1213      3     3.077
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.2569 -0.7808 -0.0242  0.7398  7.6218
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)          0.7569     0.0531   14.26  < 2e-16 ***
## FirstAuthorFemale1  -0.0145     0.0570   -0.25    0.800
## LastAuthorFemale1    0.0294     0.0569    0.52    0.606
## Year1997             0.0346     0.0758    0.46    0.648
## Year1998             0.1681     0.0732    2.30    0.022 *
## Year1999             0.0207     0.0721    0.29    0.774
## Year2000             0.0831     0.0731    1.14    0.256
## Year2001             0.0551     0.0693    0.80    0.426
## Year2002             0.3783     0.0727    5.21  2.0e-07 ***
## Year2003             0.3924     0.0711    5.52  3.5e-08 ***
## Year2004             0.4706     0.0716    6.58  5.2e-11 ***
## Year2005             0.3444     0.0697    4.94  7.9e-07 ***
## Year2006             0.3031     0.0693    4.38  1.2e-05 ***
## Year2007             0.4141     0.0710    5.83  5.6e-09 ***
## Year2008             0.3539     0.0703    5.03  4.9e-07 ***
## Year2009             0.1136     0.0693    1.64    0.101
## Year2010            -0.0470     0.0657   -0.72    0.474
## Year2011             0.1014     0.0651    1.56    0.119
## Year2012             0.0090     0.0623    0.14    0.885
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.992
```

```
## Multiple R-squared:  0.0301, Adjusted R-squared:  0.0279
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
##  18 observations
c(1497,1996,2491,2757,3089,3687,3896,3971,4037,4063,4561,5150,5151,5279,7083,
7560,7711,7713)
##   are outliers with |weight| = 0 ( < 1.3e-05);
##  706 weights are ~= 1. The remaining 7169 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0015  0.8840  0.9410  0.9010  0.9780  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.27e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##                   GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.01  1           1.005
## Year              1.01 16           1.000
```




```
## [1] "List of  219 outliers with residuals above  2.5"
##         ScopusId  NLCS Year OneField Fields residuals
## 96    0007314215 3.292 1996     1213      1     2.520
## 102  62449259298 4.627 1996     1213      1     3.870
## 362  84876320180 3.334 1997     1210      3     2.542
## 391  43349105749 3.514 1997     1213      1     2.708
## 903  79959669049 3.949 1999     1213      1     3.157
## 1019  0013039289 3.420 1999     1208      3     2.628
## 1201 33748906142 4.237 2000     1213      1     3.382
## 1202 57949091899 3.432 2000     1213      1     2.577
## 1295 85011501181 3.568 2000     1213      2     2.728
## 1742 61049504518 4.879 2002     1213      1     3.729
## 1778  3142697241 8.757 2002     1208      2     7.622
## 1780 33646911853 5.146 2002     1208      2     3.996
## 1781 33748905673 5.092 2002     1208      2     3.957
## 1785 44149119239 3.798 2002     1208      2     2.648
## 1803 60949601992 4.417 2002     1208      2     3.267
## 1821 60950259244 4.417 2002     1208      2     3.267
## 2012 78650508823 3.843 2002     1213      1     2.693
## 2013 80051741609 4.229 2002     1213      1     3.079
## 2016 84901236898 3.690 2002     1213      1     2.555
## 2019 84937380085 3.843 2002     1213      1     2.708
## 2054 61049101473 4.366 2002     1213      2     3.216
## 2059 62449109752 4.033 2002     1208      2     2.883
## 2131 84998183101 4.984 2002     1213      2     3.834
## 2182  3142779116 4.956 2002     1213      3     3.821
## 2196 60950279634 3.963 2002     1208      3     2.813
```

```
## 2280  0036526749 3.670 2002     1202      6     2.520
## 2282  1442335800 3.694 2002     1202      6     2.559
## 2315 60950075969 4.072 2003     1213      1     2.908
## 2319 60950157550 4.599 2003     1213      1     3.450
## 2321 60950180274 3.911 2003     1213      1     2.747
## 2358 26944502435 4.044 2003     1208      2     2.880
## 2359 33646681740 6.083 2003     1208      2     4.919
## 2367 34547600256 4.341 2003     1208      2     3.177
## 2402 60950066571 4.869 2003     1208      2     3.705
## 2651 33749552300 5.088 2003     1213      1     3.909
## 2656 34249351441 3.911 2003     1213      1     2.747
## 2666 77956825002 5.467 2003     1213      1     4.303
## 2690 48549096636 5.652 2003     1208      2     4.488
## 2756 84997941907 3.951 2003     1213      2     2.787
## 2885 34249345547 4.059 2004     1213      1     2.831
## 2934 61449284695 6.897 2004     1213      1     5.669
## 2966 34250312985 4.933 2004     1208      2     3.691
## 2995 60950391613 4.521 2004     1208      2     3.279
## 3021 60950719934 3.827 2004     1208      2     2.599
## 3198 60949482998 4.653 2004     1213      1     3.411
## 3237 34249135961 5.281 2004     1213      1     4.039
## 3238 34249155307 3.874 2004     1213      1     2.632
## 3240 34547752001 3.874 2004     1213      1     2.632
## 3242 60249096596 3.874 2004     1213      1     2.632
## 3245 61049303528 6.110 2004     1213      1     4.882
## 3246 61149229629 4.522 2004     1213      1     3.294
## 3270 85008579728 4.888 2004     1213      1     3.660
## 3272 85008582168 4.227 2004     1213      1     2.985
## 3278 85012439959 3.874 2004     1213      1     2.632
## 3282 85012517282 4.059 2004     1213      1     2.817
## 3284 85012545761 4.059 2004     1213      1     2.817
## 3307 60950338479 5.580 2004     1208      2     4.338
## 3322 61849165077 3.827 2004     1208      2     2.585
## 3345 70449786556 4.699 2004     1208      2     3.471
## 3347 70449824679 4.933 2004     1208      2     3.705
## 3386 84998075563 5.553 2004     1213      2     4.325
## 3633 34250008387 4.562 2005     1208      2     3.461
## 3635 38649134273 6.094 2005     1208      2     4.978
## 3641 60949254410 4.465 2005     1208      2     3.364
## 3660 60950350045 5.477 2005     1208      2     4.376
## 3689 60950562957 3.861 2005     1208      2     2.760
## 3698 60950642207 3.861 2005     1208      2     2.760
## 3750 33845250551 4.010 2005     1202      4     2.909
## 3920 60950504855 3.650 2005     1208      3     2.534
## 4024 78650499949 4.052 2005     1213      1     2.936
## 4072 44449106909 4.250 2005     1208      2     3.134
## 4075 60950423747 5.118 2005     1208      2     4.002
## 4142 84998090320 4.780 2005     1213      2     3.679
## 4163 33846333428 3.605 2005     1208      3     2.504
## 4260 34249427022 5.304 2006     1213      1     4.229
```

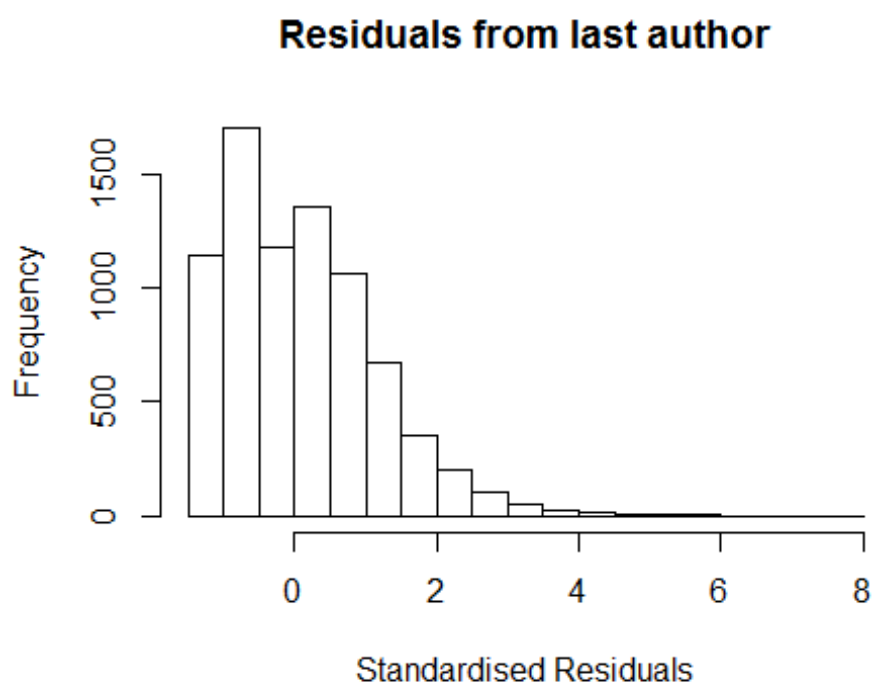
```
## 4281 61049161278 3.745 2006     1213      1     2.685
## 4284 61049322526 5.304 2006     1213      1     4.244
## 4299 63849133332 3.958 2006     1213      1     2.883
## 4320 41949093695 4.329 2006     1212      2     3.254
## 4321 43249163482 3.642 2006     1212      2     2.582
## 4353 60950491226 6.170 2006     1210      2     5.110
## 4610 60949389015 5.724 2006     1213      1     4.664
## 4700 34249142355 3.745 2006     1213      1     2.670
## 4702 34748865278 6.455 2006     1213      1     5.380
## 4711 64549121127 4.476 2006     1213      1     3.401
## 4720 84948118751 4.476 2006     1213      1     3.416
## 4722 84974819817 3.745 2006     1213      1     2.685
## 4724 84983368119 5.103 2006     1213      1     4.043
## 4732 85008525695 4.476 2006     1213      1     3.416
## 4757 34249084654 5.314 2006     1208      2     4.239
## 4765 41449090508 4.162 2006     1202      2     3.087
## 4782 44449151884 6.523 2006     1208      2     5.448
## 4783 60949270586 3.627 2006     1208      2     2.567
## 4808 61449282187 3.627 2006     1208      2     2.567
## 4811 61449530625 6.523 2006     1208      2     5.463
## 4829 70349296952 3.627 2006     1208      2     2.567
## 4832 70449775017 4.234 2006     1208      2     3.174
## 4842 70450078889 3.627 2006     1208      2     2.552
## 4864 84997860330 3.886 2006     1213      2     2.826
## 4870 84998153444 4.441 2006     1213      2     3.381
## 5006 43249136736 3.870 2007     1213      1     2.699
## 5007 43249138158 4.285 2007     1213      1     3.099
## 5036 61149150533 3.870 2007     1213      1     2.684
## 5059 43249135021 4.018 2007     1208      2     2.832
## 5069 43249146684 4.726 2007     1208      2     3.540
## 5097 60949438530 4.726 2007     1208      2     3.555
## 5127 60950643835 4.839 2007     1208      2     3.668
## 5130 60950662433 3.834 2007     1208      2     2.663
## 5359 43249145818 4.442 2007     1213      2     3.256
## 5381 84973631259 4.730 2007     1206      3     3.559
## 5396 84869154206 4.089 2007     1213      1     2.903
## 5397 36049016758 7.080 2007     1210      2     5.909
## 5421 51249143944 3.870 2007     1213      1     2.684
## 5431 63449100602 3.870 2007     1213      1     2.699
## 5436 77950873213 5.666 2007     1213      1     4.495
## 5440 84860473637 4.285 2007     1213      1     3.099
## 5451 84969630204 4.912 2007     1213      1     3.741
## 5453 85008521054 4.463 2007     1213      1     3.292
## 5468 85012475415 4.089 2007     1213      1     2.903
## 5481 43249133166 3.704 2007     1202      2     2.533
## 5538 70450078906 4.018 2007     1208      2     2.847
## 5567 84997943760 4.345 2007     1213      2     3.145
## 5644 61249239835 4.188 2007     1208      3     3.017
## 5652 61249660537 3.741 2007     1208      3     2.570
## 5775 60950096222 3.727 2008     1213      1     2.601
```

```
## 5792 61049095439 4.598 2008     1213      1     3.472
## 5796 61049355621 4.598 2008     1213      1     3.472
## 5806 61449337801 4.454 2008     1213      1     3.328
## 5817 63449139280 3.727 2008     1213      1     2.601
## 5851 60950666896 4.798 2008     1213      2     3.672
## 5863 61049425894 3.837 2008     1212      2     2.711
## 5875 61149472770 4.222 2008     1208      2     3.096
## 6056 60949119937 3.727 2008     1213      1     2.616
## 6057 60949253613 4.598 2008     1213      1     3.472
## 6101 57049157739 5.668 2008     1210      2     4.557
## 6102 57049173452 5.953 2008     1210      2     4.842
## 6103 57049179599 6.823 2008     1210      2     5.712
## 6169 64949099880 4.127 2008     1213      1     3.016
## 6174 77952917870 5.079 2008     1213      1     3.953
## 6181 84876153884 4.298 2008     1213      1     3.172
## 6226 61249468411 4.222 2008     1208      2     3.111
## 6235 70449355960 4.922 2008     1213      2     3.826
## 6237 77955354528 4.524 2008     1213      2     3.398
## 6262 84997884125 4.500 2008     1213      2     3.374
## 6263 84997901848 5.848 2008     1213      2     4.737
## 6269 84998153436 4.500 2008     1213      2     3.389
## 6485 61949462044 3.777 2009     1208      3     2.906
## 6597 70349828937 4.501 2009     1213      1     3.616
## 6638 68349157263 3.856 2009     1213      1     2.971
## 6649 77649288637 3.450 2009     1212      2     2.579
## 6654 67149097677 5.424 2009     1213      1     4.553
## 6683 65849172939 3.402 2009     1210      3     2.531
## 6736 61149318798 3.649 2009     1213      1     2.764
## 6744 67650102337 4.376 2009     1213      2     3.505
## 6753 75149146477 4.172 2009     1210      2     3.301
## 6807 79551671924 4.501 2009     1213      1     3.616
## 6808 79952872192 3.856 2009     1213      1     2.971
## 6810 79952895760 3.856 2009     1213      1     3.000
## 6812 84859341926 3.649 2009     1213      1     2.764
## 6815 84882067675 3.415 2009     1213      1     2.544
## 6817 84916206263 3.649 2009     1213      1     2.764
## 6818 84930366138 3.649 2009     1213      1     2.764
## 7013 77950787861 3.540 2010     1213      1     2.830
## 7015 77950848584 3.540 2010     1213      1     2.815
## 7142 77955800148 3.635 2010     1210      2     2.925
## 7144 77956331789 4.376 2010     1208      2     3.666
## 7147 78249251522 3.350 2010     1210      2     2.640
## 7233 79959202296 3.634 2010     1208      2     2.924
## 7417 82455232955 3.309 2010     1213      1     2.584
## 7578 77952198073 3.247 2010     1202      3     2.552
## 7579 77956500865 4.776 2010     1208      3     4.066
## 7582 78649472664 5.064 2010     1208      3     4.325
## 7600 82455228764 3.393 2010     1208      3     2.683
## 7616 77951908318 3.254 2010     1202      6     2.544
## 7654 84855781328 3.573 2011     1208      2     2.715
```

```
## 7655 84855791481 5.406 2011     1208      2     4.533
## 7664 84857779496 3.573 2011     1208      2     2.685
## 7669 84858822292 3.704 2011     1210      2     2.846
## 7725 84857935340 3.474 2011     1202      6     2.616
## 7784 80155194058 3.906 2011     1202      5     3.033
## 7861 80053198127 3.665 2011     1213      1     2.807
## 7883 84861070464 5.091 2011     1208      2     4.233
## 7956 79961064842 3.817 2011     1213      1     2.959
## 8027 79958705208 4.106 2011     1213      3     3.233
## 8147 79953030885 3.665 2011     1213      1     2.807
## 8149 79953054054 3.665 2011     1213      1     2.792
## 8181 79952719487 3.448 2011     1213      3     2.590
## 8232 82455226568 3.497 2011     1213      1     2.624
## 8347 82455232963 4.204 2011     1208      3     3.331
## 8348 82455244076 4.997 2011     1208      3     4.139
## 8500 84871394290 5.615 2012     1202      3     4.834
## 8544 84874341423 3.664 2012     1202      3     2.898
## 8620 84867392058 4.065 2012     1213      2     3.299
## 8681 84868631017 3.559 2012     1208      2     2.778
## 8682 84868671128 3.962 2012     1208      2     3.196
## 8729 84866600452 3.926 2012     1213      1     3.145
## 8739 84866634692 3.615 2012     1213      1     2.834
## 8751 84883428181 4.433 2012     1213      1     3.652
## 8790 84869152027 3.467 2012     1210      3     2.701
## 8833 84864926904 4.302 2012     1208      2     3.536
## 8891 84861970276 3.900 2012     1213      2     3.119
## 8945 84883516915 4.433 2012     1213      1     3.652
## 8996 84863094733 3.615 2012     1213      1     2.834
## 8999 84859949763 3.495 2012     1213      2     2.714
## 9070 84864138989 5.716 2012     1213      1     4.965
## 9088 84992122601 4.645 2012     1213      1     3.864
## 9108 84861350433 4.598 2012     1208      2     3.817
## 9109 84861359249 3.559 2012     1208      2     2.778
## 9112 84861395909 3.559 2012     1208      2     2.778
## 9145 84860734606 3.999 2012     1202      3     3.218
## 9153 84862115131 3.380 2012     1202      4     2.599
## 9154 84862116557 3.380 2012     1202      4     2.614
## 9182 84856817391 3.466 2012     1213      2     2.700
## 9201 84860317282 4.195 2012     1213      1     3.414
## 9246 84865710516 3.559 2012     1208      2     2.778
## 9259 84870312908 3.962 2012     1208      2     3.196
## 9261 84870334955 6.510 2012     1208      2     5.729
## 9265 84870345747 6.391 2012     1208      2     5.625
## 9363 84857285608 3.843 2012     1213      3     3.077
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
```

```
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.2416 -0.7795 -0.0247  0.7411  7.6206
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         0.75833    0.05298   14.31  < 2e-16 ***
## FirstAuthorFemale1  0.01234    0.02325    0.53    0.595
## Year1997            0.03459    0.07574    0.46    0.648
## Year1998            0.16840    0.07321    2.30    0.021 *
## Year1999            0.02012    0.07205    0.28    0.780
## Year2000            0.08316    0.07311    1.14    0.255
## Year2001            0.05518    0.06927    0.80    0.426
## Year2002            0.37808    0.07265    5.20  2.0e-07 ***
## Year2003            0.39209    0.07105    5.52  3.5e-08 ***
## Year2004            0.47095    0.07157    6.58  5.0e-11 ***
## Year2005            0.34422    0.06968    4.94  8.0e-07 ***
## Year2006            0.30298    0.06923    4.38  1.2e-05 ***
## Year2007            0.41381    0.07095    5.83  5.7e-09 ***
## Year2008            0.35399    0.07032    5.03  4.9e-07 ***
## Year2009            0.11329    0.06925    1.64    0.102
## Year2010           -0.04740    0.06571   -0.72    0.471
## Year2011            0.10099    0.06508    1.55    0.121
## Year2012            0.00884    0.06229    0.14    0.887
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.992
## Multiple R-squared:  0.03,   Adjusted R-squared:  0.028
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
##  18 observations
c(1497,1996,2491,2757,3089,3687,3896,3971,4037,4063,4561,5150,5151,5279,7083,
7560,7711,7713)
##   are outliers with |weight| = 0 ( < 1.3e-05);
##  703 weights are ~= 1. The remaining 7172 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0015  0.8840  0.9410  0.9010  0.9780  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.27e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
```

```
##                    psi           subsampling                    cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                   "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.008  1           1.004
## Year             1.008 16           1.000
```




```
## [1] "List of  219 outliers with residuals above  2.5"
##         ScopusId  NLCS Year OneField Fields residuals
## 96    0007314215 3.292 1996     1213      1     2.520
## 102  62449259298 4.627 1996     1213      1     3.870
## 362  84876320180 3.334 1997     1210      3     2.542
## 391  43349105749 3.514 1997     1213      1     2.708
## 903  79959669049 3.949 1999     1213      1     3.157
## 1019  0013039289 3.420 1999     1208      3     2.628
## 1201 33748906142 4.237 2000     1213      1     3.382
## 1202 57949091899 3.432 2000     1213      1     2.577
## 1295 85011501181 3.568 2000     1213      2     2.728
## 1742 61049504518 4.879 2002     1213      1     3.729
## 1778  3142697241 8.757 2002     1208      2     7.622
## 1780 33646911853 5.146 2002     1208      2     3.996
## 1781 33748905673 5.092 2002     1208      2     3.957
## 1785 44149119239 3.798 2002     1208      2     2.648
## 1803 60949601992 4.417 2002     1208      2     3.267
```

```
## 1821 60950259244 4.417 2002     1208      2     3.267
## 2012 78650508823 3.843 2002     1213      1     2.693
## 2013 80051741609 4.229 2002     1213      1     3.079
## 2016 84901236898 3.690 2002     1213      1     2.555
## 2019 84937380085 3.843 2002     1213      1     2.708
## 2054 61049101473 4.366 2002     1213      2     3.216
## 2059 62449109752 4.033 2002     1208      2     2.883
## 2131 84998183101 4.984 2002     1213      2     3.834
## 2182  3142779116 4.956 2002     1213      3     3.821
## 2196 60950279634 3.963 2002     1208      3     2.813
## 2280  0036526749 3.670 2002     1202      6     2.520
## 2282  1442335800 3.694 2002     1202      6     2.559
## 2315 60950075969 4.072 2003     1213      1     2.908
## 2319 60950157550 4.599 2003     1213      1     3.450
## 2321 60950180274 3.911 2003     1213      1     2.747
## 2358 26944502435 4.044 2003     1208      2     2.880
## 2359 33646681740 6.083 2003     1208      2     4.919
## 2367 34547600256 4.341 2003     1208      2     3.177
## 2402 60950066571 4.869 2003     1208      2     3.705
## 2651 33749552300 5.088 2003     1213      1     3.909
## 2656 34249351441 3.911 2003     1213      1     2.747
## 2666 77956825002 5.467 2003     1213      1     4.303
## 2690 48549096636 5.652 2003     1208      2     4.488
## 2756 84997941907 3.951 2003     1213      2     2.787
## 2885 34249345547 4.059 2004     1213      1     2.831
## 2934 61449284695 6.897 2004     1213      1     5.669
## 2966 34250312985 4.933 2004     1208      2     3.691
## 2995 60950391613 4.521 2004     1208      2     3.279
## 3021 60950719934 3.827 2004     1208      2     2.599
## 3198 60949482998 4.653 2004     1213      1     3.411
## 3237 34249135961 5.281 2004     1213      1     4.039
## 3238 34249155307 3.874 2004     1213      1     2.632
## 3240 34547752001 3.874 2004     1213      1     2.632
## 3242 60249096596 3.874 2004     1213      1     2.632
## 3245 61049303528 6.110 2004     1213      1     4.882
## 3246 61149229629 4.522 2004     1213      1     3.294
## 3270 85008579728 4.888 2004     1213      1     3.660
## 3272 85008582168 4.227 2004     1213      1     2.985
## 3278 85012439959 3.874 2004     1213      1     2.632
## 3282 85012517282 4.059 2004     1213      1     2.817
## 3284 85012545761 4.059 2004     1213      1     2.817
## 3307 60950338479 5.580 2004     1208      2     4.338
## 3322 61849165077 3.827 2004     1208      2     2.585
## 3345 70449786556 4.699 2004     1208      2     3.471
## 3347 70449824679 4.933 2004     1208      2     3.705
## 3386 84998075563 5.553 2004     1213      2     4.325
## 3633 34250008387 4.562 2005     1208      2     3.461
## 3635 38649134273 6.094 2005     1208      2     4.978
## 3641 60949254410 4.465 2005     1208      2     3.364
## 3660 60950350045 5.477 2005     1208      2     4.376
```

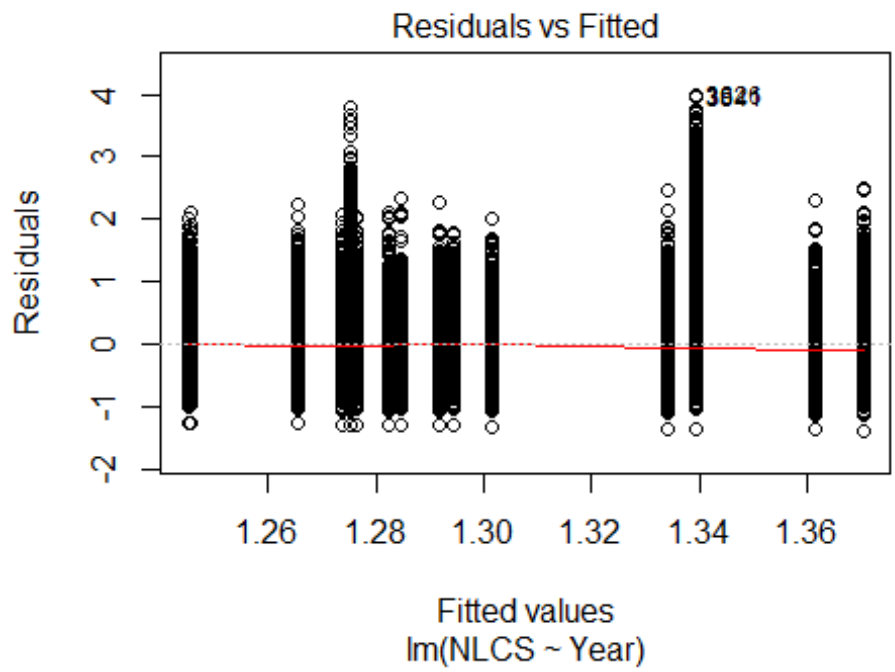
```
## 3689 60950562957 3.861 2005     1208      2     2.760
## 3698 60950642207 3.861 2005     1208      2     2.760
## 3750 33845250551 4.010 2005     1202      4     2.909
## 3920 60950504855 3.650 2005     1208      3     2.534
## 4024 78650499949 4.052 2005     1213      1     2.936
## 4072 44449106909 4.250 2005     1208      2     3.134
## 4075 60950423747 5.118 2005     1208      2     4.002
## 4142 84998090320 4.780 2005     1213      2     3.679
## 4163 33846333428 3.605 2005     1208      3     2.504
## 4260 34249427022 5.304 2006     1213      1     4.229
## 4281 61049161278 3.745 2006     1213      1     2.685
## 4284 61049322526 5.304 2006     1213      1     4.244
## 4299 63849133332 3.958 2006     1213      1     2.883
## 4320 41949093695 4.329 2006     1212      2     3.254
## 4321 43249163482 3.642 2006     1212      2     2.582
## 4353 60950491226 6.170 2006     1210      2     5.110
## 4610 60949389015 5.724 2006     1213      1     4.664
## 4700 34249142355 3.745 2006     1213      1     2.670
## 4702 34748865278 6.455 2006     1213      1     5.380
## 4711 64549121127 4.476 2006     1213      1     3.401
## 4720 84948118751 4.476 2006     1213      1     3.416
## 4722 84974819817 3.745 2006     1213      1     2.685
## 4724 84983368119 5.103 2006     1213      1     4.043
## 4732 85008525695 4.476 2006     1213      1     3.416
## 4757 34249084654 5.314 2006     1208      2     4.239
## 4765 41449090508 4.162 2006     1202      2     3.087
## 4782 44449151884 6.523 2006     1208      2     5.448
## 4783 60949270586 3.627 2006     1208      2     2.567
## 4808 61449282187 3.627 2006     1208      2     2.567
## 4811 61449530625 6.523 2006     1208      2     5.463
## 4829 70349296952 3.627 2006     1208      2     2.567
## 4832 70449775017 4.234 2006     1208      2     3.174
## 4842 70450078889 3.627 2006     1208      2     2.552
## 4864 84997860330 3.886 2006     1213      2     2.826
## 4870 84998153444 4.441 2006     1213      2     3.381
## 5006 43249136736 3.870 2007     1213      1     2.699
## 5007 43249138158 4.285 2007     1213      1     3.099
## 5036 61149150533 3.870 2007     1213      1     2.684
## 5059 43249135021 4.018 2007     1208      2     2.832
## 5069 43249146684 4.726 2007     1208      2     3.540
## 5097 60949438530 4.726 2007     1208      2     3.555
## 5127 60950643835 4.839 2007     1208      2     3.668
## 5130 60950662433 3.834 2007     1208      2     2.663
## 5359 43249145818 4.442 2007     1213      2     3.256
## 5381 84973631259 4.730 2007     1206      3     3.559
## 5396 84869154206 4.089 2007     1213      1     2.903
## 5397 36049016758 7.080 2007     1210      2     5.909
## 5421 51249143944 3.870 2007     1213      1     2.684
## 5431 63449100602 3.870 2007     1213      1     2.699
## 5436 77950873213 5.666 2007     1213      1     4.495
```

```
## 5440 84860473637 4.285 2007     1213      1     3.099
## 5451 84969630204 4.912 2007     1213      1     3.741
## 5453 85008521054 4.463 2007     1213      1     3.292
## 5468 85012475415 4.089 2007     1213      1     2.903
## 5481 43249133166 3.704 2007     1202      2     2.533
## 5538 70450078906 4.018 2007     1208      2     2.847
## 5567 84997943760 4.345 2007     1213      2     3.145
## 5644 61249239835 4.188 2007     1208      3     3.017
## 5652 61249660537 3.741 2007     1208      3     2.570
## 5775 60950096222 3.727 2008     1213      1     2.601
## 5792 61049095439 4.598 2008     1213      1     3.472
## 5796 61049355621 4.598 2008     1213      1     3.472
## 5806 61449337801 4.454 2008     1213      1     3.328
## 5817 63449139280 3.727 2008     1213      1     2.601
## 5851 60950666896 4.798 2008     1213      2     3.672
## 5863 61049425894 3.837 2008     1212      2     2.711
## 5875 61149472770 4.222 2008     1208      2     3.096
## 6056 60949119937 3.727 2008     1213      1     2.616
## 6057 60949253613 4.598 2008     1213      1     3.472
## 6101 57049157739 5.668 2008     1210      2     4.557
## 6102 57049173452 5.953 2008     1210      2     4.842
## 6103 57049179599 6.823 2008     1210      2     5.712
## 6169 64949099880 4.127 2008     1213      1     3.016
## 6174 77952917870 5.079 2008     1213      1     3.953
## 6181 84876153884 4.298 2008     1213      1     3.172
## 6226 61249468411 4.222 2008     1208      2     3.111
## 6235 70449355960 4.922 2008     1213      2     3.826
## 6237 77955354528 4.524 2008     1213      2     3.398
## 6262 84997884125 4.500 2008     1213      2     3.374
## 6263 84997901848 5.848 2008     1213      2     4.737
## 6269 84998153436 4.500 2008     1213      2     3.389
## 6485 61949462044 3.777 2009     1208      3     2.906
## 6597 70349828937 4.501 2009     1213      1     3.616
## 6638 68349157263 3.856 2009     1213      1     2.971
## 6649 77649288637 3.450 2009     1212      2     2.579
## 6654 67149097677 5.424 2009     1213      1     4.553
## 6683 65849172939 3.402 2009     1210      3     2.531
## 6736 61149318798 3.649 2009     1213      1     2.764
## 6744 67650102337 4.376 2009     1213      2     3.505
## 6753 75149146477 4.172 2009     1210      2     3.301
## 6807 79551671924 4.501 2009     1213      1     3.616
## 6808 79952872192 3.856 2009     1213      1     2.971
## 6810 79952895760 3.856 2009     1213      1     3.000
## 6812 84859341926 3.649 2009     1213      1     2.764
## 6815 84882067675 3.415 2009     1213      1     2.544
## 6817 84916206263 3.649 2009     1213      1     2.764
## 6818 84930366138 3.649 2009     1213      1     2.764
## 7013 77950787861 3.540 2010     1213      1     2.830
## 7015 77950848584 3.540 2010     1213      1     2.815
## 7142 77955800148 3.635 2010     1210      2     2.925
```

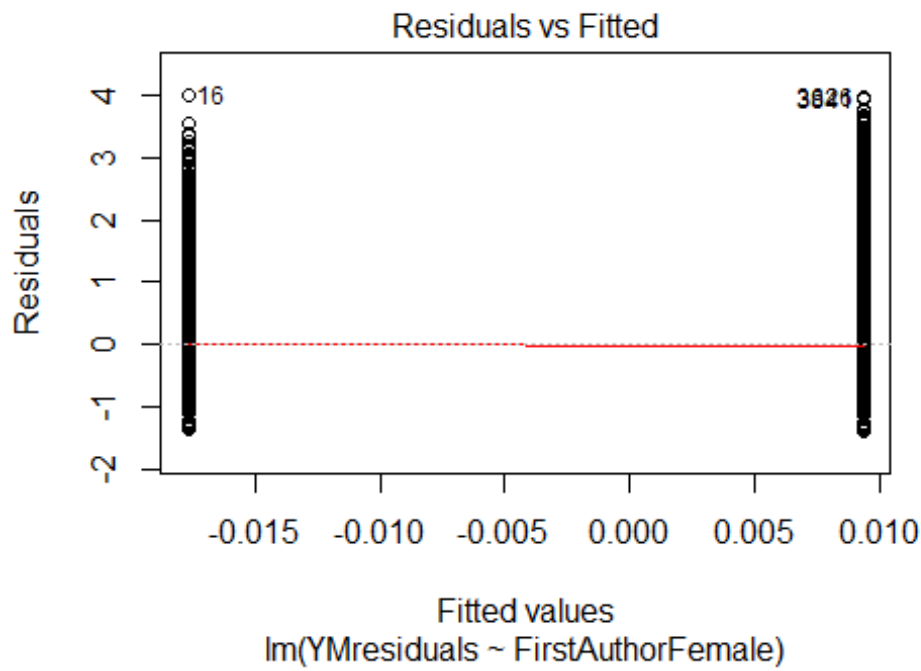
```
## 7144 77956331789 4.376 2010     1208      2     3.666
## 7147 78249251522 3.350 2010     1210      2     2.640
## 7233 79959202296 3.634 2010     1208      2     2.924
## 7417 82455232955 3.309 2010     1213      1     2.584
## 7578 77952198073 3.247 2010     1202      3     2.552
## 7579 77956500865 4.776 2010     1208      3     4.066
## 7582 78649472664 5.064 2010     1208      3     4.325
## 7600 82455228764 3.393 2010     1208      3     2.683
## 7616 77951908318 3.254 2010     1202      6     2.544
## 7654 84855781328 3.573 2011     1208      2     2.715
## 7655 84855791481 5.406 2011     1208      2     4.533
## 7664 84857779496 3.573 2011     1208      2     2.685
## 7669 84858822292 3.704 2011     1210      2     2.846
## 7725 84857935340 3.474 2011     1202      6     2.616
## 7784 80155194058 3.906 2011     1202      5     3.033
## 7861 80053198127 3.665 2011     1213      1     2.807
## 7883 84861070464 5.091 2011     1208      2     4.233
## 7956 79961064842 3.817 2011     1213      1     2.959
## 8027 79958705208 4.106 2011     1213      3     3.233
## 8147 79953030885 3.665 2011     1213      1     2.807
## 8149 79953054054 3.665 2011     1213      1     2.792
## 8181 79952719487 3.448 2011     1213      3     2.590
## 8232 82455226568 3.497 2011     1213      1     2.624
## 8347 82455232963 4.204 2011     1208      3     3.331
## 8348 82455244076 4.997 2011     1208      3     4.139
## 8500 84871394290 5.615 2012     1202      3     4.834
## 8544 84874341423 3.664 2012     1202      3     2.898
## 8620 84867392058 4.065 2012     1213      2     3.299
## 8681 84868631017 3.559 2012     1208      2     2.778
## 8682 84868671128 3.962 2012     1208      2     3.196
## 8729 84866600452 3.926 2012     1213      1     3.145
## 8739 84866634692 3.615 2012     1213      1     2.834
## 8751 84883428181 4.433 2012     1213      1     3.652
## 8790 84869152027 3.467 2012     1210      3     2.701
## 8833 84864926904 4.302 2012     1208      2     3.536
## 8891 84861970276 3.900 2012     1213      2     3.119
## 8945 84883516915 4.433 2012     1213      1     3.652
## 8996 84863094733 3.615 2012     1213      1     2.834
## 8999 84859949763 3.495 2012     1213      2     2.714
## 9070 84864138989 5.716 2012     1213      1     4.965
## 9088 84992122601 4.645 2012     1213      1     3.864
## 9108 84861350433 4.598 2012     1208      2     3.817
## 9109 84861359249 3.559 2012     1208      2     2.778
## 9112 84861395909 3.559 2012     1208      2     2.778
## 9145 84860734606 3.999 2012     1202      3     3.218
## 9153 84862115131 3.380 2012     1202      4     2.599
## 9154 84862116557 3.380 2012     1202      4     2.614
## 9182 84856817391 3.466 2012     1213      2     2.700
## 9201 84860317282 4.195 2012     1213      1     3.414
## 9246 84865710516 3.559 2012     1208      2     2.778
```

```
## 9259 84870312908 3.962 2012     1208      2     3.196
## 9261 84870334955 6.510 2012     1208      2     5.729
## 9265 84870345747 6.391 2012     1208      2     5.625
## 9363 84857285608 3.843 2012     1213      3     3.077
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.2434 -0.7814 -0.0236  0.7392  7.6224
##
## Coefficients:
##                   Estimate Std. Error t value Pr(>|t|)
## (Intercept)        0.75640    0.05305   14.26  < 2e-16 ***
## LastAuthorFemale1  0.01617    0.02322    0.70    0.486
## Year1997           0.03467    0.07577    0.46    0.647
## Year1998           0.16839    0.07320    2.30    0.021 *
## Year1999           0.02053    0.07209    0.28    0.776
## Year2000           0.08306    0.07313    1.14    0.256
## Year2001           0.05497    0.06929    0.79    0.428
## Year2002           0.37818    0.07266    5.20  2.0e-07 ***
## Year2003           0.39224    0.07106    5.52  3.5e-08 ***
## Year2004           0.47080    0.07157    6.58  5.1e-11 ***
## Year2005           0.34424    0.06970    4.94  8.0e-07 ***
## Year2006           0.30293    0.06924    4.37  1.2e-05 ***
## Year2007           0.41394    0.07096    5.83  5.6e-09 ***
## Year2008           0.35386    0.07032    5.03  5.0e-07 ***
## Year2009           0.11349    0.06926    1.64    0.101
## Year2010          -0.04735    0.06572   -0.72    0.471
## Year2011           0.10116    0.06507    1.55    0.120
## Year2012           0.00888    0.06230    0.14    0.887
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.992
## Multiple R-squared:  0.0301, Adjusted R-squared:  0.028
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
##  18 observations
c(1497,1996,2491,2757,3089,3687,3896,3971,4037,4063,4561,5150,5151,5279,7083,
7560,7711,7713)
##   are outliers with |weight| = 0 ( < 1.3e-05);
##  709 weights are ~= 1. The remaining 7166 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0015  0.8840  0.9410  0.9010  0.9780  0.9990
## Algorithmic parameters:
```

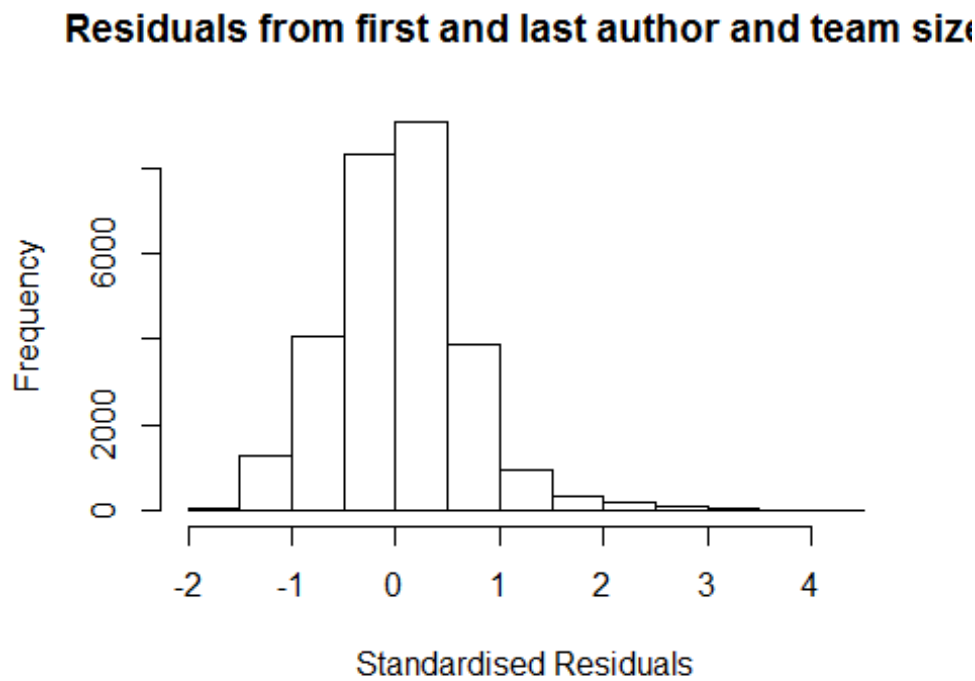
```
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.27e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Sample size for the above analysis:  7893"
## [1] ""
## [1] ""
## [1] "##################################"
## [1] "Analysis of AJSC 1300"
## [1] "##################################"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 6780 4063 3198 1767 1742 1775 1753 1678 1702 1778 1992 2650 2797 3619 4353
## 2011 2012
## 5995 8061
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 1808 1641 1487 1151  977  798 1172 1104 1134 1195 1344 1824 1906 2574 3077
## 2011 2012
## 4244 5628
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 1547 1416 1305 1016  841  698 1028  948  977 1015 1164 1551 1603 2192 2590
## 2011 2012
## 3604 4786
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
##  Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 6300, df = 16, p-value <2e-16
```




```
##
##  Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 62, df = 1, p-value = 3e-15
```




```
## [1] "Female first author team size 2018 geometric mean: 4.84778656056999"
## [1] "Male first author team size 2018 geometric mean: 4.58897846284376"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 2800000, p-value = 0.009
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 4.56232665854665"
## [1] "Male last author team size 2018 geometric mean: 4.74717280324055"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 2100000, p-value = 0.08
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.043  1           1.021
## LastAuthorFemale  1.015  1           1.007
## UniqueAuthors     1.177  4           1.021
## Year              1.173 16           1.005
```




```
## [1] "List of  164 outliers with residuals above  2.5"
##          ScopusId  NLCS Year OneField Fields residuals
## 4      0030447612 4.321 1996     1300      1     2.973
## 6      0030448994 4.349 1996     1300      1     3.052
## 8      0030451819 4.397 1996     1300      1     3.279
## 9      0030460424 5.077 1996     1300      1     3.729
## 12     0030463470 4.705 1996     1300      1     3.606
## 14     0030475008 4.182 1996     1300      1     2.834
## 16     0030480322 5.320 1996     1300      1     4.023
## 57     0030582677 4.257 1996     1300      1     3.158
## 60     0030582732 4.003 1996     1300      1     2.706
## 3626   0030606239 5.315 1996     1300      1     3.967
## 3630   0030606286 3.861 1996     1300      1     2.693
## 3634   0030606320 4.478 1996     1300      1     3.130
## 3689   0030297537 4.474 1996     1300      1     3.432
## 3690   0030297538 3.972 1996     1300      1     2.923
## 3693   0030297895 4.203 1996     1300      1     2.906
## 3694   0030297912 3.842 1996     1300      1     2.793
## 3695   0030298137 4.191 1996     1300      1     3.023
## 3698   0030298375 4.691 1996     1300      1     3.626
## 3718   0029955497 4.505 1996     1300      2     3.208
## 3755   0030592544 4.372 1996     1300      1     3.075
## 3756   0030592556 3.535 1996     1300      1     2.546
## 3758   0030592564 4.050 1996     1300      1     2.702
## 3813   0029820526 4.580 1996     1300      1     3.232
## 3839   0029798819 4.729 1996     1300      2     3.381
## 3841   0029842830 5.287 1996     1300      2     3.939
```

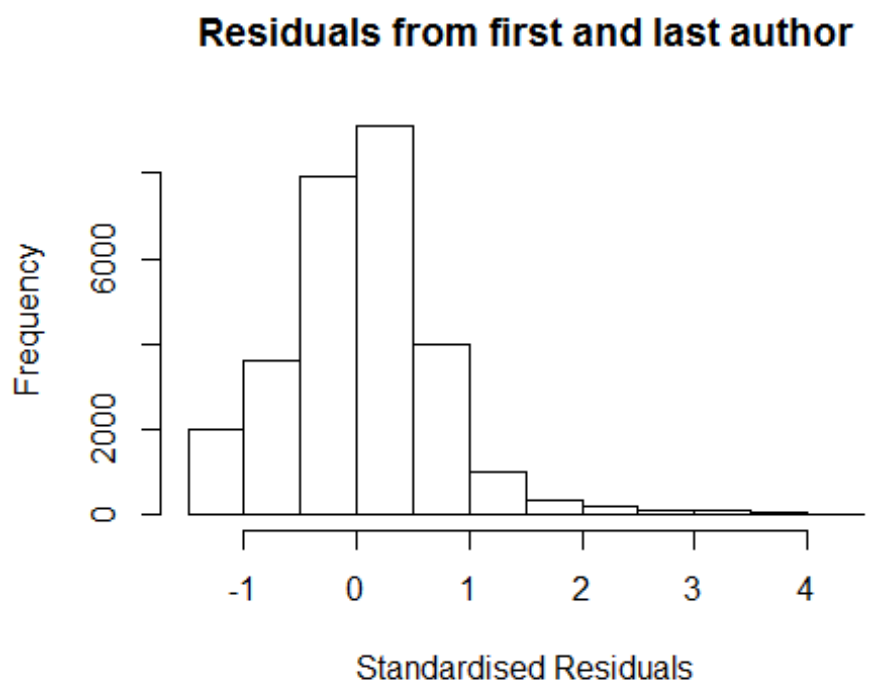
```
## 3881   0030595342 4.361 1996     1300      1     3.064
## 3886   0029817693 4.458 1996     1300      1     3.110
## 3920   0030572695 4.056 1996     1300      1     2.708
## 3922   0030572708 3.928 1996     1300      1     2.580
## 3949   9544226448 4.169 1996     1300      1     2.821
## 3951   0029791838 4.086 1996     1300      2     3.097
## 3967   0030598829 4.927 1996     1300      1     3.828
## 3970   0030598865 3.607 1996     1300      1     2.508
## 3974  16044363014 4.452 1996     1300      1     3.104
## 4005   0030576502 4.479 1996     1300      1     3.182
## 4008   0030576518 4.441 1996     1300      1     3.342
## 4079   0030602822 4.233 1996     1300      1     2.885
## 4080   0030602823 3.783 1996     1300      1     2.791
## 4083  16044364385 4.503 1996     1300      1     3.155
## 4128   0030581152 3.812 1996     1300      1     2.694
## 4131   0030581165 4.171 1996     1300      1     2.823
## 4133   0030581174 4.380 1996     1300      1     3.083
## 4136  15844386540 4.113 1996     1300      1     2.765
## 4153   0029954860 4.282 1996     1300      1     3.293
## 4157   0030010783 4.140 1996     1300      2     3.091
## 4171   0029944241 3.655 1996     1300      1     2.613
## 4173   0029999787 4.337 1996     1300      1     3.288
## 4178  15844378825 4.215 1996     1300      1     2.867
## 4188   0029944290 4.689 1996     1300      2     3.341
## 4191   0029994529 4.291 1996     1300      2     3.242
## 4224   0030604540 3.887 1996     1300      1     2.769
## 4228   0030604722 4.693 1996     1300      1     3.345
## 4231  15844367099 4.221 1996     1300      1     2.873
## 4273   0029894165 3.853 1996     1300      1     2.556
## 4274   0029899127 3.990 1996     1300      1     2.642
## 4275   0029953780 4.085 1996     1300      1     2.737
## 4302   0001506104 4.149 1996     1300      2     3.050
## 4334   0029895156 4.197 1996     1300      1     3.151
## 4336   0029939448 4.169 1996     1300      1     3.070
## 4337   0029943141 4.032 1996     1300      1     2.735
## 4342  15844384256 4.939 1996     1300      1     3.591
## 4343  15844415946 4.832 1996     1300      1     3.484
## 4344  15844420283 4.984 1996     1300      1     3.636
## 4361   0029881125 4.622 1996     1300      1     3.274
## 4362   0029892278 3.564 1996     1300      1     2.515
## 4448   0029870085 4.650 1996     1300      1     3.482
## 4451   0029993728 3.738 1996     1300      1     2.570
## 4452   0030009544 3.963 1996     1300      1     2.914
## 4453  15844372440 3.856 1996     1300      1     2.561
## 4500   0029919935 3.655 1996     1300      1     2.556
## 4501   0029980441 4.225 1996     1300      1     3.126
## 4502   0029993450 4.960 1996     1300      1     3.612
## 4565   0029876473 4.673 1996     1300      1     3.325
## 4588   0029880254 3.670 1996     1300      2     2.502
## 4589   0029880651 4.195 1996     1300      2     3.096
```

```
## 4611   0029961719 3.731 1996     1300      1     3.000
## 4613   0029965130 3.965 1996     1300      1     2.969
## 4614   0029978023 3.696 1996     1300      1     2.965
## 4615   0029991047 4.197 1996     1300      1     3.029
## 4617  13344277364 4.978 1996     1300      1     3.630
## 4703   0030030905 4.203 1996     1300      1     2.908
## 4705   0030048731 4.354 1996     1300      1     3.362
## 4707   0030058657 4.087 1996     1300      1     3.041
## 4714   0030065744 3.607 1996     1300      1     2.558
## 4715  13344261952 4.256 1996     1300      1     2.908
## 4716  13344282063 4.293 1996     1300      1     3.048
## 4727   0030020590 4.565 1996     1300      2     3.447
## 4751   0030024563 4.880 1996     1300      1     3.762
## 4797   0030033699 3.977 1996     1300      1     2.629
## 4808   0030026776 5.105 1996     1300      1     3.757
## 4809   0030026934 4.448 1996     1300      1     3.153
## 4830   0030584077 4.462 1996     1300      1     3.114
## 4831   0030584078 5.024 1996     1300      1     3.676
## 4833   0030584083 3.949 1996     1300      1     2.601
## 4854   0029671310 4.397 1996     1300      1     3.348
## 4856   0030031999 4.203 1996     1300      1     3.154
## 4858   0030034731 4.289 1996     1300      1     3.174
## 4859   0030034983 4.022 1996     1300      1     3.026
## 4860   0030050396 4.336 1996     1300      1     3.237
## 4862   0030061451 3.633 1996     1300      1     2.591
## 4887   0029664992 4.109 1996     1300      1     2.761
## 4901   0029888359 4.449 1996     1300      1     3.152
## 4903   0029890229 3.831 1996     1300      1     2.839
## 4908   0029898733 3.675 1996     1300      1     2.557
## 4917   0029936764 4.426 1996     1300      1     3.258
## 4918   0029940972 4.122 1996     1300      1     3.023
## 4922   0029949784 4.498 1996     1300      1     3.150
## 4926   0029977751 3.719 1996     1300      1     2.988
## 4929   0030014157 3.965 1996     1300      1     2.847
## 4954   0030087712 3.175 1996     1300      1     2.547
## 4974   0030131182 3.669 1996     1300      1     2.938
## 4981   0030139470 3.603 1996     1300      1     2.872
## 5009   0030271388 3.996 1996     1300      1     2.897
## 5010   0030271392 3.996 1996     1300      1     2.897
## 5012   0030271890 3.837 1996     1300      1     2.738
## 5014   0030271999 4.242 1996     1300      1     2.894
## 5015   0030272047 4.276 1996     1300      1     3.108
## 5043  15844380040 4.349 1996     1300      1     3.001
## 5044  15844417385 4.492 1996     1300      1     3.144
## 5868   0030087710 3.738 1996     1100      2     2.696
## 6075   0030606018 3.876 1996     1100      2     3.145
## 7061   0031459980 4.014 1997     1300      1     2.937
## 7229   0031451777 3.981 1997     1300      2     2.847
## 7235   0031466811 3.818 1997     1300      2     2.538
## 9659   0030613551 4.251 1997     1300      1     2.869
```

```
## 9665   0030702123 4.929 1997     1300      1     3.547
## 9696   0030886099 3.752 1997     1300      1     2.618
## 9699   0030886602 3.741 1997     1300      1     2.607
## 9702   0242421666 3.604 1997     1300      1     2.521
## 9773   0030928716 3.647 1997     1300      1     2.513
## 9776   0030954870 3.673 1997     1300      1     2.647
## 9777   0030985459 3.937 1997     1300      1     2.784
## 9838   0030865245 3.667 1997     1300      2     2.584
## 9899   0030788436 3.666 1997     1300      1     2.642
## 9906   0030848970 3.656 1997     1300      1     2.632
## 9966   0030746636 4.625 1997     1300      1     3.542
## 9975   0346613495 3.650 1997     1300      1     2.885
## 9977   0642270732 4.213 1997     1300      1     3.448
## 10027  0031586174 4.066 1997     1300      1     2.734
## 10074  0031440879 3.512 1997     1300      1     2.538
## 10095 16944366965 3.603 1997     1300      1     2.577
## 10117  0030752411 3.710 1997     1300      2     2.610
## 10314  0030687987 3.584 1997     1300      1     2.507
## 10337  0030729445 3.985 1997     1300      1     2.851
## 10344 20244377493 4.035 1997     1300      1     2.653
## 10461  0030970693 4.340 1997     1300      1     3.137
## 10462  0030982264 3.775 1997     1300      1     2.572
## 10519  0030963439 3.978 1997     1300      1     2.596
## 10616  0030944985 4.736 1997     1300      1     3.354
## 10618  0030970775 3.714 1997     1300      1     2.580
## 10620  0031000884 3.817 1997     1300      1     2.683
## 10679  0030949875 3.946 1997     1300      1     2.564
## 10685  0030614893 4.097 1997     1300      1     2.894
## 10688  0031048716 3.922 1997     1300      1     2.590
## 10712  0030893115 4.029 1997     1300      1     2.948
## 10739  0345877774 3.589 1997     1300      1     2.824
## 10740  0345877775 3.865 1997     1300      1     3.100
## 10785  0031285250 3.781 1997     1300      1     2.578
## 10802  0031015075 3.961 1997     1300      1     2.629
## 10839  0030994017 4.009 1997     1300      1     2.875
## 10863  0001679473 3.443 1997     1300      1     2.678
## 10876  0030715563 3.879 1997     1300      1     2.547
## 10884  0030866897 3.700 1997     1300      1     2.550
## 11042  0031022694 3.793 1997     1300      2     2.590
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
UniqueAuthors +
##     Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##     k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.6475 -0.3681  0.0199  0.3909  4.0225
```

```
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         0.73052    0.03927   18.60  < 2e-16 ***
## FirstAuthorFemale1 -0.05022    0.00746   -6.73  1.7e-11 ***
## LastAuthorFemale1  -0.05264    0.00838   -6.28  3.4e-10 ***
## UniqueAuthors2      0.31136    0.01472   21.16  < 2e-16 ***
## UniqueAuthors3      0.36822    0.01494   24.64  < 2e-16 ***
## UniqueAuthors4      0.43759    0.01528   28.64  < 2e-16 ***
## UniqueAuthors5      0.61716    0.01389   44.42  < 2e-16 ***
## Year1997            0.03478    0.05068    0.69  0.49253
## Year1998            0.20139    0.05055    3.98  6.8e-05 ***
## Year1999            0.29978    0.04334    6.92  4.7e-12 ***
## Year2000            0.24372    0.04366    5.58  2.4e-08 ***
## Year2001            0.28521    0.04436    6.43  1.3e-10 ***
## Year2002            0.17222    0.04220    4.08  4.5e-05 ***
## Year2003            0.18551    0.04215    4.40  1.1e-05 ***
## Year2004            0.18302    0.04225    4.33  1.5e-05 ***
## Year2005            0.17339    0.04202    4.13  3.7e-05 ***
## Year2006            0.18398    0.04140    4.44  8.9e-06 ***
## Year2007            0.16526    0.04033    4.10  4.2e-05 ***
## Year2008            0.16618    0.04026    4.13  3.7e-05 ***
## Year2009            0.16596    0.03968    4.18  2.9e-05 ***
## Year2010            0.16154    0.03928    4.11  3.9e-05 ***
## Year2011            0.13581    0.03896    3.49  0.00049 ***
## Year2012            0.09234    0.03871    2.39  0.01706 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.562
## Multiple R-squared:  0.113,  Adjusted R-squared:  0.112
## Convergence in 24 IRWLS iterations
##
## Robustness weights:
##  130 observations
c(1,2,4,5,6,7,9,33,34,495,497,500,527,528,529,530,531,532,541,561,563,592,600
,601,622,624,645,657,658,664,666,681,683,717,718,720,736,737,739,741,748,749,
758,760,764,766,783,784,785,809,810,822,838,839,840,843,844,845,853,894,898,9
26,927,957,971,984,986,987,988,989,1031,1032,1034,1039,1040,1044,1058,1079,10
84,1085,1093,1094,1105,1106,1107,1108,1109,1121,1122,1123,1127,1128,1130,1134
,1135,1162,1168,1185,1186,1188,1189,1190,1199,1200,1284,1437,1555,1624,2105,2
108,2166,2167,2226,2228,2259,2262,2263,2288,2441,2443,2508,2587,2589,2622,263
8,2647,2648,2680,2700,2711)
##   are outliers with |weight| <= 3.5e-06 ( < 3.5e-06);
##  2367 weights are ~= 1. The remaining 25784 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0001  0.8600  0.9500  0.8940  0.9860  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
```

```
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          3.54e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.020  1           1.010
## LastAuthorFemale  1.006  1           1.003
## Year              1.019 16           1.001
```




```
## [1] "List of  201 outliers with residuals above  2.5"
##          ScopusId  NLCS Year OneField Fields residuals
## 4      0030447612 4.321 1996     1300      1     3.282
## 6      0030448994 4.349 1996     1300      1     3.324
## 8      0030451819 4.397 1996     1300      1     3.372
## 9      0030460424 5.077 1996     1300      1     4.038
## 12     0030463470 4.705 1996     1300      1     3.666
```

```
## 14     0030475008 4.182 1996     1300      1     3.143
## 15     0030476713 3.576 1996     1300      1     2.537
## 16     0030480322 5.320 1996     1300      1     4.295
## 57     0030582677 4.257 1996     1300      1     3.218
## 60     0030582732 4.003 1996     1300      1     2.978
## 137   16144361909 3.762 1996     1300      1     2.723
## 3626   0030606239 5.315 1996     1300      1     4.276
## 3630   0030606286 3.861 1996     1300      1     2.822
## 3632   0030606315 3.679 1996     1300      1     2.640
## 3634   0030606320 4.478 1996     1300      1     3.439
## 3689   0030297537 4.474 1996     1300      1     3.435
## 3690   0030297538 3.972 1996     1300      1     2.947
## 3693   0030297895 4.203 1996     1300      1     3.178
## 3694   0030297912 3.842 1996     1300      1     2.817
## 3695   0030298137 4.191 1996     1300      1     3.152
## 3698   0030298375 4.691 1996     1300      1     3.730
## 3718   0029955497 4.505 1996     1300      2     3.480
## 3750   0001265782 3.556 1996     1300      1     2.517
## 3755   0030592544 4.372 1996     1300      1     3.347
## 3756   0030592556 3.535 1996     1300      1     2.560
## 3758   0030592564 4.050 1996     1300      1     3.011
## 3760   0030592578 3.622 1996     1300      1     2.583
## 3813   0029820526 4.580 1996     1300      1     3.541
## 3821  16044371587 3.665 1996     1300      1     2.640
## 3839   0029798819 4.729 1996     1300      2     3.690
## 3841   0029842830 5.287 1996     1300      2     4.248
## 3844  16044361810 3.810 1996     1300      2     2.771
## 3881   0030595342 4.361 1996     1300      1     3.336
## 3886   0029817693 4.458 1996     1300      1     3.419
## 3920   0030572695 4.056 1996     1300      1     3.017
## 3922   0030572708 3.928 1996     1300      1     2.889
## 3949   9544226448 4.169 1996     1300      1     3.130
## 3951   0029791838 4.086 1996     1300      2     3.111
## 3967   0030598829 4.927 1996     1300      1     3.888
## 3970   0030598865 3.607 1996     1300      1     2.568
## 3974  16044363014 4.452 1996     1300      1     3.413
## 4005   0030576502 4.479 1996     1300      1     3.454
## 4008   0030576518 4.441 1996     1300      1     3.402
## 4079   0030602822 4.233 1996     1300      1     3.194
## 4080   0030602823 3.783 1996     1300      1     2.758
## 4082   0030602838 3.722 1996     1300      1     2.683
## 4083  16044364385 4.503 1996     1300      1     3.464
## 4128   0030581152 3.812 1996     1300      1     2.787
## 4131   0030581165 4.171 1996     1300      1     3.132
## 4133   0030581174 4.380 1996     1300      1     3.355
## 4136  15844386540 4.113 1996     1300      1     3.074
## 4153   0029954860 4.282 1996     1300      1     3.307
## 4157   0030010783 4.140 1996     1300      2     3.115
## 4171   0029944241 3.655 1996     1300      1     2.616
## 4173   0029999787 4.337 1996     1300      1     3.312
```

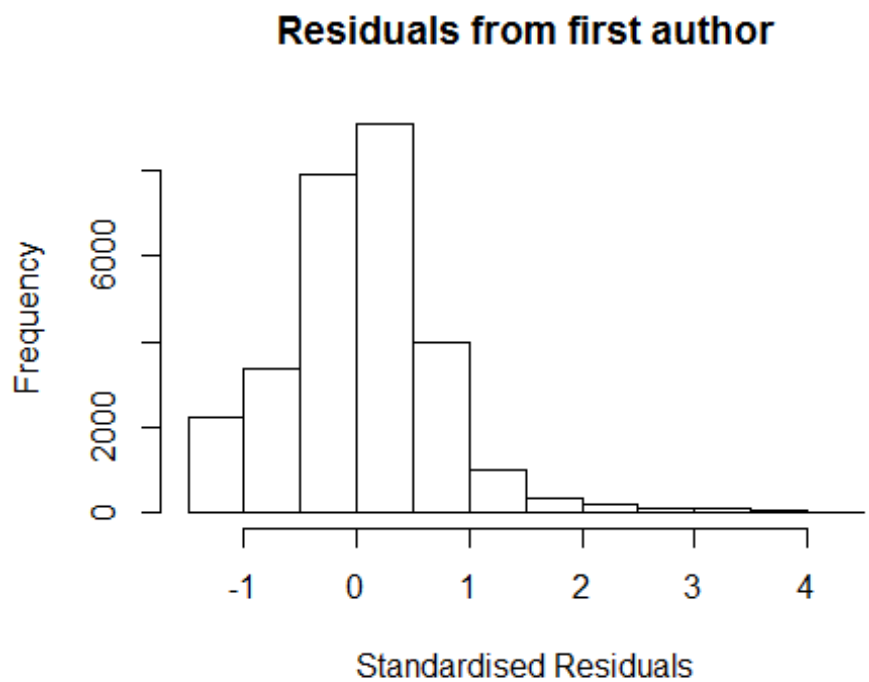
```
## 4178  15844378825 4.215 1996     1300      1     3.176
## 4188   0029944290 4.689 1996     1300      2     3.650
## 4191   0029994529 4.291 1996     1300      2     3.266
## 4224   0030604540 3.887 1996     1300      1     2.862
## 4228   0030604722 4.693 1996     1300      1     3.654
## 4231  15844367099 4.221 1996     1300      1     3.182
## 4273   0029894165 3.853 1996     1300      1     2.828
## 4274   0029899127 3.990 1996     1300      1     2.951
## 4275   0029953780 4.085 1996     1300      1     3.046
## 4302   0001506104 4.149 1996     1300      2     3.110
## 4334   0029895156 4.197 1996     1300      1     3.222
## 4336   0029939448 4.169 1996     1300      1     3.130
## 4337   0029943141 4.032 1996     1300      1     3.007
## 4340   0030010590 3.812 1996     1300      1     2.773
## 4342  15844384256 4.939 1996     1300      1     3.900
## 4343  15844415946 4.832 1996     1300      1     3.793
## 4344  15844420283 4.984 1996     1300      1     3.945
## 4361   0029881125 4.622 1996     1300      1     3.583
## 4362   0029892278 3.564 1996     1300      1     2.539
## 4448   0029870085 4.650 1996     1300      1     3.611
## 4451   0029993728 3.738 1996     1300      1     2.699
## 4452   0030009544 3.963 1996     1300      1     2.938
## 4453  15844372440 3.856 1996     1300      1     2.881
## 4500   0029919935 3.655 1996     1300      1     2.616
## 4501   0029980441 4.225 1996     1300      1     3.186
## 4502   0029993450 4.960 1996     1300      1     3.921
## 4565   0029876473 4.673 1996     1300      1     3.634
## 4588   0029880254 3.670 1996     1300      2     2.631
## 4589   0029880651 4.195 1996     1300      2     3.156
## 4606   0029669982 3.527 1996     1300      1     2.502
## 4611   0029961719 3.731 1996     1300      1     2.692
## 4613   0029965130 3.965 1996     1300      1     3.004
## 4614   0029978023 3.696 1996     1300      1     2.657
## 4615   0029991047 4.197 1996     1300      1     3.158
## 4617  13344277364 4.978 1996     1300      1     3.939
## 4703   0030030905 4.203 1996     1300      1     3.228
## 4705   0030048731 4.354 1996     1300      1     3.329
## 4706   0030053650 3.615 1996     1300      1     2.576
## 4707   0030058657 4.087 1996     1300      1     3.112
## 4714   0030065744 3.607 1996     1300      1     2.582
## 4715  13344261952 4.256 1996     1300      1     3.217
## 4716  13344282063 4.293 1996     1300      1     3.332
## 4727   0030020590 4.565 1996     1300      2     3.540
## 4751   0030024563 4.880 1996     1300      1     3.855
## 4797   0030033699 3.977 1996     1300      1     2.938
## 4808   0030026776 5.105 1996     1300      1     4.066
## 4809   0030026934 4.448 1996     1300      1     3.473
## 4830   0030584077 4.462 1996     1300      1     3.423
## 4831   0030584078 5.024 1996     1300      1     3.985
## 4833   0030584083 3.949 1996     1300      1     2.910
```

```
## 4854   0029671310 4.397 1996     1300      1     3.372
## 4856   0030031999 4.203 1996     1300      1     3.178
## 4858   0030034731 4.289 1996     1300      1     3.314
## 4859   0030034983 4.022 1996     1300      1     3.061
## 4860   0030050396 4.336 1996     1300      1     3.297
## 4862   0030061451 3.633 1996     1300      1     2.594
## 4887   0029664992 4.109 1996     1300      1     3.070
## 4901   0029888359 4.449 1996     1300      1     3.424
## 4903   0029890229 3.831 1996     1300      1     2.806
## 4908   0029898733 3.675 1996     1300      1     2.650
## 4917   0029936764 4.426 1996     1300      1     3.387
## 4918   0029940972 4.122 1996     1300      1     3.083
## 4922   0029949784 4.498 1996     1300      1     3.459
## 4926   0029977751 3.719 1996     1300      1     2.680
## 4929   0030014157 3.965 1996     1300      1     2.940
## 4934   0030030057 3.699 1996     1300      1     2.660
## 4962   0030111466 3.479 1996     1300      1     2.504
## 4974   0030131182 3.669 1996     1300      1     2.630
## 4981   0030139470 3.603 1996     1300      1     2.564
## 5009   0030271388 3.996 1996     1300      1     2.957
## 5010   0030271392 3.996 1996     1300      1     2.957
## 5012   0030271890 3.837 1996     1300      1     2.798
## 5014   0030271999 4.242 1996     1300      1     3.203
## 5015   0030272047 4.276 1996     1300      1     3.237
## 5043  15844380040 4.349 1996     1300      1     3.310
## 5044  15844417385 4.492 1996     1300      1     3.453
## 5868   0030087710 3.738 1996     1100      2     2.699
## 6075   0030606018 3.876 1996     1100      2     2.837
## 7056   0031444409 3.689 1997     1300      1     2.663
## 7057   0031456065 3.655 1997     1300      1     2.629
## 7061   0031459980 4.014 1997     1300      1     2.974
## 7128   0344936739 3.678 1997     1300      1     2.638
## 7227   0031449456 3.566 1997     1300      2     2.540
## 7229   0031451777 3.981 1997     1300      2     2.941
## 7235   0031466811 3.818 1997     1300      2     2.856
## 9659   0030613551 4.251 1997     1300      1     3.211
## 9665   0030702123 4.929 1997     1300      1     3.889
## 9696   0030886099 3.752 1997     1300      1     2.712
## 9699   0030886602 3.741 1997     1300      1     2.701
## 9702   0242421666 3.604 1997     1300      1     2.578
## 9712   0030796646 3.547 1997     1300      1     2.507
## 9773   0030928716 3.647 1997     1300      1     2.607
## 9776   0030954870 3.673 1997     1300      1     2.647
## 9777   0030985459 3.937 1997     1300      1     2.911
## 9838   0030865245 3.667 1997     1300      2     2.641
## 9899   0030788436 3.666 1997     1300      1     2.689
## 9902   0030829387 3.623 1997     1300      1     2.583
## 9906   0030848970 3.656 1997     1300      1     2.679
## 9966   0030746636 4.625 1997     1300      1     3.599
## 9975   0346613495 3.650 1997     1300      1     2.610
```

```
## 9977   0642270732 4.213 1997     1300      1     3.173
## 10022  0030755579 3.562 1997     1300      1     2.522
## 10026  0030877659 3.567 1997     1300      1     2.541
## 10027  0031586174 4.066 1997     1300      1     3.040
## 10074  0031440879 3.512 1997     1300      1     2.550
## 10075  0031444148 3.601 1997     1300      1     2.561
## 10083  0031472234 3.713 1997     1300      1     2.687
## 10095 16944366965 3.603 1997     1300      1     2.577
## 10096  0030970602 3.618 1997     1300      2     2.578
## 10117  0030752411 3.710 1997     1300      2     2.748
## 10314  0030687987 3.584 1997     1300      1     2.544
## 10317  0030712145 3.595 1997     1300      1     2.618
## 10337  0030729445 3.985 1997     1300      1     2.945
## 10343 18844476167 3.641 1997     1300      1     2.615
## 10344 20244377493 4.035 1997     1300      1     2.995
## 10376  0031007189 3.758 1997     1300      1     2.718
## 10461  0030970693 4.340 1997     1300      1     3.300
## 10462  0030982264 3.775 1997     1300      1     2.735
## 10515  0030890721 3.600 1997     1300      1     2.560
## 10519  0030963439 3.978 1997     1300      1     2.938
## 10585  0030933978 3.696 1997     1300      1     2.656
## 10616  0030944985 4.736 1997     1300      1     3.696
## 10618  0030970775 3.714 1997     1300      1     2.674
## 10620  0031000884 3.817 1997     1300      1     2.777
## 10679  0030949875 3.946 1997     1300      1     2.906
## 10685  0030614893 4.097 1997     1300      1     3.057
## 10688  0031048716 3.922 1997     1300      1     2.896
## 10712  0030893115 4.029 1997     1300      1     3.052
## 10739  0345877774 3.589 1997     1300      1     2.549
## 10740  0345877775 3.865 1997     1300      1     2.825
## 10785  0031285250 3.781 1997     1300      1     2.741
## 10802  0031015075 3.961 1997     1300      1     2.935
## 10839  0030994017 4.009 1997     1300      1     2.969
## 10876  0030715563 3.879 1997     1300      1     2.853
## 10881  0030800831 3.627 1997     1300      1     2.587
## 10884  0030866897 3.700 1997     1300      1     2.723
## 10958  0031587821 3.699 1997     1300      1     2.737
## 10959  0031587826 3.614 1997     1300      1     2.574
## 10961  0031587830 3.650 1997     1300      1     2.610
## 10974 16944364768 3.665 1997     1300      1     2.625
## 11042  0031022694 3.793 1997     1300      2     2.753
## 11356  0032446607 3.744 1998     1300      1     2.531
## 12409  0031720906 3.753 1998     1300      2     2.555
## 13176  0032540267 3.749 1998     1300      1     2.536
## 15773  0032953920 3.847 1999     1300      2     2.519
## 15917  0032976690 3.859 1999     1300      2     2.517
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
```

```
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.3804 -0.3769  0.0296  0.4029  4.2952
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.03902    0.03654   28.44  < 2e-16 ***
## FirstAuthorFemale1 -0.01422    0.00775   -1.84  0.06635 .
## LastAuthorFemale1  -0.06352    0.00889   -7.15  9.0e-13 ***
## Year1997            0.00116    0.05152    0.02  0.98199
## Year1998            0.17365    0.05097    3.41  0.00066 ***
## Year1999            0.30302    0.04273    7.09  1.4e-12 ***
## Year2000            0.28384    0.04319    6.57  5.0e-11 ***
## Year2001            0.34134    0.04370    7.81  5.8e-15 ***
## Year2002            0.22534    0.04125    5.46  4.7e-08 ***
## Year2003            0.25252    0.04124    6.12  9.3e-10 ***
## Year2004            0.25365    0.04130    6.14  8.3e-10 ***
## Year2005            0.24666    0.04124    5.98  2.2e-09 ***
## Year2006            0.27145    0.04039    6.72  1.8e-11 ***
## Year2007            0.25004    0.03917    6.38  1.8e-10 ***
## Year2008            0.26412    0.03927    6.73  1.8e-11 ***
## Year2009            0.27345    0.03847    7.11  1.2e-12 ***
## Year2010            0.28635    0.03794    7.55  4.6e-14 ***
## Year2011            0.26542    0.03755    7.07  1.6e-12 ***
## Year2012            0.23460    0.03721    6.30  2.9e-10 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.578
## Multiple R-squared:  0.0181, Adjusted R-squared:  0.0175
## Convergence in 22 IRWLS iterations
##
## Robustness weights:
##  138 observations
c(1,2,4,5,6,7,9,33,34,67,495,497,500,527,528,529,530,531,532,541,561,563,592,
600,601,602,622,624,645,646,657,658,664,666,681,683,717,718,720,736,737,739,7
41,748,749,758,760,764,766,783,784,785,808,809,810,822,838,839,840,841,843,84
4,845,853,894,898,899,926,927,957,971,986,988,989,1031,1032,1034,1039,1040,10
44,1058,1079,1084,1085,1093,1094,1095,1105,1106,1107,1108,1109,1121,1122,1123
,1127,1128,1130,1135,1185,1186,1188,1189,1190,1199,1200,1437,1555,1624,1627,2
105,2108,2125,2167,2259,2263,2288,2331,2441,2443,2462,2508,2509,2539,2587,258
9,2617,2622,2624,2638,2648,2672,2680,2700,2721,2726,2772,2819)
##   are outliers with |weight| = 0 ( < 3.5e-06);
##  2426 weights are ~= 1. The remaining 25717 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.000   0.854   0.950   0.891   0.986   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
```

```
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          3.54e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.016  1           1.008
## Year              1.016 16           1.000
```




```
## [1] "List of  201 outliers with residuals above  2.5"
##          ScopusId  NLCS Year OneField Fields residuals
## 4      0030447612 4.321 1996     1300      1     3.282
## 6      0030448994 4.349 1996     1300      1     3.324
## 8      0030451819 4.397 1996     1300      1     3.372
## 9      0030460424 5.077 1996     1300      1     4.038
## 12     0030463470 4.705 1996     1300      1     3.666
## 14     0030475008 4.182 1996     1300      1     3.143
```

```
## 15     0030476713 3.576 1996     1300      1     2.537
## 16     0030480322 5.320 1996     1300      1     4.295
## 57     0030582677 4.257 1996     1300      1     3.218
## 60     0030582732 4.003 1996     1300      1     2.978
## 137   16144361909 3.762 1996     1300      1     2.723
## 3626   0030606239 5.315 1996     1300      1     4.276
## 3630   0030606286 3.861 1996     1300      1     2.822
## 3632   0030606315 3.679 1996     1300      1     2.640
## 3634   0030606320 4.478 1996     1300      1     3.439
## 3689   0030297537 4.474 1996     1300      1     3.435
## 3690   0030297538 3.972 1996     1300      1     2.947
## 3693   0030297895 4.203 1996     1300      1     3.178
## 3694   0030297912 3.842 1996     1300      1     2.817
## 3695   0030298137 4.191 1996     1300      1     3.152
## 3698   0030298375 4.691 1996     1300      1     3.730
## 3718   0029955497 4.505 1996     1300      2     3.480
## 3750   0001265782 3.556 1996     1300      1     2.517
## 3755   0030592544 4.372 1996     1300      1     3.347
## 3756   0030592556 3.535 1996     1300      1     2.560
## 3758   0030592564 4.050 1996     1300      1     3.011
## 3760   0030592578 3.622 1996     1300      1     2.583
## 3813   0029820526 4.580 1996     1300      1     3.541
## 3821  16044371587 3.665 1996     1300      1     2.640
## 3839   0029798819 4.729 1996     1300      2     3.690
## 3841   0029842830 5.287 1996     1300      2     4.248
## 3844  16044361810 3.810 1996     1300      2     2.771
## 3881   0030595342 4.361 1996     1300      1     3.336
## 3886   0029817693 4.458 1996     1300      1     3.419
## 3920   0030572695 4.056 1996     1300      1     3.017
## 3922   0030572708 3.928 1996     1300      1     2.889
## 3949   9544226448 4.169 1996     1300      1     3.130
## 3951   0029791838 4.086 1996     1300      2     3.111
## 3967   0030598829 4.927 1996     1300      1     3.888
## 3970   0030598865 3.607 1996     1300      1     2.568
## 3974  16044363014 4.452 1996     1300      1     3.413
## 4005   0030576502 4.479 1996     1300      1     3.454
## 4008   0030576518 4.441 1996     1300      1     3.402
## 4079   0030602822 4.233 1996     1300      1     3.194
## 4080   0030602823 3.783 1996     1300      1     2.758
## 4082   0030602838 3.722 1996     1300      1     2.683
## 4083  16044364385 4.503 1996     1300      1     3.464
## 4128   0030581152 3.812 1996     1300      1     2.787
## 4131   0030581165 4.171 1996     1300      1     3.132
## 4133   0030581174 4.380 1996     1300      1     3.355
## 4136  15844386540 4.113 1996     1300      1     3.074
## 4153   0029954860 4.282 1996     1300      1     3.307
## 4157   0030010783 4.140 1996     1300      2     3.115
## 4171   0029944241 3.655 1996     1300      1     2.616
## 4173   0029999787 4.337 1996     1300      1     3.312
## 4178  15844378825 4.215 1996     1300      1     3.176
```

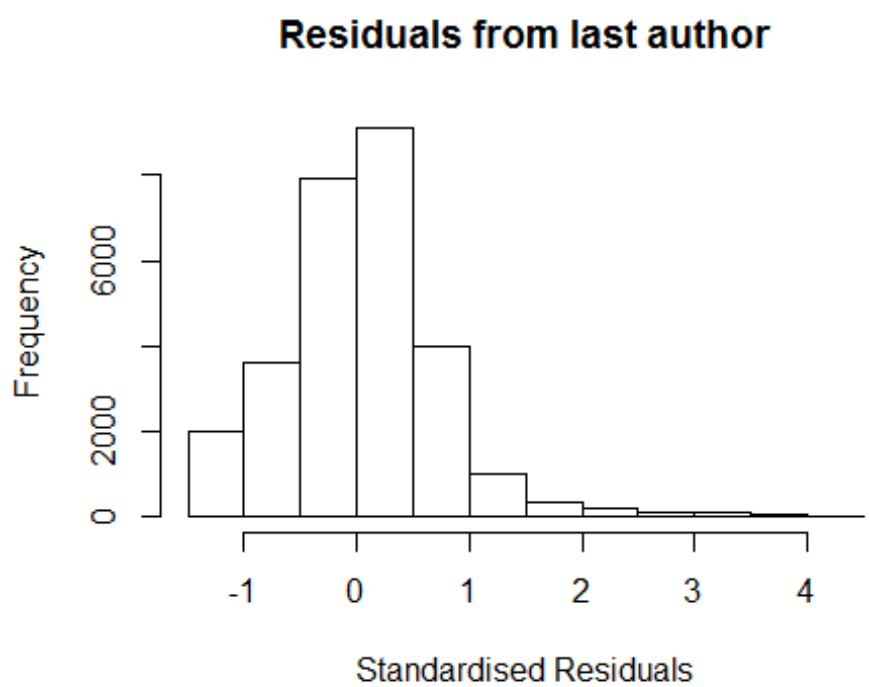
```
## 4188   0029944290 4.689 1996     1300      2     3.650
## 4191   0029994529 4.291 1996     1300      2     3.266
## 4224   0030604540 3.887 1996     1300      1     2.862
## 4228   0030604722 4.693 1996     1300      1     3.654
## 4231  15844367099 4.221 1996     1300      1     3.182
## 4273   0029894165 3.853 1996     1300      1     2.828
## 4274   0029899127 3.990 1996     1300      1     2.951
## 4275   0029953780 4.085 1996     1300      1     3.046
## 4302   0001506104 4.149 1996     1300      2     3.110
## 4334   0029895156 4.197 1996     1300      1     3.222
## 4336   0029939448 4.169 1996     1300      1     3.130
## 4337   0029943141 4.032 1996     1300      1     3.007
## 4340   0030010590 3.812 1996     1300      1     2.773
## 4342  15844384256 4.939 1996     1300      1     3.900
## 4343  15844415946 4.832 1996     1300      1     3.793
## 4344  15844420283 4.984 1996     1300      1     3.945
## 4361   0029881125 4.622 1996     1300      1     3.583
## 4362   0029892278 3.564 1996     1300      1     2.539
## 4448   0029870085 4.650 1996     1300      1     3.611
## 4451   0029993728 3.738 1996     1300      1     2.699
## 4452   0030009544 3.963 1996     1300      1     2.938
## 4453  15844372440 3.856 1996     1300      1     2.881
## 4500   0029919935 3.655 1996     1300      1     2.616
## 4501   0029980441 4.225 1996     1300      1     3.186
## 4502   0029993450 4.960 1996     1300      1     3.921
## 4565   0029876473 4.673 1996     1300      1     3.634
## 4588   0029880254 3.670 1996     1300      2     2.631
## 4589   0029880651 4.195 1996     1300      2     3.156
## 4606   0029669982 3.527 1996     1300      1     2.502
## 4611   0029961719 3.731 1996     1300      1     2.692
## 4613   0029965130 3.965 1996     1300      1     3.004
## 4614   0029978023 3.696 1996     1300      1     2.657
## 4615   0029991047 4.197 1996     1300      1     3.158
## 4617  13344277364 4.978 1996     1300      1     3.939
## 4703   0030030905 4.203 1996     1300      1     3.228
## 4705   0030048731 4.354 1996     1300      1     3.329
## 4706   0030053650 3.615 1996     1300      1     2.576
## 4707   0030058657 4.087 1996     1300      1     3.112
## 4714   0030065744 3.607 1996     1300      1     2.582
## 4715  13344261952 4.256 1996     1300      1     3.217
## 4716  13344282063 4.293 1996     1300      1     3.332
## 4727   0030020590 4.565 1996     1300      2     3.540
## 4751   0030024563 4.880 1996     1300      1     3.855
## 4797   0030033699 3.977 1996     1300      1     2.938
## 4808   0030026776 5.105 1996     1300      1     4.066
## 4809   0030026934 4.448 1996     1300      1     3.473
## 4830   0030584077 4.462 1996     1300      1     3.423
## 4831   0030584078 5.024 1996     1300      1     3.985
## 4833   0030584083 3.949 1996     1300      1     2.910
## 4854   0029671310 4.397 1996     1300      1     3.372
```

```
## 4856   0030031999 4.203 1996     1300      1     3.178
## 4858   0030034731 4.289 1996     1300      1     3.314
## 4859   0030034983 4.022 1996     1300      1     3.061
## 4860   0030050396 4.336 1996     1300      1     3.297
## 4862   0030061451 3.633 1996     1300      1     2.594
## 4887   0029664992 4.109 1996     1300      1     3.070
## 4901   0029888359 4.449 1996     1300      1     3.424
## 4903   0029890229 3.831 1996     1300      1     2.806
## 4908   0029898733 3.675 1996     1300      1     2.650
## 4917   0029936764 4.426 1996     1300      1     3.387
## 4918   0029940972 4.122 1996     1300      1     3.083
## 4922   0029949784 4.498 1996     1300      1     3.459
## 4926   0029977751 3.719 1996     1300      1     2.680
## 4929   0030014157 3.965 1996     1300      1     2.940
## 4934   0030030057 3.699 1996     1300      1     2.660
## 4962   0030111466 3.479 1996     1300      1     2.504
## 4974   0030131182 3.669 1996     1300      1     2.630
## 4981   0030139470 3.603 1996     1300      1     2.564
## 5009   0030271388 3.996 1996     1300      1     2.957
## 5010   0030271392 3.996 1996     1300      1     2.957
## 5012   0030271890 3.837 1996     1300      1     2.798
## 5014   0030271999 4.242 1996     1300      1     3.203
## 5015   0030272047 4.276 1996     1300      1     3.237
## 5043  15844380040 4.349 1996     1300      1     3.310
## 5044  15844417385 4.492 1996     1300      1     3.453
## 5868   0030087710 3.738 1996     1100      2     2.699
## 6075   0030606018 3.876 1996     1100      2     2.837
## 7056   0031444409 3.689 1997     1300      1     2.663
## 7057   0031456065 3.655 1997     1300      1     2.629
## 7061   0031459980 4.014 1997     1300      1     2.974
## 7128   0344936739 3.678 1997     1300      1     2.638
## 7227   0031449456 3.566 1997     1300      2     2.540
## 7229   0031451777 3.981 1997     1300      2     2.941
## 7235   0031466811 3.818 1997     1300      2     2.856
## 9659   0030613551 4.251 1997     1300      1     3.211
## 9665   0030702123 4.929 1997     1300      1     3.889
## 9696   0030886099 3.752 1997     1300      1     2.712
## 9699   0030886602 3.741 1997     1300      1     2.701
## 9702   0242421666 3.604 1997     1300      1     2.578
## 9712   0030796646 3.547 1997     1300      1     2.507
## 9773   0030928716 3.647 1997     1300      1     2.607
## 9776   0030954870 3.673 1997     1300      1     2.647
## 9777   0030985459 3.937 1997     1300      1     2.911
## 9838   0030865245 3.667 1997     1300      2     2.641
## 9899   0030788436 3.666 1997     1300      1     2.689
## 9902   0030829387 3.623 1997     1300      1     2.583
## 9906   0030848970 3.656 1997     1300      1     2.679
## 9966   0030746636 4.625 1997     1300      1     3.599
## 9975   0346613495 3.650 1997     1300      1     2.610
## 9977   0642270732 4.213 1997     1300      1     3.173
```

```
## 10022  0030755579 3.562 1997     1300      1     2.522
## 10026  0030877659 3.567 1997     1300      1     2.541
## 10027  0031586174 4.066 1997     1300      1     3.040
## 10074  0031440879 3.512 1997     1300      1     2.550
## 10075  0031444148 3.601 1997     1300      1     2.561
## 10083  0031472234 3.713 1997     1300      1     2.687
## 10095 16944366965 3.603 1997     1300      1     2.577
## 10096  0030970602 3.618 1997     1300      2     2.578
## 10117  0030752411 3.710 1997     1300      2     2.748
## 10314  0030687987 3.584 1997     1300      1     2.544
## 10317  0030712145 3.595 1997     1300      1     2.618
## 10337  0030729445 3.985 1997     1300      1     2.945
## 10343 18844476167 3.641 1997     1300      1     2.615
## 10344 20244377493 4.035 1997     1300      1     2.995
## 10376  0031007189 3.758 1997     1300      1     2.718
## 10461  0030970693 4.340 1997     1300      1     3.300
## 10462  0030982264 3.775 1997     1300      1     2.735
## 10515  0030890721 3.600 1997     1300      1     2.560
## 10519  0030963439 3.978 1997     1300      1     2.938
## 10585  0030933978 3.696 1997     1300      1     2.656
## 10616  0030944985 4.736 1997     1300      1     3.696
## 10618  0030970775 3.714 1997     1300      1     2.674
## 10620  0031000884 3.817 1997     1300      1     2.777
## 10679  0030949875 3.946 1997     1300      1     2.906
## 10685  0030614893 4.097 1997     1300      1     3.057
## 10688  0031048716 3.922 1997     1300      1     2.896
## 10712  0030893115 4.029 1997     1300      1     3.052
## 10739  0345877774 3.589 1997     1300      1     2.549
## 10740  0345877775 3.865 1997     1300      1     2.825
## 10785  0031285250 3.781 1997     1300      1     2.741
## 10802  0031015075 3.961 1997     1300      1     2.935
## 10839  0030994017 4.009 1997     1300      1     2.969
## 10876  0030715563 3.879 1997     1300      1     2.853
## 10881  0030800831 3.627 1997     1300      1     2.587
## 10884  0030866897 3.700 1997     1300      1     2.723
## 10958  0031587821 3.699 1997     1300      1     2.737
## 10959  0031587826 3.614 1997     1300      1     2.574
## 10961  0031587830 3.650 1997     1300      1     2.610
## 10974 16944364768 3.665 1997     1300      1     2.625
## 11042  0031022694 3.793 1997     1300      2     2.753
## 11356  0032446607 3.744 1998     1300      1     2.531
## 12409  0031720906 3.753 1998     1300      2     2.555
## 13176  0032540267 3.749 1998     1300      1     2.536
## 15773  0032953920 3.847 1999     1300      2     2.519
## 15917  0032976690 3.859 1999     1300      2     2.517
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
```

```
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.3690 -0.3804  0.0319  0.4021  4.3133
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.02945    0.03734   27.57  < 2e-16 ***
## FirstAuthorFemale1 -0.02276    0.00779   -2.92   0.0035 **
## Year1997           -0.00238    0.05157   -0.05   0.9632
## Year1998            0.17194    0.05129    3.35   0.0008 ***
## Year1999            0.30365    0.04333    7.01  2.5e-12 ***
## Year2000            0.28244    0.04384    6.44  1.2e-10 ***
## Year2001            0.33953    0.04441    7.65  2.1e-14 ***
## Year2002            0.22566    0.04197    5.38  7.7e-08 ***
## Year2003            0.25081    0.04193    5.98  2.2e-09 ***
## Year2004            0.25196    0.04195    6.01  1.9e-09 ***
## Year2005            0.24485    0.04196    5.83  5.4e-09 ***
## Year2006            0.26895    0.04116    6.53  6.5e-11 ***
## Year2007            0.24841    0.03997    6.21  5.2e-10 ***
## Year2008            0.26224    0.04010    6.54  6.3e-11 ***
## Year2009            0.27158    0.03930    6.91  5.0e-12 ***
## Year2010            0.28450    0.03878    7.34  2.3e-13 ***
## Year2011            0.26272    0.03838    6.84  7.8e-12 ***
## Year2012            0.23154    0.03805    6.09  1.2e-09 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.576
## Multiple R-squared:  0.0164, Adjusted R-squared:  0.0159
## Convergence in 23 IRWLS iterations
##
## Robustness weights:
##  141 observations
c(1,2,4,5,6,7,9,33,34,67,495,497,500,527,528,529,530,531,532,541,561,563,592,
600,601,602,622,624,645,646,657,658,664,666,681,683,717,718,720,736,737,739,7
41,748,749,758,760,764,766,783,784,785,808,809,810,822,838,839,840,841,843,84
4,845,853,894,897,898,899,926,927,957,971,984,986,988,989,1031,1032,1034,1039
,1040,1044,1058,1079,1084,1085,1093,1094,1095,1105,1106,1107,1108,1109,1121,1
122,1123,1127,1128,1130,1135,1185,1186,1188,1189,1190,1199,1200,1284,1437,155
5,1624,1627,2105,2108,2125,2128,2167,2259,2263,2288,2316,2331,2441,2443,2462,
2508,2509,2539,2587,2589,2617,2622,2624,2638,2648,2672,2680,2700,2721,2819)
##   are outliers with |weight| = 0 ( < 3.5e-06);
##  2460 weights are ~= 1. The remaining 25680 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.000   0.853   0.949   0.890   0.985   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
```

```
##          1.00e-07          1.00e-07          3.54e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##                   GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.003  1           1.001
## Year             1.003 16           1.000
```




```
## [1] "List of  201 outliers with residuals above  2.5"
##          ScopusId  NLCS Year OneField Fields residuals
## 4      0030447612 4.321 1996     1300      1     3.282
## 6      0030448994 4.349 1996     1300      1     3.324
## 8      0030451819 4.397 1996     1300      1     3.372
## 9      0030460424 5.077 1996     1300      1     4.038
## 12     0030463470 4.705 1996     1300      1     3.666
## 14     0030475008 4.182 1996     1300      1     3.143
## 15     0030476713 3.576 1996     1300      1     2.537
## 16     0030480322 5.320 1996     1300      1     4.295
```

```
## 57     0030582677 4.257 1996     1300      1     3.218
## 60     0030582732 4.003 1996     1300      1     2.978
## 137   16144361909 3.762 1996     1300      1     2.723
## 3626   0030606239 5.315 1996     1300      1     4.276
## 3630   0030606286 3.861 1996     1300      1     2.822
## 3632   0030606315 3.679 1996     1300      1     2.640
## 3634   0030606320 4.478 1996     1300      1     3.439
## 3689   0030297537 4.474 1996     1300      1     3.435
## 3690   0030297538 3.972 1996     1300      1     2.947
## 3693   0030297895 4.203 1996     1300      1     3.178
## 3694   0030297912 3.842 1996     1300      1     2.817
## 3695   0030298137 4.191 1996     1300      1     3.152
## 3698   0030298375 4.691 1996     1300      1     3.730
## 3718   0029955497 4.505 1996     1300      2     3.480
## 3750   0001265782 3.556 1996     1300      1     2.517
## 3755   0030592544 4.372 1996     1300      1     3.347
## 3756   0030592556 3.535 1996     1300      1     2.560
## 3758   0030592564 4.050 1996     1300      1     3.011
## 3760   0030592578 3.622 1996     1300      1     2.583
## 3813   0029820526 4.580 1996     1300      1     3.541
## 3821  16044371587 3.665 1996     1300      1     2.640
## 3839   0029798819 4.729 1996     1300      2     3.690
## 3841   0029842830 5.287 1996     1300      2     4.248
## 3844  16044361810 3.810 1996     1300      2     2.771
## 3881   0030595342 4.361 1996     1300      1     3.336
## 3886   0029817693 4.458 1996     1300      1     3.419
## 3920   0030572695 4.056 1996     1300      1     3.017
## 3922   0030572708 3.928 1996     1300      1     2.889
## 3949   9544226448 4.169 1996     1300      1     3.130
## 3951   0029791838 4.086 1996     1300      2     3.111
## 3967   0030598829 4.927 1996     1300      1     3.888
## 3970   0030598865 3.607 1996     1300      1     2.568
## 3974  16044363014 4.452 1996     1300      1     3.413
## 4005   0030576502 4.479 1996     1300      1     3.454
## 4008   0030576518 4.441 1996     1300      1     3.402
## 4079   0030602822 4.233 1996     1300      1     3.194
## 4080   0030602823 3.783 1996     1300      1     2.758
## 4082   0030602838 3.722 1996     1300      1     2.683
## 4083  16044364385 4.503 1996     1300      1     3.464
## 4128   0030581152 3.812 1996     1300      1     2.787
## 4131   0030581165 4.171 1996     1300      1     3.132
## 4133   0030581174 4.380 1996     1300      1     3.355
## 4136  15844386540 4.113 1996     1300      1     3.074
## 4153   0029954860 4.282 1996     1300      1     3.307
## 4157   0030010783 4.140 1996     1300      2     3.115
## 4171   0029944241 3.655 1996     1300      1     2.616
## 4173   0029999787 4.337 1996     1300      1     3.312
## 4178  15844378825 4.215 1996     1300      1     3.176
## 4188   0029944290 4.689 1996     1300      2     3.650
## 4191   0029994529 4.291 1996     1300      2     3.266
```

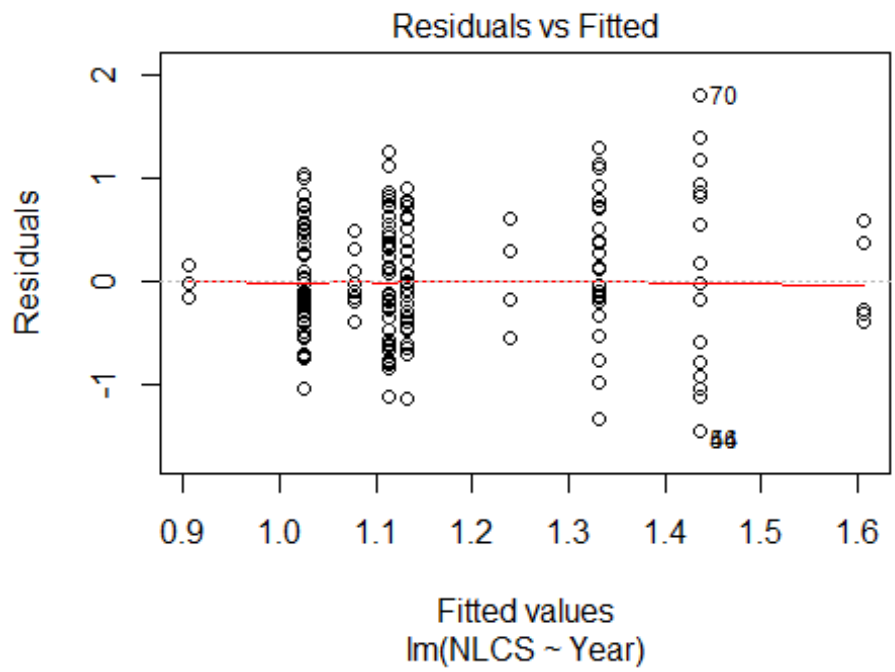
```
## 4224   0030604540 3.887 1996     1300      1     2.862
## 4228   0030604722 4.693 1996     1300      1     3.654
## 4231  15844367099 4.221 1996     1300      1     3.182
## 4273   0029894165 3.853 1996     1300      1     2.828
## 4274   0029899127 3.990 1996     1300      1     2.951
## 4275   0029953780 4.085 1996     1300      1     3.046
## 4302   0001506104 4.149 1996     1300      2     3.110
## 4334   0029895156 4.197 1996     1300      1     3.222
## 4336   0029939448 4.169 1996     1300      1     3.130
## 4337   0029943141 4.032 1996     1300      1     3.007
## 4340   0030010590 3.812 1996     1300      1     2.773
## 4342  15844384256 4.939 1996     1300      1     3.900
## 4343  15844415946 4.832 1996     1300      1     3.793
## 4344  15844420283 4.984 1996     1300      1     3.945
## 4361   0029881125 4.622 1996     1300      1     3.583
## 4362   0029892278 3.564 1996     1300      1     2.539
## 4448   0029870085 4.650 1996     1300      1     3.611
## 4451   0029993728 3.738 1996     1300      1     2.699
## 4452   0030009544 3.963 1996     1300      1     2.938
## 4453  15844372440 3.856 1996     1300      1     2.881
## 4500   0029919935 3.655 1996     1300      1     2.616
## 4501   0029980441 4.225 1996     1300      1     3.186
## 4502   0029993450 4.960 1996     1300      1     3.921
## 4565   0029876473 4.673 1996     1300      1     3.634
## 4588   0029880254 3.670 1996     1300      2     2.631
## 4589   0029880651 4.195 1996     1300      2     3.156
## 4606   0029669982 3.527 1996     1300      1     2.502
## 4611   0029961719 3.731 1996     1300      1     2.692
## 4613   0029965130 3.965 1996     1300      1     3.004
## 4614   0029978023 3.696 1996     1300      1     2.657
## 4615   0029991047 4.197 1996     1300      1     3.158
## 4617  13344277364 4.978 1996     1300      1     3.939
## 4703   0030030905 4.203 1996     1300      1     3.228
## 4705   0030048731 4.354 1996     1300      1     3.329
## 4706   0030053650 3.615 1996     1300      1     2.576
## 4707   0030058657 4.087 1996     1300      1     3.112
## 4714   0030065744 3.607 1996     1300      1     2.582
## 4715  13344261952 4.256 1996     1300      1     3.217
## 4716  13344282063 4.293 1996     1300      1     3.332
## 4727   0030020590 4.565 1996     1300      2     3.540
## 4751   0030024563 4.880 1996     1300      1     3.855
## 4797   0030033699 3.977 1996     1300      1     2.938
## 4808   0030026776 5.105 1996     1300      1     4.066
## 4809   0030026934 4.448 1996     1300      1     3.473
## 4830   0030584077 4.462 1996     1300      1     3.423
## 4831   0030584078 5.024 1996     1300      1     3.985
## 4833   0030584083 3.949 1996     1300      1     2.910
## 4854   0029671310 4.397 1996     1300      1     3.372
## 4856   0030031999 4.203 1996     1300      1     3.178
## 4858   0030034731 4.289 1996     1300      1     3.314
```

```
## 4859   0030034983 4.022 1996     1300      1     3.061
## 4860   0030050396 4.336 1996     1300      1     3.297
## 4862   0030061451 3.633 1996     1300      1     2.594
## 4887   0029664992 4.109 1996     1300      1     3.070
## 4901   0029888359 4.449 1996     1300      1     3.424
## 4903   0029890229 3.831 1996     1300      1     2.806
## 4908   0029898733 3.675 1996     1300      1     2.650
## 4917   0029936764 4.426 1996     1300      1     3.387
## 4918   0029940972 4.122 1996     1300      1     3.083
## 4922   0029949784 4.498 1996     1300      1     3.459
## 4926   0029977751 3.719 1996     1300      1     2.680
## 4929   0030014157 3.965 1996     1300      1     2.940
## 4934   0030030057 3.699 1996     1300      1     2.660
## 4962   0030111466 3.479 1996     1300      1     2.504
## 4974   0030131182 3.669 1996     1300      1     2.630
## 4981   0030139470 3.603 1996     1300      1     2.564
## 5009   0030271388 3.996 1996     1300      1     2.957
## 5010   0030271392 3.996 1996     1300      1     2.957
## 5012   0030271890 3.837 1996     1300      1     2.798
## 5014   0030271999 4.242 1996     1300      1     3.203
## 5015   0030272047 4.276 1996     1300      1     3.237
## 5043  15844380040 4.349 1996     1300      1     3.310
## 5044  15844417385 4.492 1996     1300      1     3.453
## 5868   0030087710 3.738 1996     1100      2     2.699
## 6075   0030606018 3.876 1996     1100      2     2.837
## 7056   0031444409 3.689 1997     1300      1     2.663
## 7057   0031456065 3.655 1997     1300      1     2.629
## 7061   0031459980 4.014 1997     1300      1     2.974
## 7128   0344936739 3.678 1997     1300      1     2.638
## 7227   0031449456 3.566 1997     1300      2     2.540
## 7229   0031451777 3.981 1997     1300      2     2.941
## 7235   0031466811 3.818 1997     1300      2     2.856
## 9659   0030613551 4.251 1997     1300      1     3.211
## 9665   0030702123 4.929 1997     1300      1     3.889
## 9696   0030886099 3.752 1997     1300      1     2.712
## 9699   0030886602 3.741 1997     1300      1     2.701
## 9702   0242421666 3.604 1997     1300      1     2.578
## 9712   0030796646 3.547 1997     1300      1     2.507
## 9773   0030928716 3.647 1997     1300      1     2.607
## 9776   0030954870 3.673 1997     1300      1     2.647
## 9777   0030985459 3.937 1997     1300      1     2.911
## 9838   0030865245 3.667 1997     1300      2     2.641
## 9899   0030788436 3.666 1997     1300      1     2.689
## 9902   0030829387 3.623 1997     1300      1     2.583
## 9906   0030848970 3.656 1997     1300      1     2.679
## 9966   0030746636 4.625 1997     1300      1     3.599
## 9975   0346613495 3.650 1997     1300      1     2.610
## 9977   0642270732 4.213 1997     1300      1     3.173
## 10022  0030755579 3.562 1997     1300      1     2.522
## 10026  0030877659 3.567 1997     1300      1     2.541
```

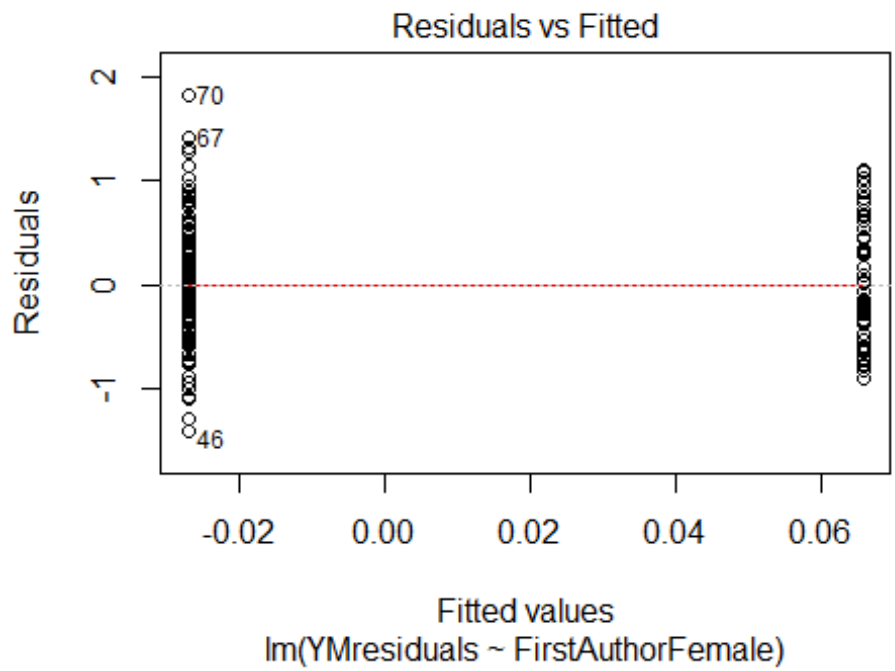
```
## 10027  0031586174 4.066 1997     1300      1     3.040
## 10074  0031440879 3.512 1997     1300      1     2.550
## 10075  0031444148 3.601 1997     1300      1     2.561
## 10083  0031472234 3.713 1997     1300      1     2.687
## 10095 16944366965 3.603 1997     1300      1     2.577
## 10096  0030970602 3.618 1997     1300      2     2.578
## 10117  0030752411 3.710 1997     1300      2     2.748
## 10314  0030687987 3.584 1997     1300      1     2.544
## 10317  0030712145 3.595 1997     1300      1     2.618
## 10337  0030729445 3.985 1997     1300      1     2.945
## 10343 18844476167 3.641 1997     1300      1     2.615
## 10344 20244377493 4.035 1997     1300      1     2.995
## 10376  0031007189 3.758 1997     1300      1     2.718
## 10461  0030970693 4.340 1997     1300      1     3.300
## 10462  0030982264 3.775 1997     1300      1     2.735
## 10515  0030890721 3.600 1997     1300      1     2.560
## 10519  0030963439 3.978 1997     1300      1     2.938
## 10585  0030933978 3.696 1997     1300      1     2.656
## 10616  0030944985 4.736 1997     1300      1     3.696
## 10618  0030970775 3.714 1997     1300      1     2.674
## 10620  0031000884 3.817 1997     1300      1     2.777
## 10679  0030949875 3.946 1997     1300      1     2.906
## 10685  0030614893 4.097 1997     1300      1     3.057
## 10688  0031048716 3.922 1997     1300      1     2.896
## 10712  0030893115 4.029 1997     1300      1     3.052
## 10739  0345877774 3.589 1997     1300      1     2.549
## 10740  0345877775 3.865 1997     1300      1     2.825
## 10785  0031285250 3.781 1997     1300      1     2.741
## 10802  0031015075 3.961 1997     1300      1     2.935
## 10839  0030994017 4.009 1997     1300      1     2.969
## 10876  0030715563 3.879 1997     1300      1     2.853
## 10881  0030800831 3.627 1997     1300      1     2.587
## 10884  0030866897 3.700 1997     1300      1     2.723
## 10958  0031587821 3.699 1997     1300      1     2.737
## 10959  0031587826 3.614 1997     1300      1     2.574
## 10961  0031587830 3.650 1997     1300      1     2.610
## 10974 16944364768 3.665 1997     1300      1     2.625
## 11042  0031022694 3.793 1997     1300      2     2.753
## 11356  0032446607 3.744 1998     1300      1     2.531
## 12409  0031720906 3.753 1998     1300      2     2.555
## 13176  0032540267 3.749 1998     1300      1     2.536
## 15773  0032953920 3.847 1999     1300      2     2.519
## 15917  0032976690 3.859 1999     1300      2     2.517
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
```

```
##     Min      1Q  Median      3Q     Max
## -1.3769 -0.3775  0.0286  0.4036  4.2845
##
## Coefficients:
##                   Estimate Std. Error t value Pr(>|t|)
## (Intercept)        1.03553    0.03652   28.36  < 2e-16 ***
## LastAuthorFemale1 -0.06601    0.00892   -7.40  1.4e-13 ***
## Year1997           0.00172    0.05155    0.03  0.97337
## Year1998           0.17354    0.05097    3.40  0.00066 ***
## Year1999           0.30297    0.04275    7.09  1.4e-12 ***
## Year2000           0.28381    0.04320    6.57  5.1e-11 ***
## Year2001           0.34133    0.04371    7.81  6.0e-15 ***
## Year2002           0.22480    0.04126    5.45  5.1e-08 ***
## Year2003           0.25205    0.04125    6.11  1.0e-09 ***
## Year2004           0.25315    0.04132    6.13  9.1e-10 ***
## Year2005           0.24565    0.04125    5.95  2.6e-09 ***
## Year2006           0.27073    0.04041    6.70  2.1e-11 ***
## Year2007           0.24897    0.03918    6.35  2.1e-10 ***
## Year2008           0.26323    0.03929    6.70  2.1e-11 ***
## Year2009           0.27263    0.03848    7.09  1.4e-12 ***
## Year2010           0.28481    0.03794    7.51  6.2e-14 ***
## Year2011           0.26382    0.03754    7.03  2.2e-12 ***
## Year2012           0.23289    0.03721    6.26  3.9e-10 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.578
## Multiple R-squared:  0.0179, Adjusted R-squared:  0.0174
## Convergence in 22 IRWLS iterations
##
## Robustness weights:
##  138 observations
c(1,2,4,5,6,7,9,33,34,67,495,497,500,527,528,529,530,531,532,541,561,563,592,
600,601,602,622,624,645,646,657,658,664,666,681,683,717,718,720,736,737,739,7
41,748,749,758,760,764,766,783,784,785,808,809,810,822,838,839,840,841,843,84
4,845,853,894,898,899,926,927,957,971,986,988,989,1031,1032,1034,1039,1040,10
44,1058,1079,1084,1085,1093,1094,1095,1105,1106,1107,1108,1109,1121,1122,1123
,1127,1128,1130,1135,1185,1186,1188,1189,1190,1199,1200,1437,1555,1624,1627,2
105,2108,2125,2167,2259,2263,2288,2331,2441,2443,2462,2508,2509,2539,2587,258
9,2617,2622,2624,2638,2648,2672,2680,2700,2721,2726,2772,2819)
##   are outliers with |weight| = 0 ( < 3.5e-06);
##  2422 weights are ~= 1. The remaining 25721 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.000   0.854   0.950   0.891   0.986   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          3.54e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
```

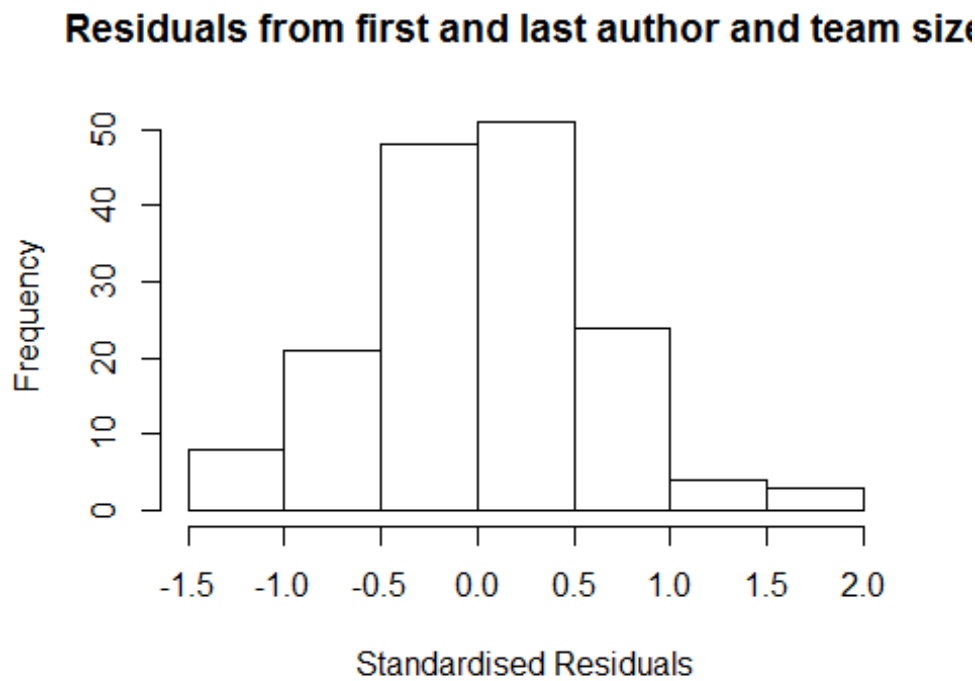
```
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Sample size for the above analysis:  28281"
## [1] ""
## [1] ""
## [1] "####################################"
## [1] "Analysis of AJSC 1301"
## [1] "####################################"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 2004 2005 2006 2007 2008 2009 2010 2011 2012
##   14   11    4   13   31   37   49   60   77
##
## 2004 2005 2006 2007 2008 2009 2010 2011 2012
##   10    5    3    5   19   29   31   45   49
##
## 2004 2005 2006 2007 2008 2009 2010 2011 2012
##    7    5    3    4   16   24   26   38   36
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
##  Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 32, df = 8, p-value = 1e-04
```




```
##
##  Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 0.27, df = 1, p-value = 0.6
```

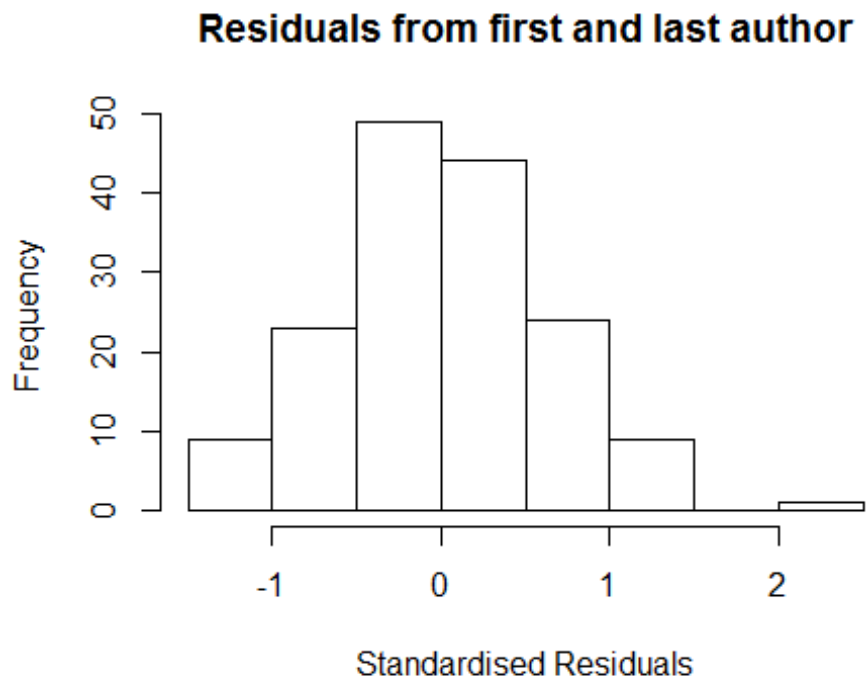



```
## [1] "Female first author team size 2018 geometric mean: 3.94043476191405"
## [1] "Male first author team size 2018 geometric mean: 4.07563698803087"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 1500, p-value = 0.7
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 3.70390539698511"
## [1] "Male last author team size 2018 geometric mean: 4.17422485688719"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 1300, p-value = 0.8
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##                     GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale  1.668  1           1.292
## LastAuthorFemale   1.668  1           1.291
## UniqueAuthors      7.075  4           1.277
## Year              11.774  8           1.167
```



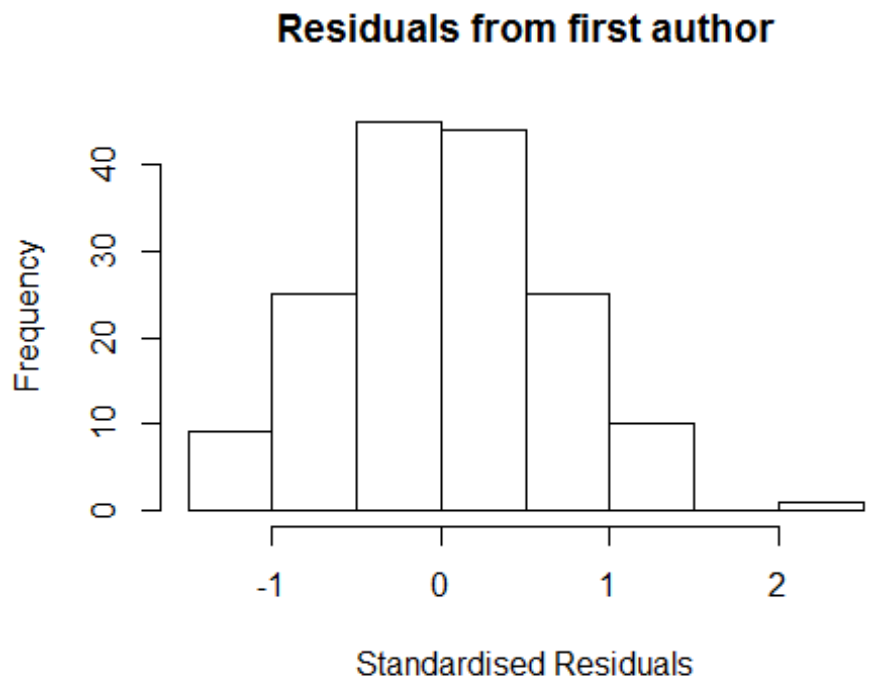

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
UniqueAuthors +
##     Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##     k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.2895 -0.3755  0.0254  0.3828  1.7628
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)          0.8043     0.1827    4.40  2.1e-05 ***
## FirstAuthorFemale1   0.1859     0.1291    1.44   0.1520
## LastAuthorFemale1   -0.2415     0.1457   -1.66   0.0997 .
## UniqueAuthors2       0.3387     0.1385    2.45   0.0157 *
## UniqueAuthors3       0.2997     0.1938    1.55   0.1242
## UniqueAuthors4       0.1605     0.2685    0.60   0.5509
## UniqueAuthors5       0.4723     0.1574    3.00   0.0032 **
## Year2005             0.0312     0.1967    0.16   0.8741
## Year2006            -0.1124     0.1689   -0.67   0.5068
## Year2007             0.6798     0.2562    2.65   0.0089 **
```

```
## Year2008             0.1986     0.4092    0.49   0.6281
## Year2009             0.1465     0.2482    0.59   0.5558
## Year2010            -0.0323     0.1855   -0.17   0.8621
## Year2011            -0.0959     0.1766   -0.54   0.5878
## Year2012             0.0991     0.1882    0.53   0.5992
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.605
## Multiple R-squared:  0.134,  Adjusted R-squared:  0.0494
## Convergence in 21 IRWLS iterations
##
## Robustness weights:
##  12 weights are ~= 1. The remaining 147 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.376   0.870   0.959   0.909   0.987   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          6.29e-04          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.348  1           1.161
## LastAuthorFemale  1.472  1           1.213
## Year              1.957  8           1.043
```



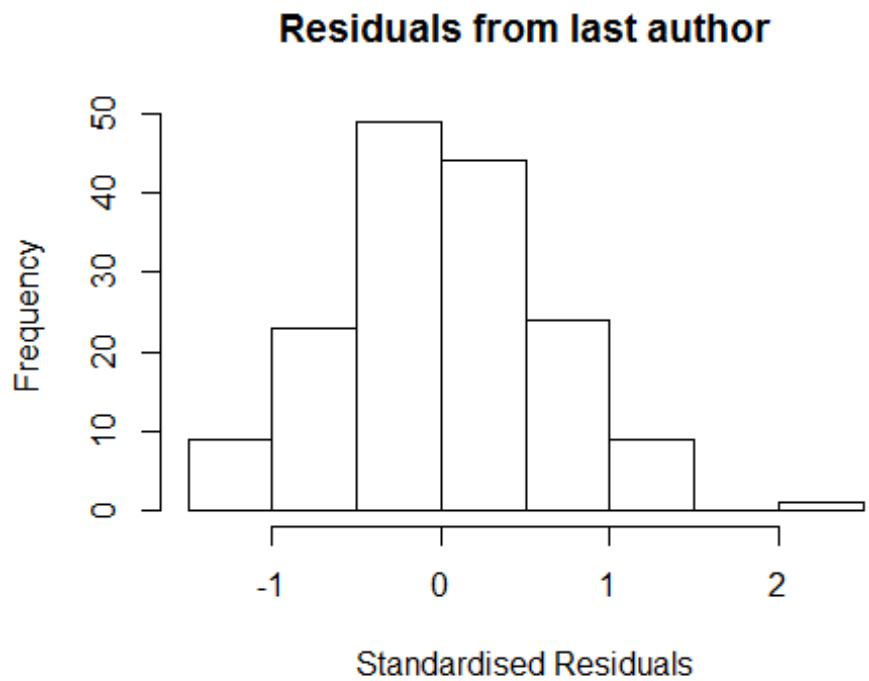

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.20431 -0.37951 -0.00822  0.43252  2.03369
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)           1.0461     0.1250    8.37    4e-14 ***
## FirstAuthorFemale1    0.2263     0.1264    1.79  0.07543 .
## LastAuthorFemale1    -0.2889     0.1415   -2.04  0.04299 *
## Year2005              0.1607     0.2066    0.78  0.43775
## Year2006             -0.1415     0.1434   -0.99  0.32534
## Year2007              0.7803     0.2007    3.89  0.00015 ***
## Year2008              0.1582     0.3823    0.41  0.67961
## Year2009              0.1550     0.2136    0.73  0.46934
## Year2010              0.0257     0.1675    0.15  0.87823
## Year2011             -0.0379     0.1531   -0.25  0.80497
## Year2012              0.1555     0.1732    0.90  0.37067
## ---
```

```
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.632
## Multiple R-squared:  0.0781, Adjusted R-squared:  0.0158
## Convergence in 19 IRWLS iterations
##
## Robustness weights:
##  13 weights are ~= 1. The remaining 146 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.279   0.860   0.955   0.911   0.985   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          6.29e-04          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.314  1           1.146
## Year              1.314  8           1.017
```




```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##       Min        1Q    Median        3Q       Max
## -1.174942 -0.384346  0.000058  0.437098  2.118794
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.05221    0.12257    8.58  1.1e-14 ***
## FirstAuthorFemale1  0.18379    0.13028    1.41    0.160
## Year2005            0.11016    0.21630    0.51    0.611
## Year2006           -0.14757    0.14123   -1.04    0.298
## Year2007            0.62931    0.25224    2.49    0.014 *
## Year2008            0.06700    0.39672    0.17    0.866
## Year2009            0.12274    0.20906    0.59    0.558
## Year2010            0.00825    0.16285    0.05    0.960
## Year2011           -0.08415    0.15176   -0.55    0.580
## Year2012            0.08766    0.16803    0.52    0.603
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

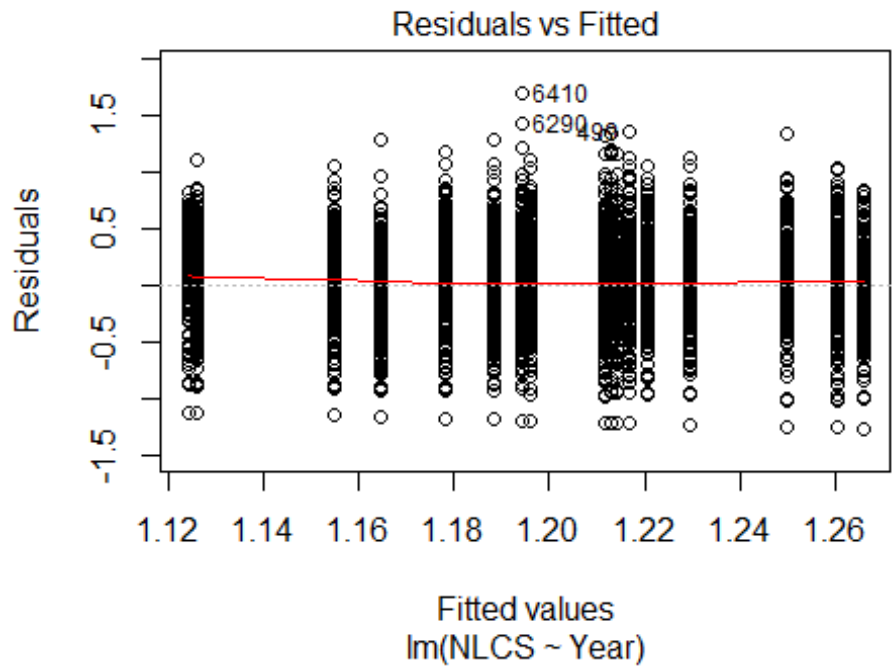
```
##
## Robust residual standard error: 0.638
## Multiple R-squared:  0.0518, Adjusted R-squared:  -0.00552
## Convergence in 19 IRWLS iterations
##
## Robustness weights:
##  12 weights are ~= 1. The remaining 147 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.247   0.865   0.953   0.910   0.984   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          6.29e-04          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##                   GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.482  1           1.217
## Year             1.482  8           1.025
```



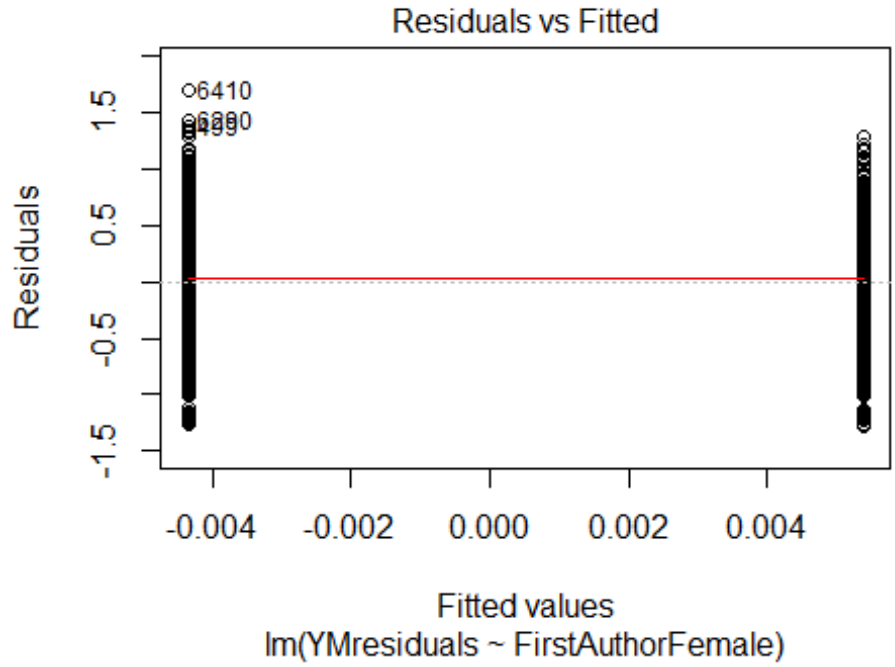

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.2990 -0.3971 -0.0138  0.4285  1.9939
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.0791     0.1127    9.57  < 2e-16 ***
## LastAuthorFemale1  -0.2422     0.1424   -1.70  0.09096 .
## Year2005            0.2088     0.2209    0.95  0.34617
## Year2006           -0.1745     0.1328   -1.31  0.19081
## Year2007            0.7239     0.2020    3.58  0.00046 ***
## Year2008            0.1650     0.4364    0.38  0.70585
## Year2009            0.2199     0.2080    1.06  0.29208
## Year2010            0.0589     0.1546    0.38  0.70368
## Year2011           -0.0233     0.1467   -0.16  0.87406
## Year2012            0.1718     0.1613    1.06  0.28864
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
##
## Robust residual standard error: 0.634
## Multiple R-squared:  0.0581, Adjusted R-squared:  0.00116
## Convergence in 21 IRWLS iterations
##
## Robustness weights:
##  17 weights are ~= 1. The remaining 142 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.302   0.878   0.946   0.907   0.986   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          6.29e-04          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Sample size for the above analysis:  159"
## [1] ""
## [1] ""
## [1] "####################################"
## [1] "Analysis of AJSC 1302"
## [1] "####################################"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  332  282  302  275  303  335  340  270  305  351  385  428  444  418  460
## 2011 2012
##  466  564
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  238  168  219  198  159  138  260  197  231  274  292  323  330  328  335
## 2011 2012
##  339  412
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  220  155  206  176  143  121  228  172  201  244  262  287  285  291  291
## 2011 2012
##  309  368
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
```

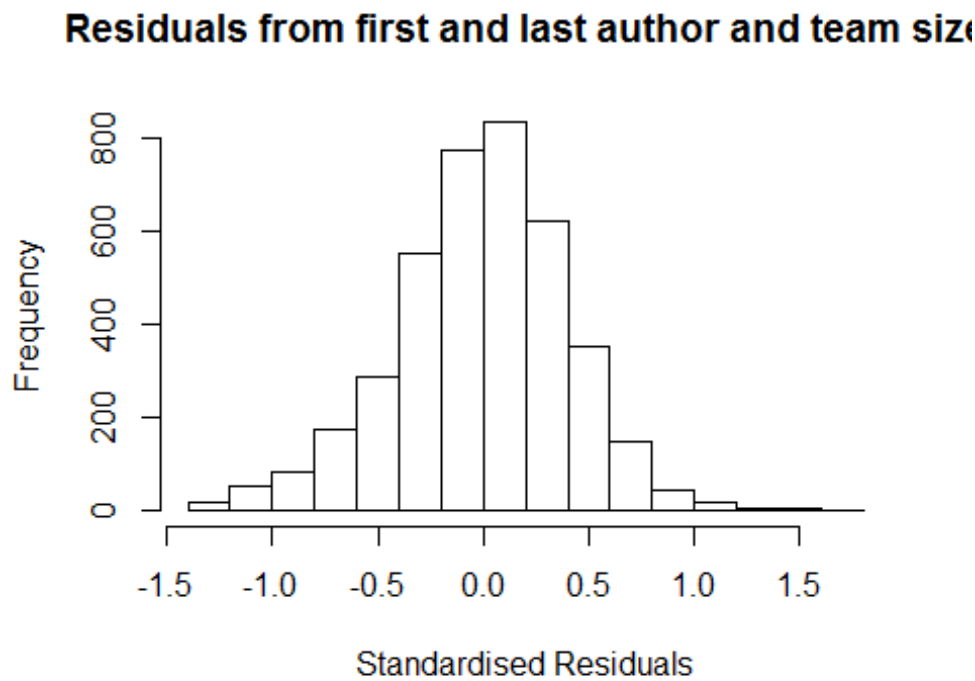
```
##
##  Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 67, df = 16, p-value = 3e-08
```




```
##
##  Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 3.5, df = 1, p-value = 0.06
```

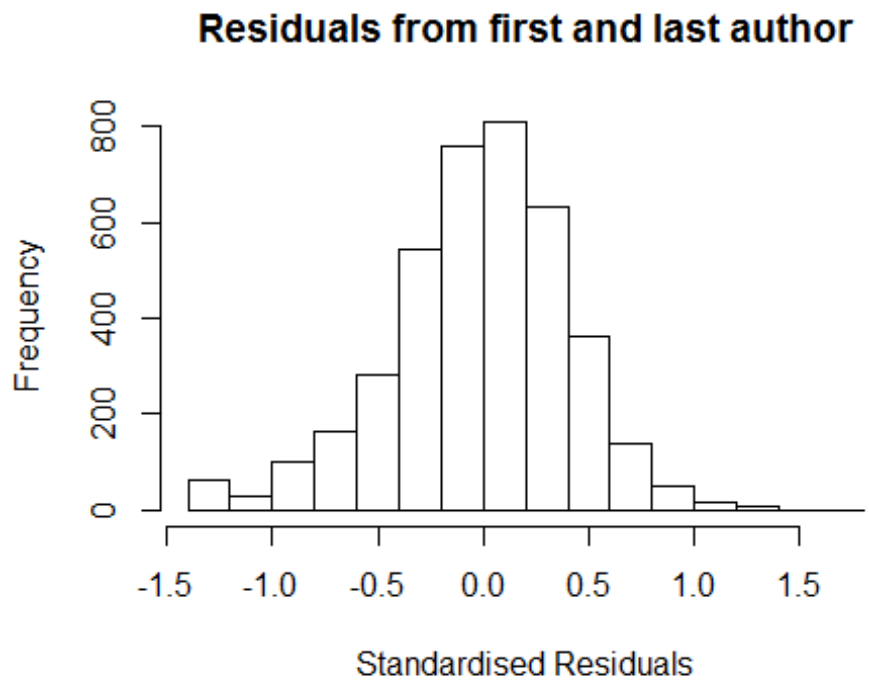



```
## [1] "Female first author team size 2018 geometric mean: 5.4209222658251"
## [1] "Male first author team size 2018 geometric mean: 5.49338633227761"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 11000, p-value = 0.7
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 5.48481049458571"
## [1] "Male last author team size 2018 geometric mean: 5.44611392165555"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 11000, p-value = 0.6
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.041  1           1.020
## LastAuthorFemale  1.045  1           1.022
## UniqueAuthors     1.127  4           1.015
## Year              1.156 16           1.005
```



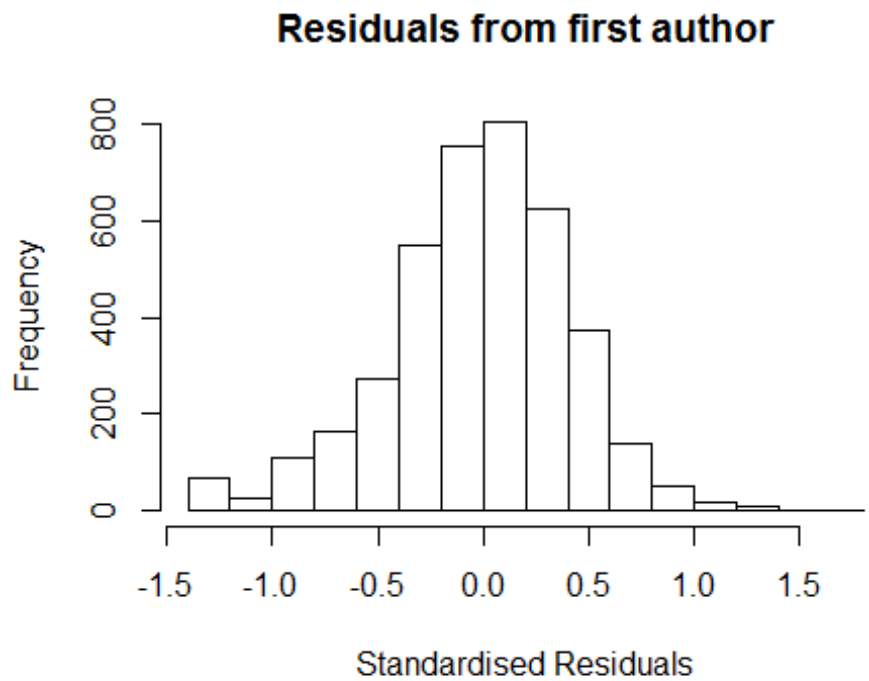

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
UniqueAuthors +
##     Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##     k.max = 1000))
##  \--> method = "MM"
## Residuals:
##    Min     1Q Median     3Q    Max
## -1.397 -0.262  0.013  0.253  1.602
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.05506    0.04023   26.22  < 2e-16 ***
## FirstAuthorFemale1 -0.00661    0.01299   -0.51  0.61091
## LastAuthorFemale1   0.03152    0.01386    2.27  0.02296 *
## UniqueAuthors2      0.15023    0.02962    5.07  4.1e-07 ***
## UniqueAuthors3      0.23160    0.02918    7.94  2.7e-15 ***
## UniqueAuthors4      0.27002    0.02980    9.06  < 2e-16 ***
## UniqueAuthors5      0.34017    0.02690   12.64  < 2e-16 ***
## Year1997           -0.04339    0.04726   -0.92  0.35863
## Year1998           -0.10926    0.04696   -2.33  0.02003 *
## Year1999           -0.10645    0.04491   -2.37  0.01781 *
```

```
## Year2000           -0.02272    0.04774   -0.48  0.63411
## Year2001           -0.08157    0.05806   -1.40  0.16015
## Year2002           -0.03205    0.04113   -0.78  0.43590
## Year2003           -0.06077    0.04507   -1.35  0.17763
## Year2004           -0.01979    0.04129   -0.48  0.63164
## Year2005           -0.07281    0.04087   -1.78  0.07489 .
## Year2006           -0.07010    0.03931   -1.78  0.07460 .
## Year2007           -0.10734    0.03929   -2.73  0.00633 **
## Year2008           -0.14613    0.03992   -3.66  0.00025 ***
## Year2009           -0.07254    0.04042   -1.79  0.07282 .
## Year2010           -0.08361    0.03791   -2.21  0.02750 *
## Year2011           -0.11502    0.03951   -2.91  0.00362 **
## Year2012           -0.13054    0.03775   -3.46  0.00055 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.379
## Multiple R-squared:  0.0767, Adjusted R-squared:  0.0715
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
##  336 weights are ~= 1. The remaining 3623 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0349  0.8660  0.9490  0.8950  0.9860  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          2.53e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.031  1           1.016
## LastAuthorFemale  1.039  1           1.019
## Year              1.041 16           1.001
```



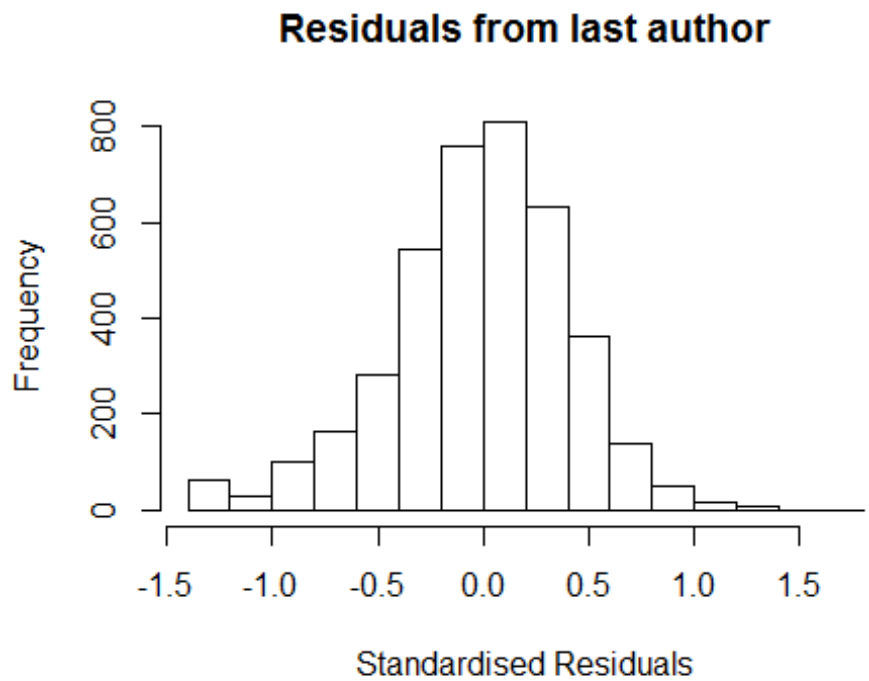

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.29408 -0.26097  0.00976  0.26123  1.68298
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.26179    0.03236   38.99   <2e-16 ***
## FirstAuthorFemale1  0.00288    0.01332    0.22    0.829
## LastAuthorFemale1   0.01914    0.01420    1.35    0.178
## Year1997           -0.06104    0.04828   -1.26    0.206
## Year1998           -0.10777    0.04699   -2.29    0.022 *
## Year1999           -0.09888    0.04560   -2.17    0.030 *
## Year2000            0.01026    0.04808    0.21    0.831
## Year2001           -0.05608    0.05868   -0.96    0.339
## Year2002           -0.02105    0.04222   -0.50    0.618
## Year2003           -0.02554    0.04606   -0.55    0.579
## Year2004            0.00667    0.04087    0.16    0.870
## Year2005           -0.02357    0.04072   -0.58    0.563
```

```
## Year2006           -0.03378    0.03985   -0.85    0.397
## Year2007           -0.07287    0.04019   -1.81    0.070 .
## Year2008           -0.10166    0.03970   -2.56    0.010 *
## Year2009           -0.02924    0.04064   -0.72    0.472
## Year2010           -0.03100    0.03772   -0.82    0.411
## Year2011           -0.06278    0.03873   -1.62    0.105
## Year2012           -0.06570    0.03734   -1.76    0.079 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.388
## Multiple R-squared:  0.00773,    Adjusted R-squared:  0.0032
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
##  335 weights are ~= 1. The remaining 3624 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0208  0.8670  0.9500  0.8930  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          2.53e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.015  1           1.008
## Year              1.015 16           1.000
```




```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.2815 -0.2642  0.0101  0.2603  1.6791
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)          1.26658    0.03221   39.32   <2e-16 ***
## FirstAuthorFemale1  0.00627    0.01326    0.47    0.636
## Year1997           -0.06032    0.04831   -1.25    0.212
## Year1998           -0.10696    0.04700   -2.28    0.023 *
## Year1999           -0.09837    0.04555   -2.16    0.031 *
## Year2000            0.00869    0.04806    0.18    0.857
## Year2001           -0.05662    0.05872   -0.96    0.335
## Year2002           -0.02022    0.04213   -0.48    0.631
## Year2003           -0.02465    0.04601   -0.54    0.592
## Year2004            0.00725    0.04084    0.18    0.859
## Year2005           -0.02298    0.04072   -0.56    0.573
## Year2006           -0.03445    0.03980   -0.87    0.387
```

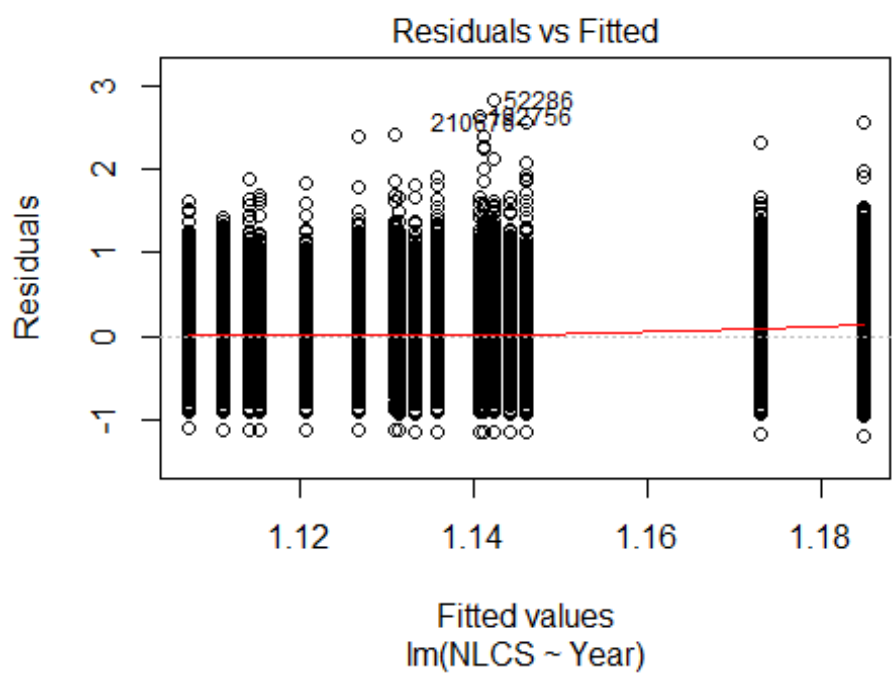
```
## Year2007           -0.07260    0.04014   -1.81    0.071 .
## Year2008           -0.10199    0.03966   -2.57    0.010 *
## Year2009           -0.02911    0.04061   -0.72    0.474
## Year2010           -0.03207    0.03765   -0.85    0.394
## Year2011           -0.06363    0.03869   -1.64    0.100
## Year2012           -0.06578    0.03728   -1.76    0.078 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.388
## Multiple R-squared:  0.0072, Adjusted R-squared:  0.00292
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
##  321 weights are ~= 1. The remaining 3638 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0219  0.8670  0.9510  0.8930  0.9860  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          2.53e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##                   GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.023  1           1.011
## Year             1.023 16           1.001
```



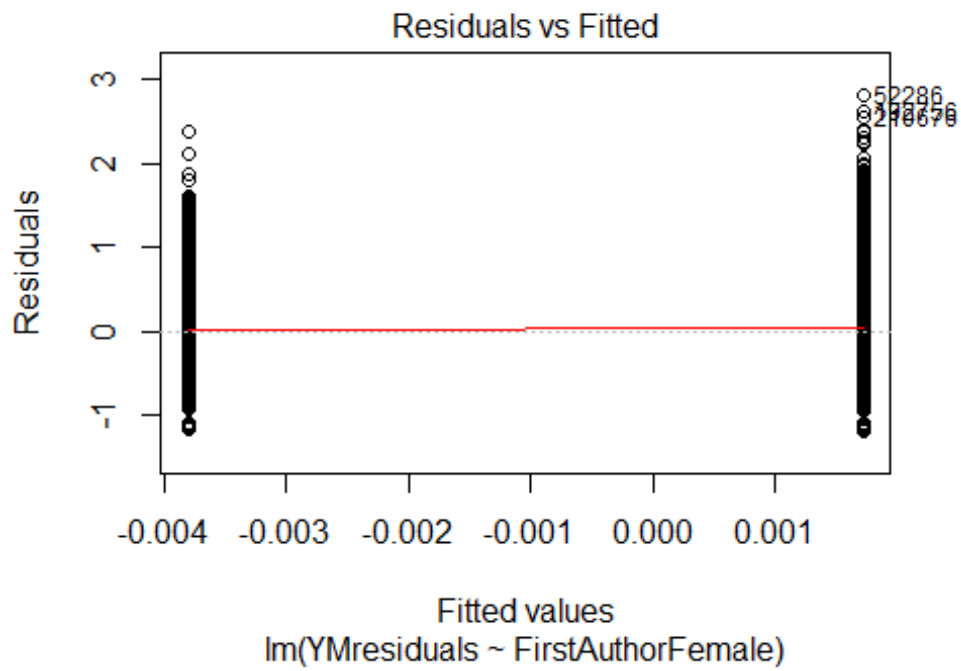

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.2931 -0.2603  0.0101  0.2616  1.6819
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.26292    0.03185   39.66   <2e-16 ***
## LastAuthorFemale1   0.01972    0.01412    1.40    0.163
## Year1997           -0.06127    0.04829   -1.27    0.205
## Year1998           -0.10794    0.04698   -2.30    0.022 *
## Year1999           -0.09908    0.04556   -2.17    0.030 *
## Year2000            0.01042    0.04806    0.22    0.828
## Year2001           -0.05603    0.05869   -0.95    0.340
## Year2002           -0.02122    0.04222   -0.50    0.615
## Year2003           -0.02544    0.04604   -0.55    0.581
## Year2004            0.00665    0.04087    0.16    0.871
## Year2005           -0.02356    0.04072   -0.58    0.563
## Year2006           -0.03392    0.03984   -0.85    0.395
```

```
## Year2007          -0.07286    0.04020   -1.81    0.070 .
## Year2008          -0.10148    0.03969   -2.56    0.011 *
## Year2009          -0.02910    0.04066   -0.72    0.474
## Year2010          -0.03093    0.03772   -0.82    0.412
## Year2011          -0.06284    0.03873   -1.62    0.105
## Year2012          -0.06554    0.03734   -1.76    0.079 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.388
## Multiple R-squared:  0.00772,    Adjusted R-squared:  0.00344
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
##  331 weights are ~= 1. The remaining 3628 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.021   0.867   0.950   0.893   0.985   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          2.53e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Sample size for the above analysis:  3959"
## [1] ""
## [1] ""
## [1] "####################################"
## [1] "Analysis of AJSC 1303"
## [1] "####################################"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
##  1996  1997  1998  1999  2000  2001  2002  2003  2004  2005  2006  2007
## 18428 15695 13893 11177 12021 10612 12399 10949 11780 11498 11621 11420
##  2008  2009  2010  2011  2012
## 11661 12214 12137 12441 11513
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 8766 8130 7578 7183 6605 5186 7984 6859 7286 7123 7150 7156 7398 7654 7557
## 2011 2012
```

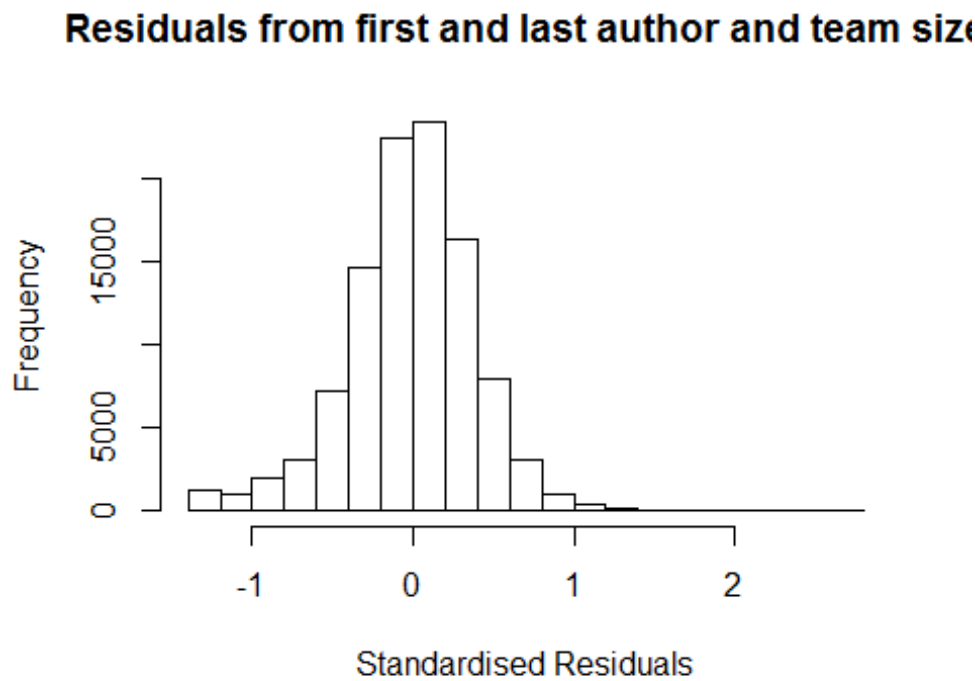
```
## 7815 7342
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 7362 6852 6421 6003 5565 4339 6668 5774 6073 5880 5888 5885 6120 6327 6193
## 2011 2012
## 6411 6056
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
##  Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 7400, df = 16, p-value <2e-16
```




```
##
##  Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 110, df = 1, p-value <2e-16
```



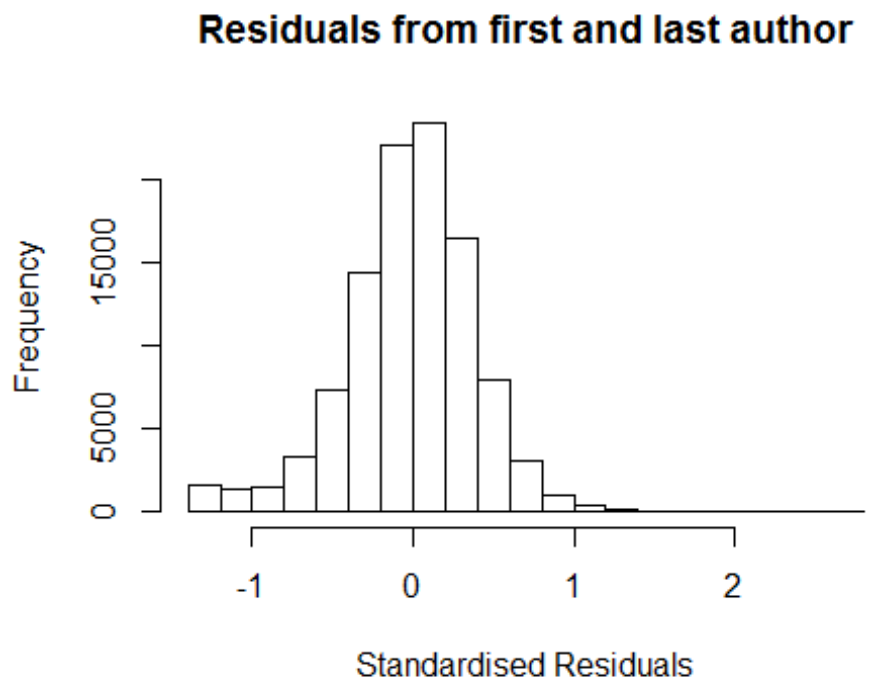

```
## [1] "Female first author team size 2018 geometric mean: 4.78301339300549"
## [1] "Male first author team size 2018 geometric mean: 4.29102125968701"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 1600000, p-value = 1e-05
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 4.52351897014074"
## [1] "Male last author team size 2018 geometric mean: 4.44268004707956"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 1100000, p-value = 0.4
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##                     GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.012  1           1.006
## LastAuthorFemale  1.009  1           1.005
## UniqueAuthors     1.032  4           1.004
## Year              1.035 16           1.001
```




```
## [1] "List of  2 outliers with residuals above  2.5"
##           ScopusId  NLCS Year OneField Fields residuals
## 52286   0031773680 3.959 1998     1303      6     2.784
## 192756 65449136284 3.759 2009     1303      7     2.640
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
UniqueAuthors +
##     Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##     k.max = 1000))
##  \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.37323 -0.23504  0.00266  0.22808  2.78366
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.01377    0.01133   89.48  < 2e-16 ***
## FirstAuthorFemale1 -0.01192    0.00242   -4.92  8.7e-07 ***
## LastAuthorFemale1  -0.01551    0.00299   -5.19  2.1e-07 ***
## UniqueAuthors2      0.23725    0.00989   23.98  < 2e-16 ***
## UniqueAuthors3      0.26901    0.00985   27.30  < 2e-16 ***
## UniqueAuthors4      0.28970    0.00988   29.33  < 2e-16 ***
## UniqueAuthors5      0.35946    0.00968   37.13  < 2e-16 ***
## Year1997           -0.01873    0.00880   -2.13    0.033 *
## Year1998           -0.07569    0.00819   -9.24  < 2e-16 ***
```

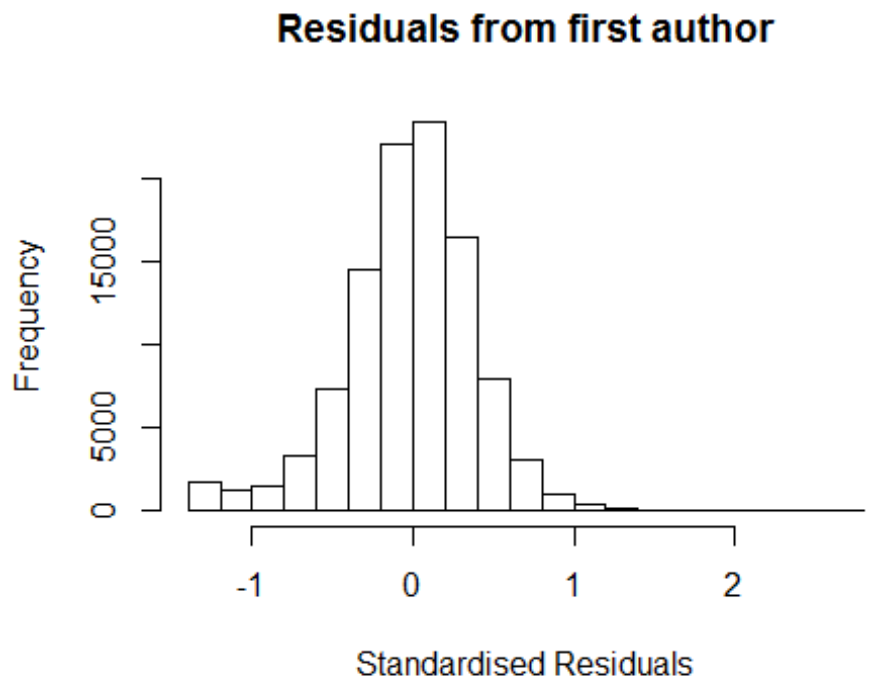
```
## Year1999           -0.13167    0.00755  -17.44  < 2e-16 ***
## Year2000           -0.13849    0.00798  -17.36  < 2e-16 ***
## Year2001           -0.15229    0.00793  -19.21  < 2e-16 ***
## Year2002           -0.16174    0.00730  -22.15  < 2e-16 ***
## Year2003           -0.18356    0.00741  -24.78  < 2e-16 ***
## Year2004           -0.17893    0.00727  -24.63  < 2e-16 ***
## Year2005           -0.18379    0.00738  -24.92  < 2e-16 ***
## Year2006           -0.19231    0.00741  -25.96  < 2e-16 ***
## Year2007           -0.16893    0.00753  -22.43  < 2e-16 ***
## Year2008           -0.16662    0.00746  -22.34  < 2e-16 ***
## Year2009           -0.16389    0.00752  -21.78  < 2e-16 ***
## Year2010           -0.16618    0.00755  -22.03  < 2e-16 ***
## Year2011           -0.16589    0.00755  -21.97  < 2e-16 ***
## Year2012           -0.17313    0.00780  -22.18  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.344
## Multiple R-squared:  0.0616, Adjusted R-squared:  0.0614
## Convergence in 18 IRWLS iterations
##
## Robustness weights:
##  39 observations
c(1530,5840,6105,6343,7901,11042,11126,19496,20540,22105,24502,36491,36492,43
140,49606,51031,53827,58501,59800,59870,68247,69185,72728,75431,82415,82663,8
7850,89177,90315,92480,93903,95083,96640,96642,97121,101331,101580,101956,102
611)
##   are outliers with |weight| = 0 ( < 9.6e-07);
##  8743 weights are ~= 1. The remaining 95035 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.000   0.863   0.951   0.887   0.986   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          9.63e-07          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##                    GVIF Df GVIF^(1/(2*Df))
```

```
## FirstAuthorFemale 1.007  1           1.004
## LastAuthorFemale  1.009  1           1.005
## Year              1.007 16           1.000
```



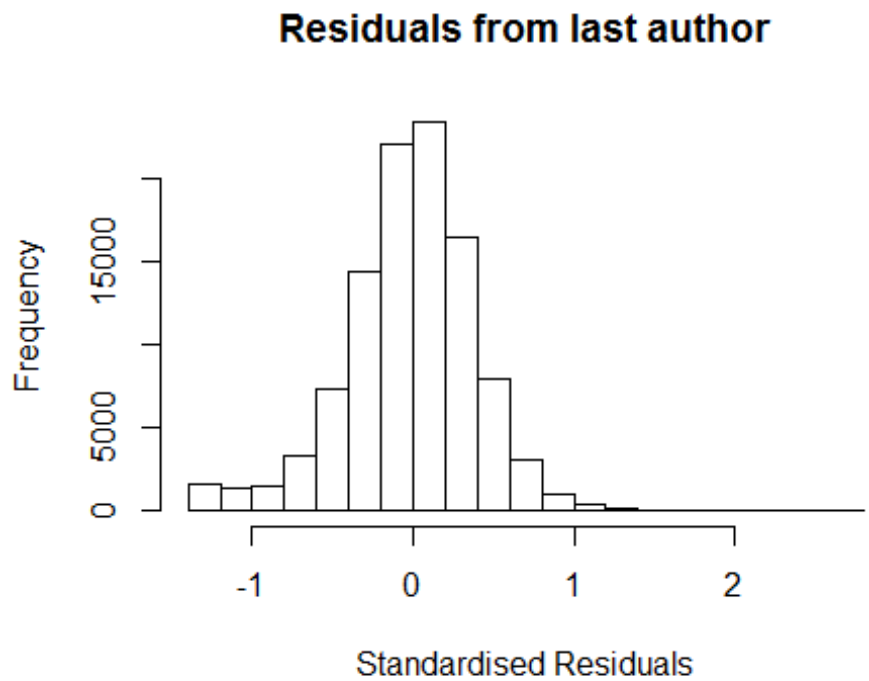

```
## [1] "List of  2 outliers with residuals above  2.5"
##           ScopusId  NLCS Year OneField Fields residuals
## 52286   0031773680 3.959 1998     1303      6     2.744
## 192756 65449136284 3.759 2009     1303      7     2.611
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.28914 -0.23928  0.00344  0.22844  2.74440
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.28914    0.00619  208.21  < 2e-16 ***
## FirstAuthorFemale1 -0.00208    0.00245   -0.85     0.40
## LastAuthorFemale1  -0.01526    0.00302   -5.05  4.5e-07 ***
## Year1997           -0.01332    0.00889   -1.50     0.13
## Year1998           -0.07454    0.00829   -8.99  < 2e-16 ***
## Year1999           -0.12968    0.00764  -16.98  < 2e-16 ***
## Year2000           -0.13348    0.00808  -16.52  < 2e-16 ***
```

```
## Year2001           -0.14250    0.00804  -17.72  < 2e-16 ***
## Year2002           -0.15010    0.00739  -20.32  < 2e-16 ***
## Year2003           -0.16808    0.00750  -22.41  < 2e-16 ***
## Year2004           -0.16272    0.00737  -22.09  < 2e-16 ***
## Year2005           -0.16758    0.00748  -22.40  < 2e-16 ***
## Year2006           -0.17326    0.00750  -23.12  < 2e-16 ***
## Year2007           -0.14909    0.00763  -19.54  < 2e-16 ***
## Year2008           -0.14465    0.00756  -19.13  < 2e-16 ***
## Year2009           -0.14078    0.00762  -18.47  < 2e-16 ***
## Year2010           -0.14100    0.00765  -18.44  < 2e-16 ***
## Year2011           -0.14076    0.00765  -18.40  < 2e-16 ***
## Year2012           -0.14670    0.00792  -18.52  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.349
## Multiple R-squared:  0.0195, Adjusted R-squared:  0.0193
## Convergence in 15 IRWLS iterations
##
## Robustness weights:
##  30 observations
c(1530,5840,6105,7901,19496,20540,22105,36491,36492,43140,44801,51031,59870,6
8247,69185,72728,80702,82663,87850,89177,90315,92480,93903,95083,96640,97121,
101331,101580,101956,102611)
##   are outliers with |weight| = 0 ( < 9.6e-07);
##  8924 weights are ~= 1. The remaining 94863 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.000   0.862   0.950   0.885   0.986   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          9.63e-07          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.003  1           1.002
## Year              1.003 16           1.000
```




```
## [1] "List of  2 outliers with residuals above  2.5"
##           ScopusId  NLCS Year OneField Fields residuals
## 52286   0031773680 3.959 1998     1303      6     2.744
## 192756 65449136284 3.759 2009     1303      7     2.611
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.28730 -0.23936  0.00318  0.22840  2.74640
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.28730    0.00619  208.04   <2e-16 ***
## FirstAuthorFemale1 -0.00324    0.00245   -1.32     0.19
## Year1997           -0.01338    0.00889   -1.50     0.13
## Year1998           -0.07470    0.00829   -9.01   <2e-16 ***
## Year1999           -0.13007    0.00764  -17.03   <2e-16 ***
## Year2000           -0.13365    0.00808  -16.54   <2e-16 ***
## Year2001           -0.14315    0.00804  -17.80   <2e-16 ***
## Year2002           -0.15054    0.00739  -20.37   <2e-16 ***
## Year2003           -0.16849    0.00750  -22.46   <2e-16 ***
## Year2004           -0.16325    0.00737  -22.16   <2e-16 ***
## Year2005           -0.16796    0.00748  -22.44   <2e-16 ***
```

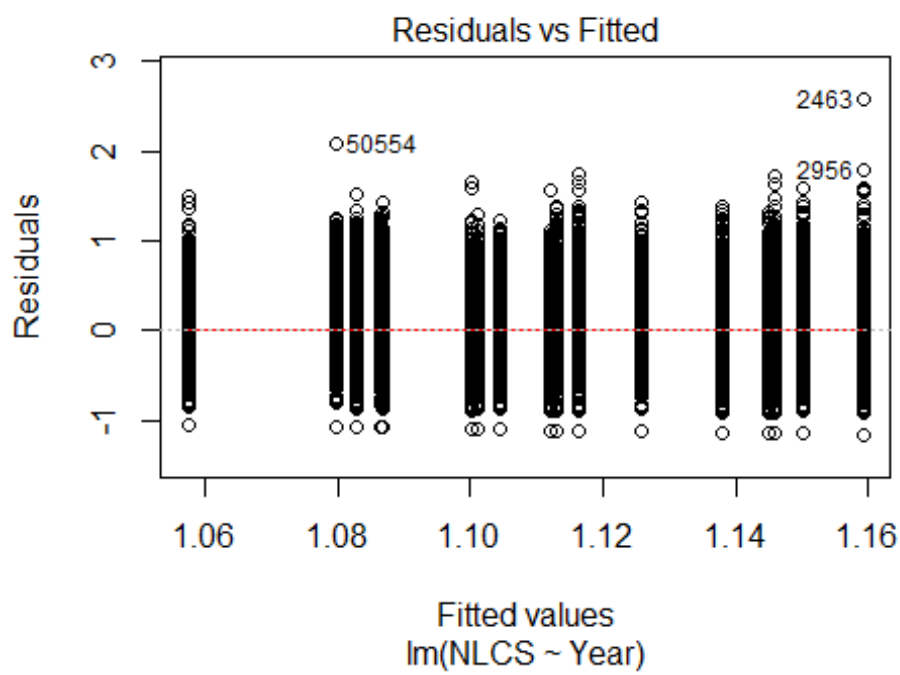
```
## Year2006           -0.17395    0.00750  -23.20   <2e-16 ***
## Year2007           -0.14982    0.00763  -19.63   <2e-16 ***
## Year2008           -0.14555    0.00756  -19.25   <2e-16 ***
## Year2009           -0.14146    0.00762  -18.56   <2e-16 ***
## Year2010           -0.14167    0.00765  -18.52   <2e-16 ***
## Year2011           -0.14178    0.00765  -18.54   <2e-16 ***
## Year2012           -0.14749    0.00792  -18.62   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.349
## Multiple R-squared:  0.0192, Adjusted R-squared:  0.0191
## Convergence in 15 IRWLS iterations
##
## Robustness weights:
##  30 observations
c(1530,5840,6105,7901,19496,20540,22105,36491,36492,43140,44801,51031,59870,6
8247,69185,72728,80702,82663,87850,89177,90315,92480,93903,95083,96640,97121,
101331,101580,101956,102611)
##   are outliers with |weight| = 0 ( < 9.6e-07);
##  8934 weights are ~= 1. The remaining 94853 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.000   0.862   0.950   0.885   0.986   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          9.63e-07          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##                   GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.005  1           1.002
## Year             1.005 16           1.000
```



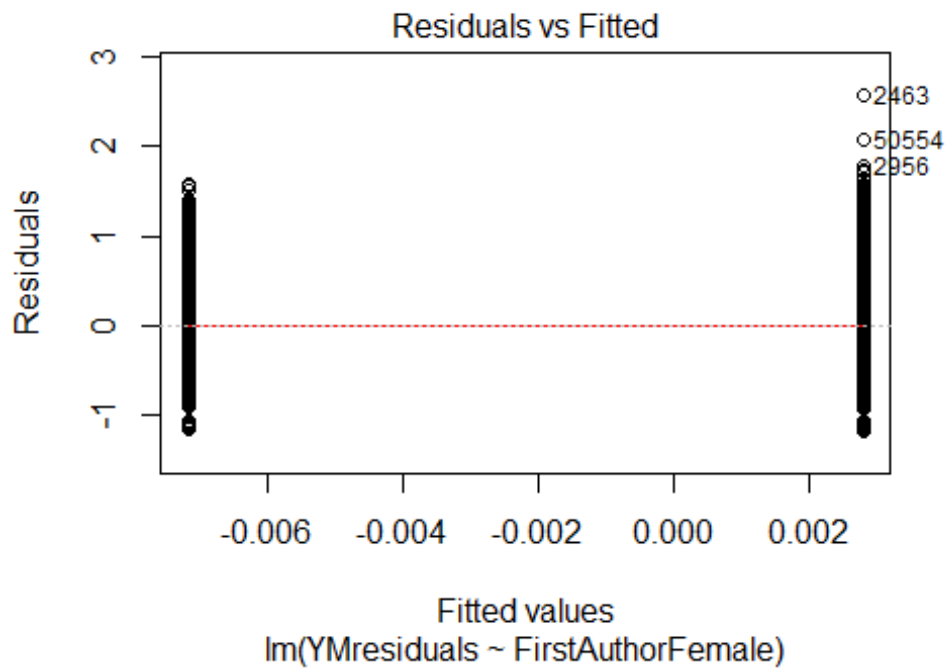

```
## [1] "List of  2 outliers with residuals above  2.5"
##           ScopusId  NLCS Year OneField Fields residuals
## 52286   0031773680 3.959 1998     1303      6     2.744
## 192756 65449136284 3.759 2009     1303      7     2.611
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.2886 -0.2394  0.0033  0.2285  2.7450
##
## Coefficients:
##                   Estimate Std. Error t value Pr(>|t|)
## (Intercept)        1.28857    0.00615  209.44  < 2e-16 ***
## LastAuthorFemale1 -0.01550    0.00302   -5.13  2.8e-07 ***
## Year1997          -0.01334    0.00889   -1.50     0.13
## Year1998          -0.07455    0.00829   -8.99  < 2e-16 ***
## Year1999          -0.12970    0.00764  -16.98  < 2e-16 ***
## Year2000          -0.13352    0.00808  -16.53  < 2e-16 ***
## Year2001          -0.14254    0.00804  -17.73  < 2e-16 ***
## Year2002          -0.15016    0.00739  -20.32  < 2e-16 ***
## Year2003          -0.16814    0.00750  -22.42  < 2e-16 ***
## Year2004          -0.16279    0.00736  -22.11  < 2e-16 ***
## Year2005          -0.16765    0.00748  -22.41  < 2e-16 ***
```

```
## Year2006          -0.17334    0.00749  -23.13  < 2e-16 ***
## Year2007          -0.14919    0.00763  -19.55  < 2e-16 ***
## Year2008          -0.14477    0.00756  -19.15  < 2e-16 ***
## Year2009          -0.14087    0.00762  -18.48  < 2e-16 ***
## Year2010          -0.14112    0.00765  -18.46  < 2e-16 ***
## Year2011          -0.14086    0.00765  -18.42  < 2e-16 ***
## Year2012          -0.14683    0.00792  -18.54  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.349
## Multiple R-squared:  0.0195, Adjusted R-squared:  0.0193
## Convergence in 15 IRWLS iterations
##
## Robustness weights:
##  30 observations
c(1530,5840,6105,7901,19496,20540,22105,36491,36492,43140,44801,51031,59870,6
8247,69185,72728,80702,82663,87850,89177,90315,92480,93903,95083,96640,97121,
101331,101580,101956,102611)
##   are outliers with |weight| = 0 ( < 9.6e-07);
##  8890 weights are ~= 1. The remaining 94897 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.000   0.862   0.950   0.885   0.986   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          9.63e-07          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Sample size for the above analysis:  103817"
## [1] ""
## [1] ""
## [1] "##################################"
## [1] "Analysis of AJSC 1304"
## [1] "##################################"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 2885 2778 2749 2712 2827 2452 2411 2415 2408 3360 2733 3794 2758 2638 2827
```

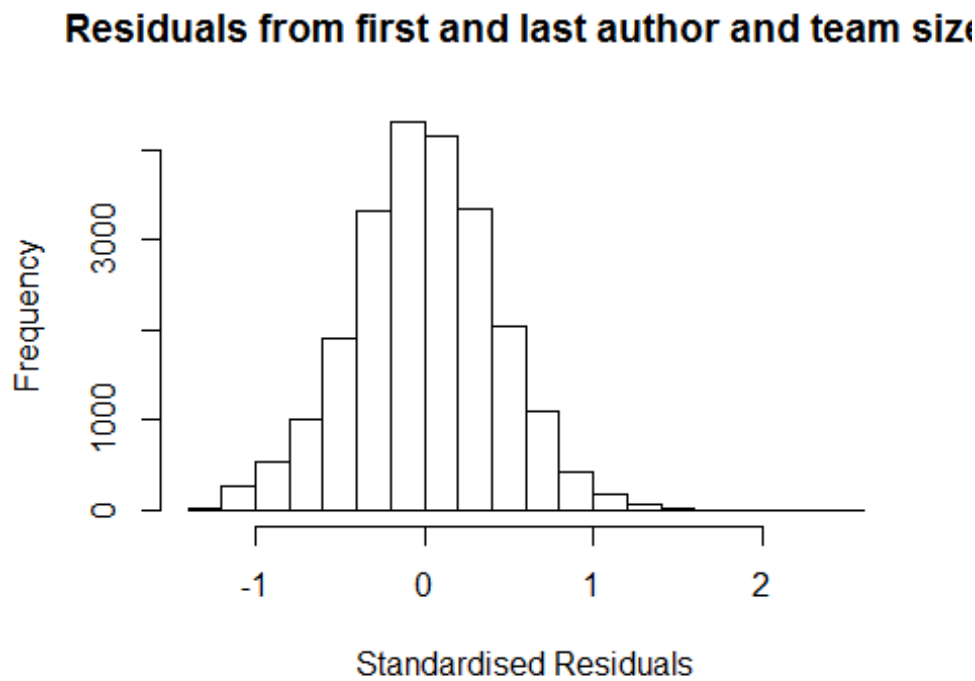
```
## 2011 2012
## 2931 2877
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 1709 1655 1631 1626 1525 1282 1496 1458 1419 1473 1636 1645 1704 1573 1758
## 2011 2012
## 1885 1816
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 1423 1390 1369 1337 1241 1090 1265 1204 1191 1209 1333 1350 1410 1309 1452
## 2011 2012
## 1572 1505
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
##  Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 130, df = 16, p-value <2e-16
```




```
##
##  Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 30, df = 1, p-value = 4e-08
```

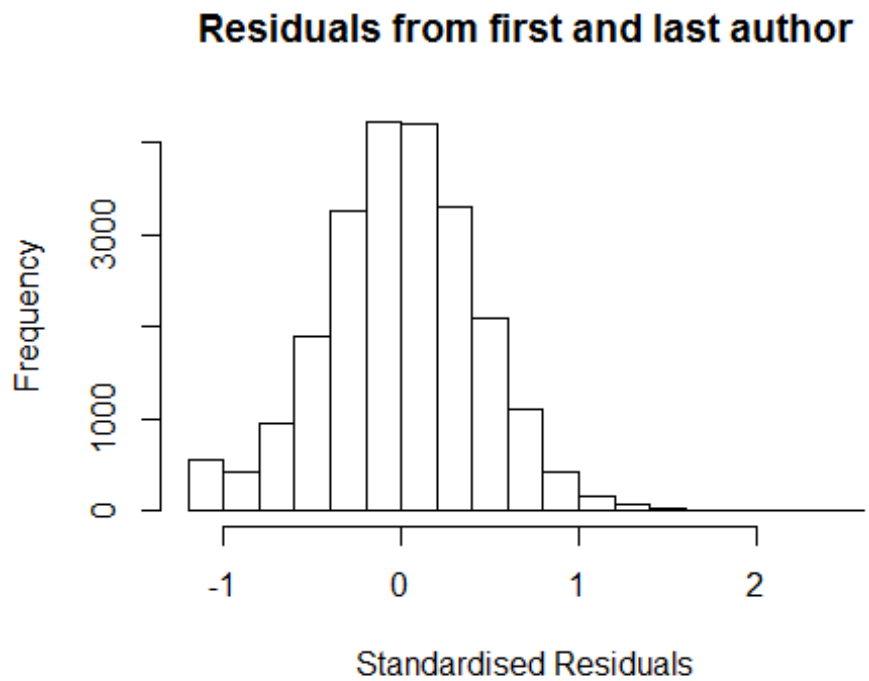



```
## [1] "Female first author team size 2018 geometric mean: 4.24003280404902"
## [1] "Male first author team size 2018 geometric mean: 4.02503475882146"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 150000, p-value = 0.4
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 4.22132304640045"
## [1] "Male last author team size 2018 geometric mean: 4.05159064771839"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 120000, p-value = 0.7
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.022  1           1.011
## LastAuthorFemale  1.013  1           1.007
## UniqueAuthors     1.055  4           1.007
## Year              1.046 16           1.001
```



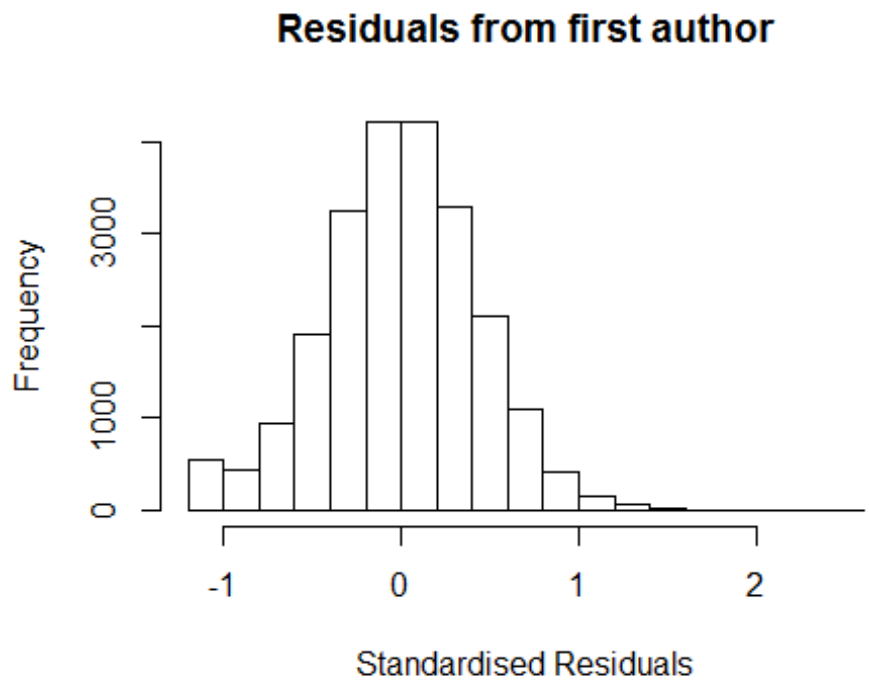

```
## [1] "List of  1 outliers with residuals above  2.5"
##        ScopusId  NLCS Year OneField Fields residuals
## 2463 0029878720 3.734 1996     1303      2     2.546
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
UniqueAuthors +
##     Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##     k.max = 1000))
##  \--> method = "MM"
## Residuals:
##       Min        1Q    Median        3Q       Max
## -1.261142 -0.276444 -0.000831  0.279212  2.545781
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         0.93054    0.01890   49.22  < 2e-16 ***
## FirstAuthorFemale1 -0.02510    0.00628   -4.00  6.4e-05 ***
## LastAuthorFemale1  -0.02884    0.00785   -3.68  0.00024 ***
## UniqueAuthors2      0.22522    0.01612   13.97  < 2e-16 ***
## UniqueAuthors3      0.25768    0.01600   16.10  < 2e-16 ***
## UniqueAuthors4      0.29155    0.01622   17.98  < 2e-16 ***
## UniqueAuthors5      0.33060    0.01571   21.04  < 2e-16 ***
## Year1997           -0.01201    0.01769   -0.68  0.49706
## Year1998           -0.02833    0.01674   -1.69  0.09056 .
## Year1999           -0.09550    0.01659   -5.75  8.8e-09 ***
```

```
## Year2000           -0.06563    0.01757   -3.73  0.00019 ***
## Year2001           -0.06684    0.01725   -3.87  0.00011 ***
## Year2002           -0.11913    0.01731   -6.88  6.1e-12 ***
## Year2003           -0.08961    0.01736   -5.16  2.5e-07 ***
## Year2004           -0.09301    0.01669   -5.57  2.5e-08 ***
## Year2005           -0.06033    0.01728   -3.49  0.00048 ***
## Year2006           -0.07009    0.01693   -4.14  3.5e-05 ***
## Year2007           -0.03958    0.01719   -2.30  0.02131 *
## Year2008           -0.07933    0.01641   -4.83  1.4e-06 ***
## Year2009           -0.08191    0.01693   -4.84  1.3e-06 ***
## Year2010           -0.06131    0.01666   -3.68  0.00023 ***
## Year2011           -0.09300    0.01670   -5.57  2.6e-08 ***
## Year2012           -0.10540    0.01704   -6.18  6.3e-10 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.412
## Multiple R-squared:  0.0437, Adjusted R-squared:  0.0427
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
##  4 observations c(1115,1323,6401,21345)
##   are outliers with |weight| = 0 ( < 4.4e-06);
##  1854 weights are ~= 1. The remaining 20792 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0003  0.8690  0.9510  0.9020  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          4.42e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.007  1           1.004
## LastAuthorFemale  1.007  1           1.004
## Year              1.007 16           1.000
```



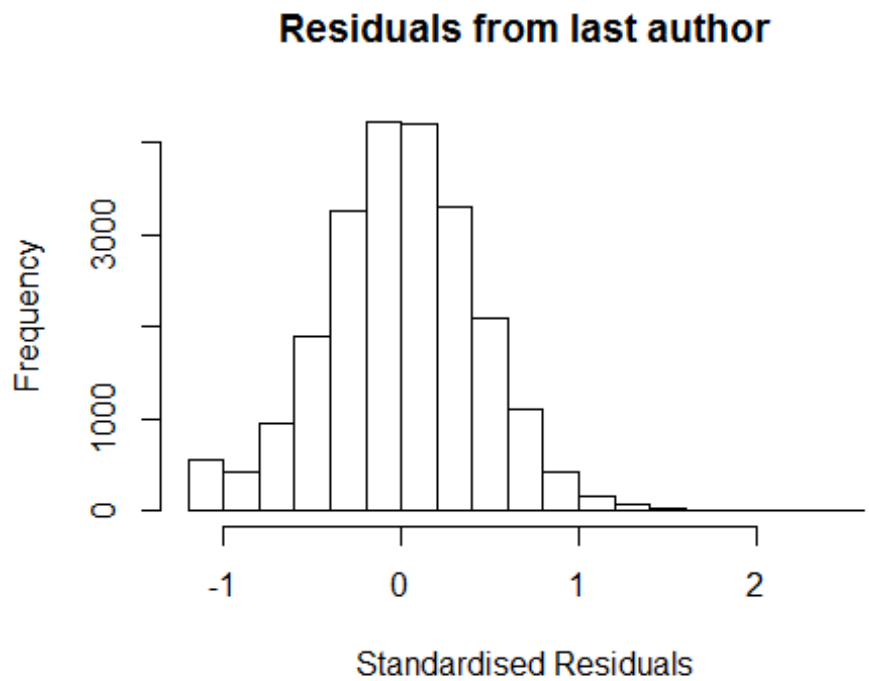

```
## [1] "List of  1 outliers with residuals above  2.5"
##        ScopusId  NLCS Year OneField Fields residuals
## 2463 0029878720 3.734 1996     1303      2     2.556
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##       Min        1Q    Median        3Q       Max
## -1.178200 -0.276951  0.000584  0.280970  2.555800
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)          1.17820    0.01253   94.03  < 2e-16 ***
## FirstAuthorFemale1 -0.00833    0.00632   -1.32  0.18711
## LastAuthorFemale1  -0.02831    0.00797   -3.55  0.00038 ***
## Year1997           -0.01317    0.01798   -0.73  0.46391
## Year1998           -0.03125    0.01692   -1.85  0.06478 .
## Year1999           -0.09584    0.01666   -5.75  9.0e-09 ***
## Year2000           -0.05573    0.01770   -3.15  0.00164 **
## Year2001           -0.06447    0.01734   -3.72  0.00020 ***
## Year2002           -0.10881    0.01743   -6.24  4.4e-10 ***
## Year2003           -0.08137    0.01762   -4.62  3.9e-06 ***
## Year2004           -0.08420    0.01686   -4.99  6.0e-07 ***
## Year2005           -0.04272    0.01746   -2.45  0.01441 *
```

```
## Year2006           -0.05284    0.01713   -3.09  0.00203 **
## Year2007           -0.02396    0.01728   -1.39  0.16563
## Year2008           -0.06290    0.01665   -3.78  0.00016 ***
## Year2009           -0.06268    0.01708   -3.67  0.00024 ***
## Year2010           -0.04114    0.01685   -2.44  0.01463 *
## Year2011           -0.06858    0.01694   -4.05  5.2e-05 ***
## Year2012           -0.08217    0.01736   -4.73  2.2e-06 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.415
## Multiple R-squared:  0.00531,    Adjusted R-squared:  0.00451
## Convergence in 10 IRWLS iterations
##
## Robustness weights:
##  2 observations c(1115,21345) are outliers with |weight| = 0 ( < 4.4e-06);
##  1958 weights are ~= 1. The remaining 20690 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0282  0.8670  0.9500  0.9000  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          4.42e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.004  1           1.002
## Year              1.004 16           1.000
```




```
## [1] "List of  1 outliers with residuals above  2.5"
##        ScopusId  NLCS Year OneField Fields residuals
## 2463 0029878720 3.734 1996     1303      2     2.556
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.1749 -0.2763  0.0016  0.2806  2.5591
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.17490    0.01249   94.05  < 2e-16 ***
## FirstAuthorFemale1 -0.01076    0.00632   -1.70  0.08866 .
## Year1997           -0.01304    0.01798   -0.72  0.46847
## Year1998           -0.03152    0.01691   -1.86  0.06238 .
## Year1999           -0.09583    0.01667   -5.75  9.1e-09 ***
## Year2000           -0.05569    0.01769   -3.15  0.00165 **
## Year2001           -0.06570    0.01734   -3.79  0.00015 ***
## Year2002           -0.10969    0.01743   -6.29  3.2e-10 ***
## Year2003           -0.08156    0.01762   -4.63  3.7e-06 ***
## Year2004           -0.08419    0.01686   -4.99  6.0e-07 ***
## Year2005           -0.04294    0.01747   -2.46  0.01397 *
## Year2006           -0.05360    0.01712   -3.13  0.00175 **
```

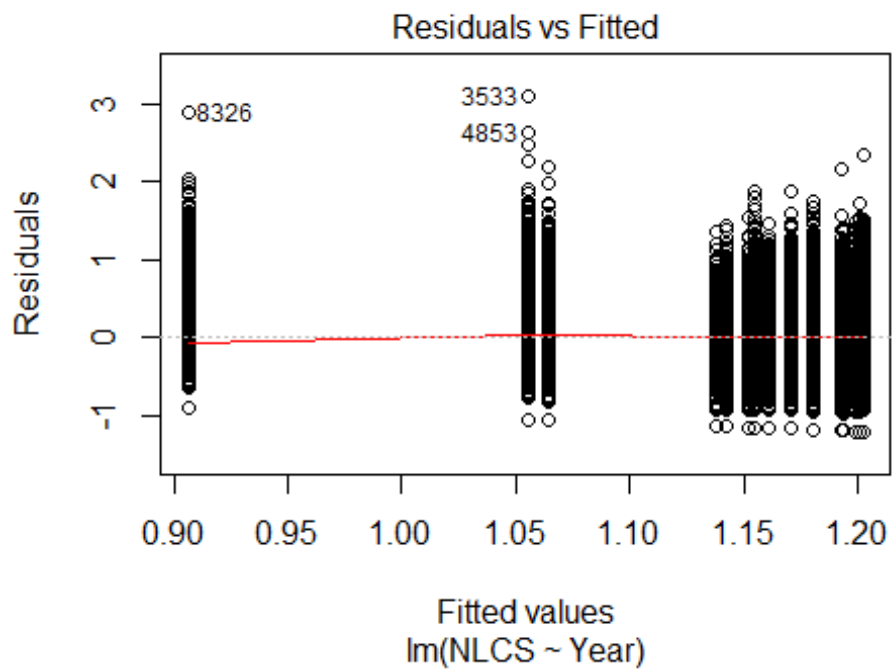
```
## Year2007           -0.02442    0.01729   -1.41  0.15780
## Year2008           -0.06359    0.01664   -3.82  0.00013 ***
## Year2009           -0.06412    0.01708   -3.75  0.00018 ***
## Year2010           -0.04199    0.01684   -2.49  0.01267 *
## Year2011           -0.06981    0.01694   -4.12  3.8e-05 ***
## Year2012           -0.08350    0.01735   -4.81  1.5e-06 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.415
## Multiple R-squared:  0.00471,    Adjusted R-squared:  0.00397
## Convergence in 10 IRWLS iterations
##
## Robustness weights:
##  2 observations c(1115,21345) are outliers with |weight| = 0 ( < 4.4e-06);
##  1899 weights are ~= 1. The remaining 20749 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0273  0.8680  0.9510  0.9000  0.9860  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          4.42e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##                   GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.004  1           1.002
## Year             1.004 16           1.000
```



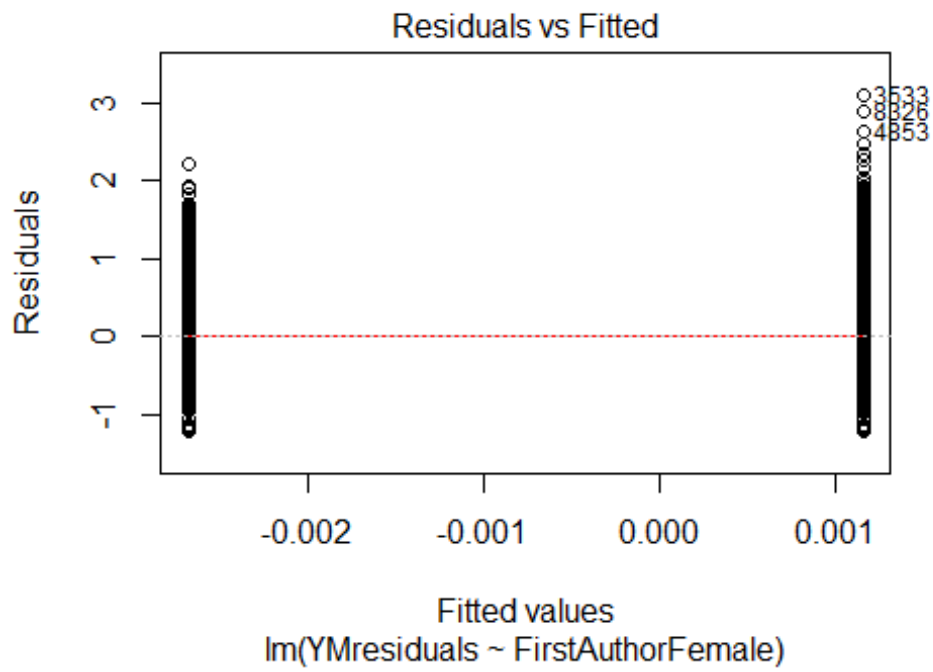

```
## [1] "List of  1 outliers with residuals above  2.5"
##        ScopusId  NLCS Year OneField Fields residuals
## 2463 0029878720 3.734 1996     1303      2     2.556
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##       Min        1Q    Median        3Q       Max
## -1.176216 -0.276281  0.000894  0.280771  2.557784
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.17622    0.01241   94.80  < 2e-16 ***
## LastAuthorFemale1 -0.02941    0.00796   -3.69  0.00022 ***
## Year1997          -0.01334    0.01799   -0.74  0.45821
## Year1998          -0.03128    0.01693   -1.85  0.06469 .
## Year1999          -0.09587    0.01667   -5.75  8.9e-09 ***
## Year2000          -0.05599    0.01770   -3.16  0.00156 **
## Year2001          -0.06458    0.01735   -3.72  0.00020 ***
## Year2002          -0.10893    0.01744   -6.24  4.3e-10 ***
## Year2003          -0.08156    0.01762   -4.63  3.7e-06 ***
## Year2004          -0.08455    0.01686   -5.01  5.4e-07 ***
## Year2005          -0.04308    0.01746   -2.47  0.01362 *
## Year2006          -0.05319    0.01713   -3.10  0.00191 **
```

```
## Year2007          -0.02410    0.01729   -1.39  0.16328
## Year2008          -0.06329    0.01665   -3.80  0.00014 ***
## Year2009          -0.06314    0.01708   -3.70  0.00022 ***
## Year2010          -0.04159    0.01686   -2.47  0.01363 *
## Year2011          -0.06892    0.01695   -4.07  4.8e-05 ***
## Year2012          -0.08282    0.01736   -4.77  1.9e-06 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.415
## Multiple R-squared:  0.00522,    Adjusted R-squared:  0.00447
## Convergence in 10 IRWLS iterations
##
## Robustness weights:
##  2 observations c(1115,21345) are outliers with |weight| = 0 ( < 4.4e-06);
##  1935 weights are ~= 1. The remaining 20713 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0277  0.8670  0.9510  0.9000  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          4.42e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Sample size for the above analysis:  22650"
## [1] ""
## [1] ""
## [1] "####################################"
## [1] "Analysis of AJSC 1305"
## [1] "####################################"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 4708 4042 2997 1631 1724 1707 1848 1676 1910 2297 2344 2511 2669 2767 2924
## 2011 2012
## 3112 3072
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 1363 1585 1187  909  727  668 1035  993 1152 1389 1374 1573 1636 1726 1771
```

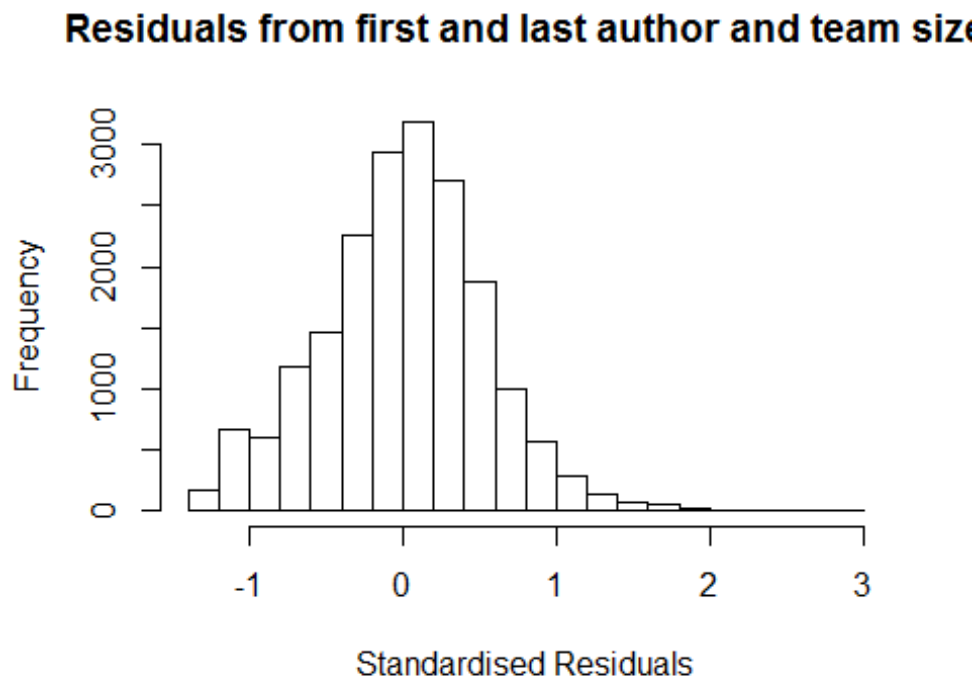
```
## 2011 2012
## 1907 1848
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 1142 1362 1014  782  622  571  889  846  994 1150 1163 1306 1351 1445 1449
## 2011 2012
## 1567 1511
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
##  Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 1900, df = 16, p-value <2e-16
```




```
##
##  Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 52, df = 1, p-value = 5e-13
```

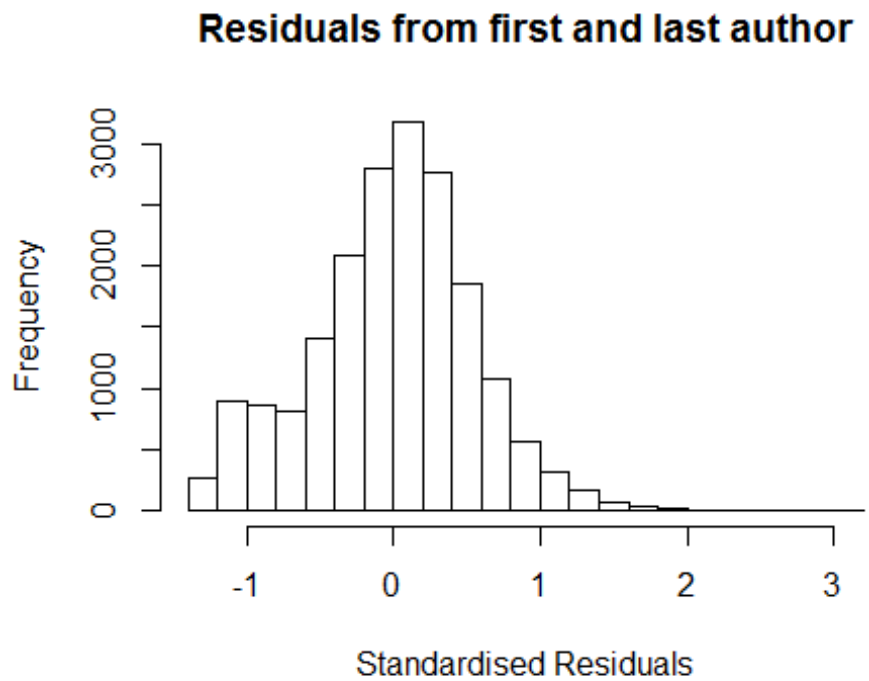



```
## [1] "Female first author team size 2018 geometric mean: 4.44473799334926"
## [1] "Male first author team size 2018 geometric mean: 3.95527217068427"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 230000, p-value = 0.003
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 3.92963310504279"
## [1] "Male last author team size 2018 geometric mean: 4.16075124387709"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 160000, p-value = 0.2
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.030  1           1.015
## LastAuthorFemale  1.027  1           1.014
## UniqueAuthors     1.158  4           1.019
## Year              1.168 16           1.005
```



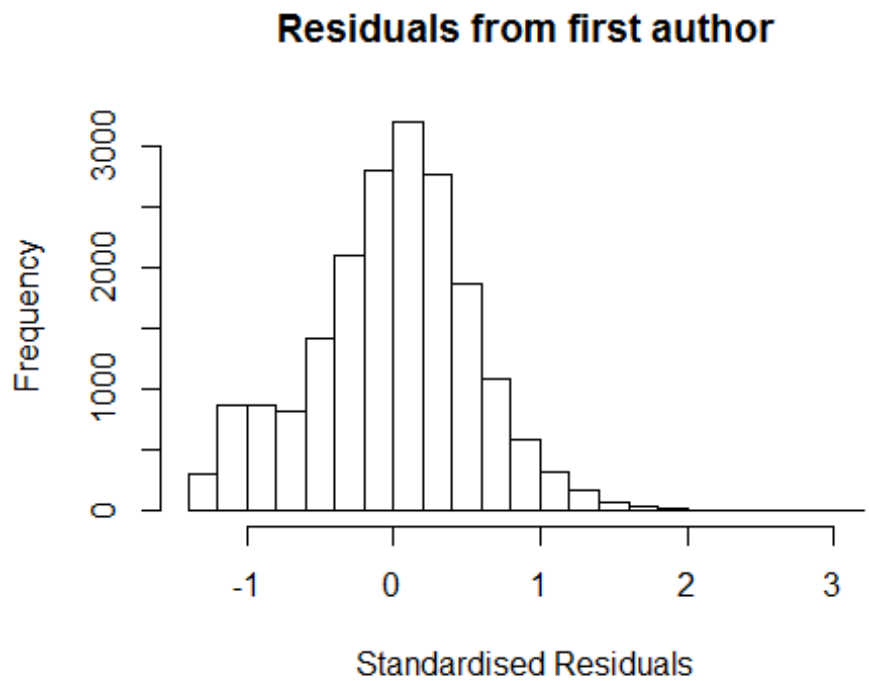

```
## [1] "List of  3 outliers with residuals above  2.5"
##        ScopusId  NLCS Year OneField Fields residuals
## 3533 0001506104 4.149 1996     1300      2     2.992
## 4853 0029670262 3.675 1996     1305      6     2.560
## 8326 0031022694 3.793 1997     1300      2     2.786
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
UniqueAuthors +
##     Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##     k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.3968 -0.3316  0.0194  0.3299  2.9922
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         0.67245    0.03465   19.40  < 2e-16 ***
## FirstAuthorFemale1 -0.01843    0.00820   -2.25   0.0246 *
## LastAuthorFemale1  -0.01408    0.00962   -1.46   0.1436
## UniqueAuthors2      0.44287    0.01885   23.49  < 2e-16 ***
## UniqueAuthors3      0.48431    0.01884   25.71  < 2e-16 ***
## UniqueAuthors4      0.52592    0.01912   27.51  < 2e-16 ***
## UniqueAuthors5      0.62494    0.01795   34.81  < 2e-16 ***
## Year1997           -0.19110    0.04613   -4.14  3.5e-05 ***
```

```
## Year1998           -0.00544    0.04179   -0.13   0.8965
## Year1999            0.09939    0.03822    2.60   0.0093 **
## Year2000            0.07581    0.03909    1.94   0.0525 .
## Year2001            0.05163    0.03912    1.32   0.1870
## Year2002           -0.00148    0.03720   -0.04   0.9683
## Year2003           -0.01805    0.03662   -0.49   0.6220
## Year2004           -0.00820    0.03628   -0.23   0.8211
## Year2005           -0.01156    0.03592   -0.32   0.7477
## Year2006           -0.00844    0.03593   -0.24   0.8142
## Year2007            0.02352    0.03569    0.66   0.5100
## Year2008           -0.00277    0.03595   -0.08   0.9386
## Year2009            0.01247    0.03575    0.35   0.7273
## Year2010            0.01767    0.03567    0.50   0.6203
## Year2011            0.02087    0.03590    0.58   0.5610
## Year2012            0.03021    0.03562    0.85   0.3963
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.483
## Multiple R-squared:  0.12,   Adjusted R-squared:  0.119
## Convergence in 22 IRWLS iterations
##
## Robustness weights:
##  7 observations c(554,630,1076,2084,2092,5469,18661)
##   are outliers with |weight| = 0 ( < 5.2e-06);
##  1556 weights are ~= 1. The remaining 17601 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0132  0.8520  0.9480  0.8920  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          5.22e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.014  1           1.007
## LastAuthorFemale  1.014  1           1.007
## Year              1.018 16           1.001
```



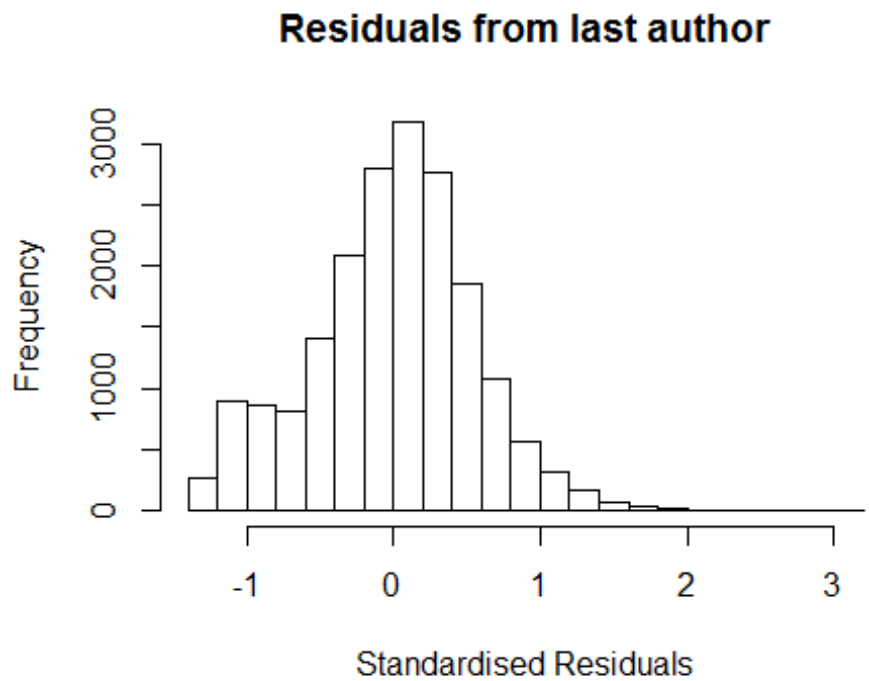

```
## [1] "List of  3 outliers with residuals above  2.5"
##        ScopusId  NLCS Year OneField Fields residuals
## 3533 0001506104 4.149 1996     1300      2     3.095
## 4853 0029670262 3.675 1996     1305      6     2.621
## 8326 0031022694 3.793 1997     1300      2     2.937
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.2242 -0.3411  0.0302  0.3460  3.0948
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.05417    0.03556   29.64  < 2e-16 ***
## FirstAuthorFemale1  0.00716    0.00857    0.84  0.40328
## LastAuthorFemale1  -0.02383    0.01024   -2.33  0.01993 *
## Year1997           -0.19767    0.04993   -3.96  7.6e-05 ***
## Year1998            0.04292    0.04459    0.96  0.33575
## Year1999            0.16291    0.04058    4.01  6.0e-05 ***
## Year2000            0.14956    0.04128    3.62  0.00029 ***
## Year2001            0.12291    0.04143    2.97  0.00302 **
## Year2002            0.09617    0.03955    2.43  0.01504 *
## Year2003            0.09122    0.03870    2.36  0.01844 *
```

```
## Year2004            0.10594    0.03830    2.77  0.00568 **
## Year2005            0.10439    0.03795    2.75  0.00596 **
## Year2006            0.11178    0.03784    2.95  0.00314 **
## Year2007            0.14675    0.03758    3.90  9.5e-05 ***
## Year2008            0.12115    0.03797    3.19  0.00142 **
## Year2009            0.13122    0.03790    3.46  0.00054 ***
## Year2010            0.15448    0.03768    4.10  4.2e-05 ***
## Year2011            0.15085    0.03796    3.97  7.1e-05 ***
## Year2012            0.15700    0.03765    4.17  3.1e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.488
## Multiple R-squared:  0.0258, Adjusted R-squared:  0.0249
## Convergence in 23 IRWLS iterations
##
## Robustness weights:
##  5 observations c(554,630,1076,2084,18661)
##   are outliers with |weight| = 0 ( < 5.2e-06);
##  1548 weights are ~= 1. The remaining 17611 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0002  0.8400  0.9460  0.8860  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          5.22e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##                   GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.01  1           1.005
## Year              1.01 16           1.000
```




```
## [1] "List of  3 outliers with residuals above  2.5"
##        ScopusId  NLCS Year OneField Fields residuals
## 3533 0001506104 4.149 1996     1300      2     3.095
## 4853 0029670262 3.675 1996     1305      6     2.621
## 8326 0031022694 3.793 1997     1300      2     2.937
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.2176 -0.3418  0.0303  0.3446  3.0977
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.05133    0.03530   29.79  < 2e-16 ***
## FirstAuthorFemale1  0.00433    0.00857    0.51  0.61312
## Year1997           -0.19702    0.04923   -4.00  6.3e-05 ***
## Year1998            0.04223    0.04470    0.94  0.34478
## Year1999            0.16199    0.04050    4.00  6.4e-05 ***
## Year2000            0.14993    0.04113    3.65  0.00027 ***
## Year2001            0.12210    0.04130    2.96  0.00311 **
## Year2002            0.09506    0.03943    2.41  0.01593 *
## Year2003            0.09056    0.03854    2.35  0.01879 *
## Year2004            0.10477    0.03814    2.75  0.00603 **
```

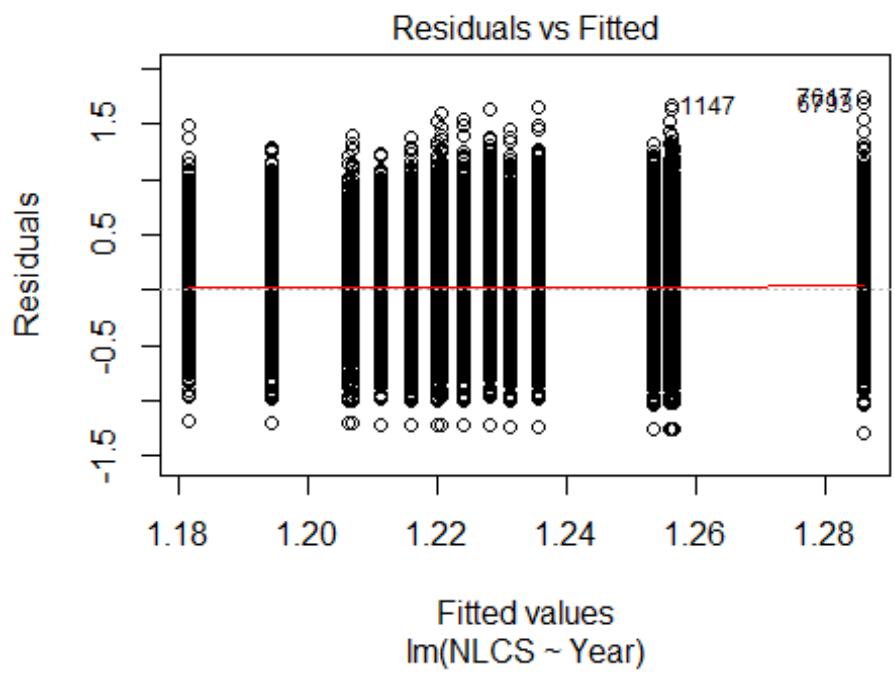
```
## Year2005            0.10343    0.03782    2.73  0.00625 **
## Year2006            0.11091    0.03768    2.94  0.00325 **
## Year2007            0.14556    0.03742    3.89  0.00010 ***
## Year2008            0.12010    0.03781    3.18  0.00149 **
## Year2009            0.12980    0.03778    3.44  0.00059 ***
## Year2010            0.15302    0.03755    4.07  4.6e-05 ***
## Year2011            0.14944    0.03785    3.95  7.9e-05 ***
## Year2012            0.15563    0.03745    4.16  3.3e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.491
## Multiple R-squared:  0.0254, Adjusted R-squared:  0.0245
## Convergence in 24 IRWLS iterations
##
## Robustness weights:
##  5 observations c(554,630,1076,2084,18661)
##   are outliers with |weight| = 0 ( < 5.2e-06);
##  1559 weights are ~= 1. The remaining 17600 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0004  0.8420  0.9470  0.8860  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          5.22e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                    psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##                   GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.009  1           1.005
## Year             1.009 16           1.000
```



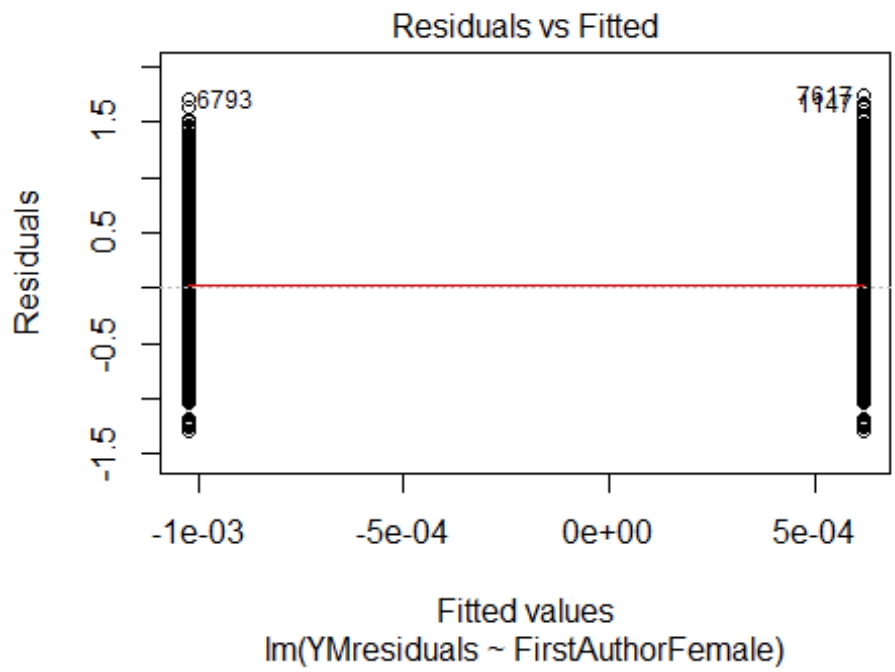

```
## [1] "List of  3 outliers with residuals above  2.5"
##        ScopusId  NLCS Year OneField Fields residuals
## 3533 0001506104 4.149 1996     1300      2     3.095
## 4853 0029670262 3.675 1996     1305      6     2.621
## 8326 0031022694 3.793 1997     1300      2     2.937
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.2184 -0.3412  0.0307  0.3456  3.0934
##
## Coefficients:
##                   Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.0556     0.0353   29.94  < 2e-16 ***
## LastAuthorFemale1  -0.0227     0.0102   -2.22  0.02617 *
## Year1997           -0.1969     0.0492   -4.00  6.3e-05 ***
## Year1998            0.0425     0.0446    0.95  0.34027
## Year1999            0.1628     0.0405    4.02  5.7e-05 ***
## Year2000            0.1497     0.0411    3.64  0.00027 ***
## Year2001            0.1232     0.0412    2.99  0.00281 **
## Year2002            0.0964     0.0394    2.45  0.01438 *
## Year2003            0.0916     0.0385    2.38  0.01735 *
## Year2004            0.1066     0.0381    2.80  0.00513 **
```

```
## Year2005            0.1050     0.0377    2.78  0.00543 **
## Year2006            0.1123     0.0376    2.98  0.00284 **
## Year2007            0.1474     0.0373    3.95  7.9e-05 ***
## Year2008            0.1218     0.0377    3.23  0.00125 **
## Year2009            0.1318     0.0377    3.49  0.00048 ***
## Year2010            0.1548     0.0375    4.13  3.7e-05 ***
## Year2011            0.1513     0.0378    4.00  6.3e-05 ***
## Year2012            0.1576     0.0374    4.21  2.5e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.491
## Multiple R-squared:  0.0256, Adjusted R-squared:  0.0248
## Convergence in 24 IRWLS iterations
##
## Robustness weights:
##  5 observations c(554,630,1076,2084,18661)
##   are outliers with |weight| = 0 ( < 5.2e-06);
##  1546 weights are ~= 1. The remaining 17613 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0006  0.8420  0.9470  0.8870  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          5.22e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                    psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Sample size for the above analysis:  19164"
## [1] ""
## [1] ""
## [1] "##################################"
## [1] "Analysis of AJSC 1306"
## [1] "##################################"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 3818 3609 3793 3652 3758 4070 4153 3893 4057 4473 4649 4836 5023 4848 4769
## 2011 2012
## 5048 5012
```

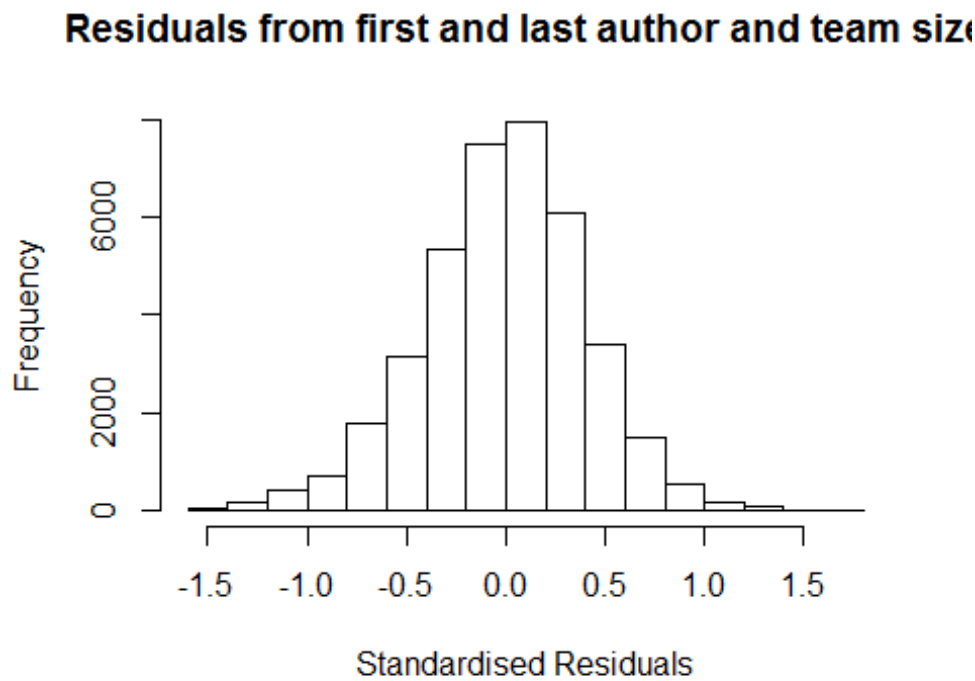
```
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 2172 2134 2352 2282 2046 1811 2785 2668 2684 2985 2950 3090 3269 3155 3003
## 2011 2012
## 3300 3314
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 1851 1851 2022 1954 1717 1527 2352 2244 2230 2519 2492 2599 2747 2650 2551
## 2011 2012
## 2770 2829
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
##  Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 650, df = 16, p-value <2e-16
```




```
##
##  Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 100, df = 1, p-value <2e-16
```

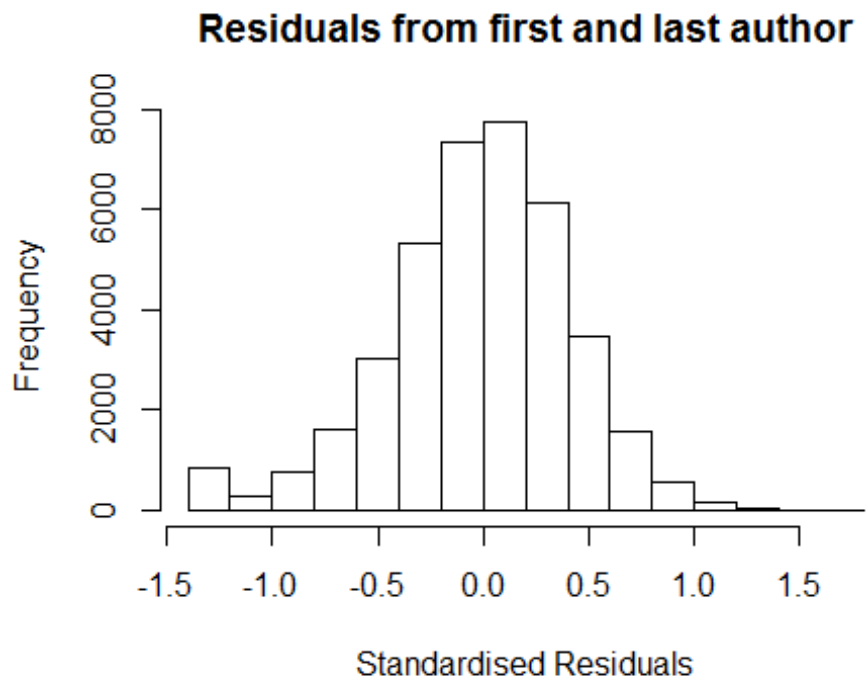



```
## [1] "Female first author team size 2018 geometric mean: 7.31631121645218"
## [1] "Male first author team size 2018 geometric mean: 7.0655421192179"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 870000, p-value = 0.7
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 7.0032665891806"
## [1] "Male last author team size 2018 geometric mean: 7.26601268878691"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 730000, p-value = 0.02
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.024  1           1.012
## LastAuthorFemale  1.025  1           1.012
## UniqueAuthors     1.035  4           1.004
## Year              1.047 16           1.001
```



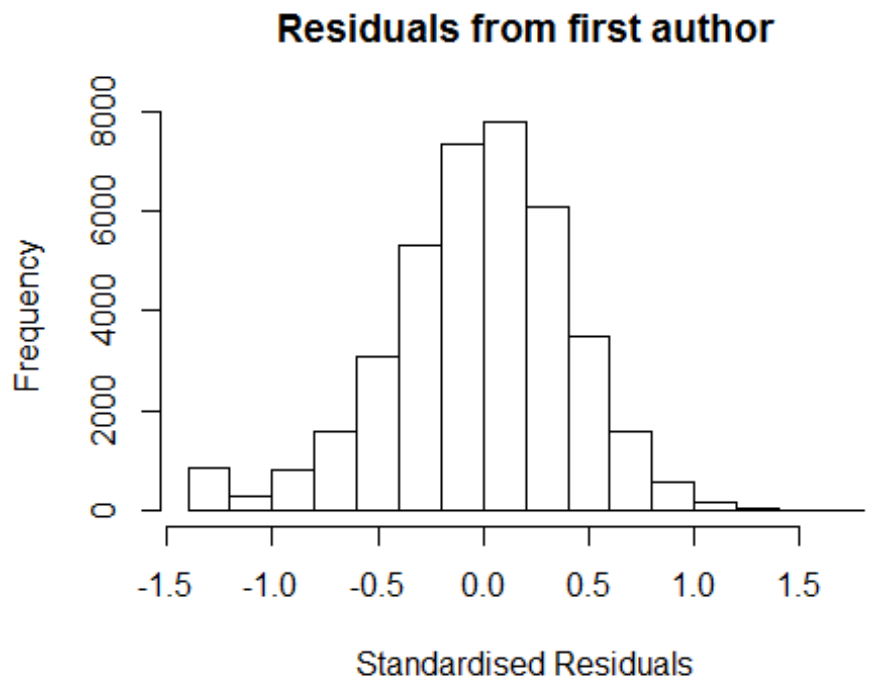

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
UniqueAuthors +
##     Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##     k.max = 1000))
##  \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.42425 -0.26088  0.00736  0.26001  1.71832
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         0.86583    0.02076   41.71  < 2e-16 ***
## FirstAuthorFemale1 -0.01067    0.00425   -2.51    0.012 *
## LastAuthorFemale1  -0.01076    0.00489   -2.20    0.028 *
## UniqueAuthors2      0.31452    0.01874   16.79  < 2e-16 ***
## UniqueAuthors3      0.39454    0.01797   21.95  < 2e-16 ***
## UniqueAuthors4      0.42290    0.01767   23.93  < 2e-16 ***
## UniqueAuthors5      0.55842    0.01682   33.20  < 2e-16 ***
## Year1997           -0.01047    0.01607   -0.65    0.515
## Year1998           -0.06186    0.01553   -3.98  6.8e-05 ***
## Year1999           -0.09852    0.01537   -6.41  1.5e-10 ***
```

```
## Year2000           -0.11370    0.01570   -7.24  4.5e-13 ***
## Year2001           -0.10683    0.01681   -6.36  2.1e-10 ***
## Year2002           -0.09425    0.01495   -6.30  2.9e-10 ***
## Year2003           -0.13026    0.01426   -9.13  < 2e-16 ***
## Year2004           -0.15994    0.01442  -11.09  < 2e-16 ***
## Year2005           -0.13693    0.01440   -9.51  < 2e-16 ***
## Year2006           -0.12179    0.01443   -8.44  < 2e-16 ***
## Year2007           -0.09055    0.01443   -6.27  3.5e-10 ***
## Year2008           -0.08993    0.01430   -6.29  3.3e-10 ***
## Year2009           -0.08950    0.01446   -6.19  6.0e-10 ***
## Year2010           -0.11307    0.01459   -7.75  9.3e-15 ***
## Year2011           -0.12268    0.01458   -8.41  < 2e-16 ***
## Year2012           -0.11901    0.01464   -8.13  4.5e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.386
## Multiple R-squared:  0.104,  Adjusted R-squared:  0.103
## Convergence in 16 IRWLS iterations
##
## Robustness weights:
##  3233 weights are ~= 1. The remaining 35672 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0096  0.8660  0.9500  0.8970  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          2.57e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.019  1           1.009
## LastAuthorFemale  1.020  1           1.010
## Year              1.016 16           1.001
```



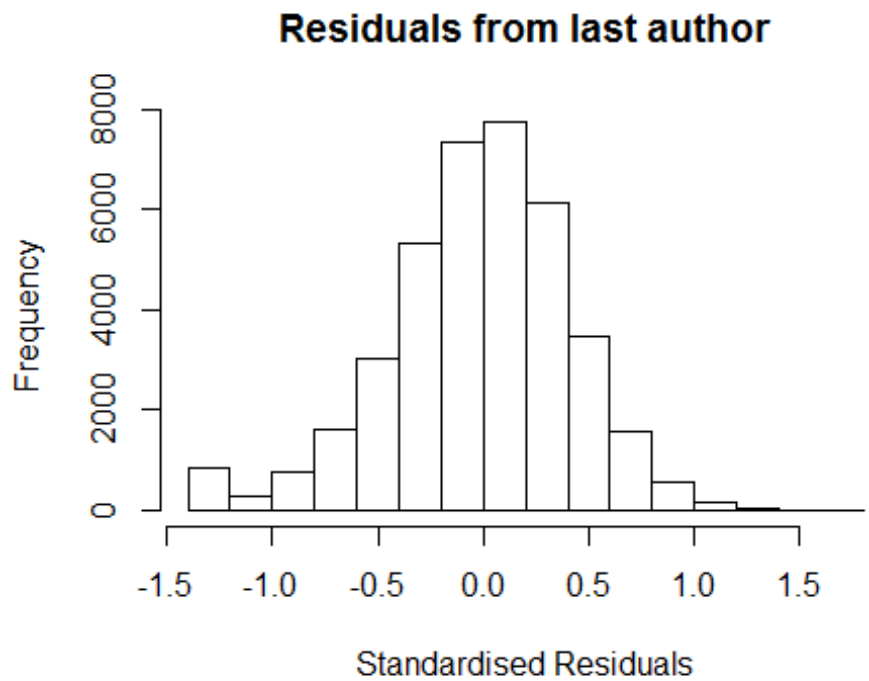

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.32006 -0.26909  0.00706  0.26361  1.71339
##
## Coefficients:
##                      Estimate Std. Error t value Pr(>|t|)
## (Intercept)          1.320062   0.012712  103.85  < 2e-16 ***
## FirstAuthorFemale1 -0.000764   0.004362   -0.18   0.8611
## LastAuthorFemale1  -0.020149   0.005028   -4.01  6.1e-05 ***
## Year1997           -0.009449   0.016515   -0.57   0.5672
## Year1998           -0.054967   0.016010   -3.43   0.0006 ***
## Year1999           -0.085879   0.015757   -5.45  5.1e-08 ***
## Year2000           -0.091161   0.016144   -5.65  1.6e-08 ***
## Year2001           -0.082426   0.017252   -4.78  1.8e-06 ***
## Year2002           -0.066603   0.015365   -4.33  1.5e-05 ***
## Year2003           -0.095432   0.014762   -6.46  1.0e-10 ***
## Year2004           -0.121834   0.014971   -8.14  4.1e-16 ***
## Year2005           -0.088628   0.014868   -5.96  2.5e-09 ***
```

```
## Year2006           -0.078538   0.014877   -5.28  1.3e-07 ***
## Year2007           -0.045762   0.014905   -3.07   0.0021 **
## Year2008           -0.040451   0.014730   -2.75   0.0060 **
## Year2009           -0.042762   0.014870   -2.88   0.0040 **
## Year2010           -0.064890   0.015050   -4.31  1.6e-05 ***
## Year2011           -0.078069   0.015049   -5.19  2.1e-07 ***
## Year2012           -0.062207   0.015111   -4.12  3.9e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.396
## Multiple R-squared:  0.00525,    Adjusted R-squared:  0.00479
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
##  3220 weights are ~= 1. The remaining 35685 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0213  0.8660  0.9510  0.8930  0.9860  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          2.57e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.008  1           1.004
## Year              1.008 16           1.000
```




```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.3170 -0.2677  0.0069  0.2633  1.7166
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.31695    0.01268  103.85  < 2e-16 ***
## FirstAuthorFemale1 -0.00310    0.00434   -0.71  0.47486
## Year1997           -0.00957    0.01651   -0.58  0.56228
## Year1998           -0.05536    0.01601   -3.46  0.00054 ***
## Year1999           -0.08566    0.01575   -5.44  5.4e-08 ***
## Year2000           -0.09132    0.01615   -5.66  1.6e-08 ***
## Year2001           -0.08230    0.01725   -4.77  1.8e-06 ***
## Year2002           -0.06685    0.01536   -4.35  1.4e-05 ***
## Year2003           -0.09584    0.01476   -6.49  8.5e-11 ***
## Year2004           -0.12228    0.01497   -8.17  3.2e-16 ***
## Year2005           -0.08944    0.01487   -6.02  1.8e-09 ***
## Year2006           -0.07918    0.01488   -5.32  1.0e-07 ***
```

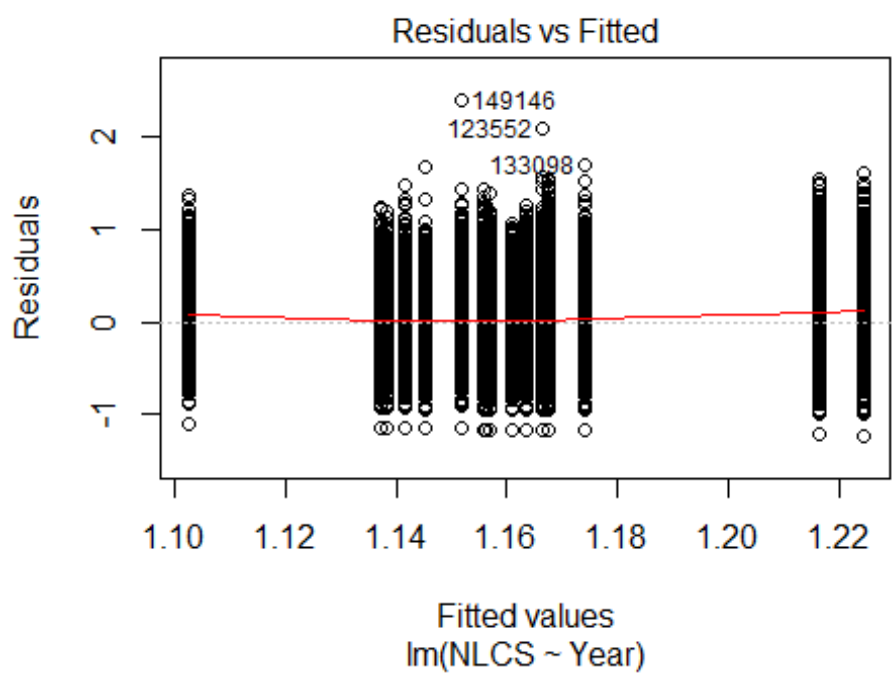
```
## Year2007           -0.04678    0.01490   -3.14  0.00169 **
## Year2008           -0.04108    0.01473   -2.79  0.00530 **
## Year2009           -0.04386    0.01486   -2.95  0.00317 **
## Year2010           -0.06595    0.01505   -4.38  1.2e-05 ***
## Year2011           -0.07969    0.01504   -5.30  1.2e-07 ***
## Year2012           -0.06397    0.01510   -4.24  2.3e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.396
## Multiple R-squared:  0.00483,    Adjusted R-squared:  0.00439
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
##  3215 weights are ~= 1. The remaining 35690 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0204  0.8660  0.9510  0.8930  0.9860  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          2.57e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##                   GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.009  1           1.005
## Year             1.009 16           1.000
```



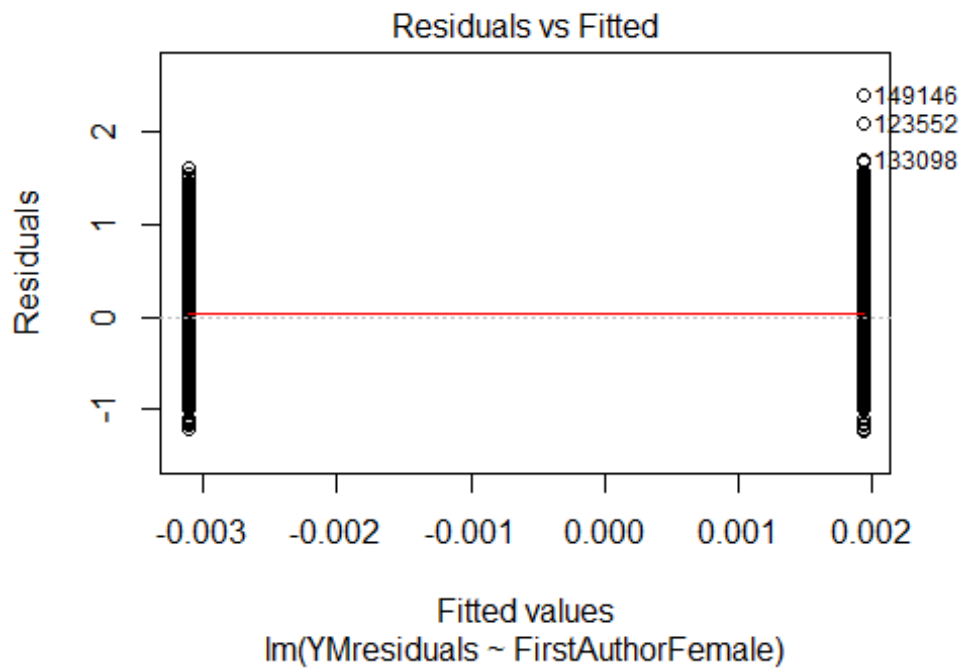

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.31984 -0.26916  0.00687  0.26363  1.71362
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.31984    0.01263  104.48  < 2e-16 ***
## LastAuthorFemale1 -0.02027    0.00500   -4.05  5.1e-05 ***
## Year1997          -0.00946    0.01651   -0.57  0.56679
## Year1998          -0.05499    0.01601   -3.44  0.00059 ***
## Year1999          -0.08590    0.01576   -5.45  5.0e-08 ***
## Year2000          -0.09117    0.01614   -5.65  1.6e-08 ***
## Year2001          -0.08245    0.01725   -4.78  1.8e-06 ***
## Year2002          -0.06663    0.01536   -4.34  1.5e-05 ***
## Year2003          -0.09547    0.01476   -6.47  1.0e-10 ***
## Year2004          -0.12187    0.01497   -8.14  4.1e-16 ***
## Year2005          -0.08868    0.01487   -5.97  2.5e-09 ***
## Year2006          -0.07861    0.01487   -5.29  1.3e-07 ***
```

```
## Year2007          -0.04582    0.01490   -3.08  0.00210 **
## Year2008          -0.04052    0.01472   -2.75  0.00592 **
## Year2009          -0.04285    0.01486   -2.88  0.00394 **
## Year2010          -0.06497    0.01504   -4.32  1.6e-05 ***
## Year2011          -0.07815    0.01504   -5.19  2.1e-07 ***
## Year2012          -0.06229    0.01510   -4.13  3.7e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.396
## Multiple R-squared:  0.00525,    Adjusted R-squared:  0.00481
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
##  3214 weights are ~= 1. The remaining 35691 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0212  0.8660  0.9510  0.8930  0.9860  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          2.57e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Sample size for the above analysis:  38905"
## [1] ""
## [1] ""
## [1] "####################################"
## [1] "Analysis of AJSC 1307"
## [1] "####################################"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
##  1996  1997  1998  1999  2000  2001  2002  2003  2004  2005  2006  2007
## 11449  9734  8957  7061  8341  7151  7470  6310  6984  6851  6844  7122
##  2008  2009  2010  2011  2012
##  7269  7287  7219  7360  7142
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 5601 5236 4947 4467 4239 3326 4691 3893 4256 4206 4176 4370 4504 4482 4456
## 2011 2012
```

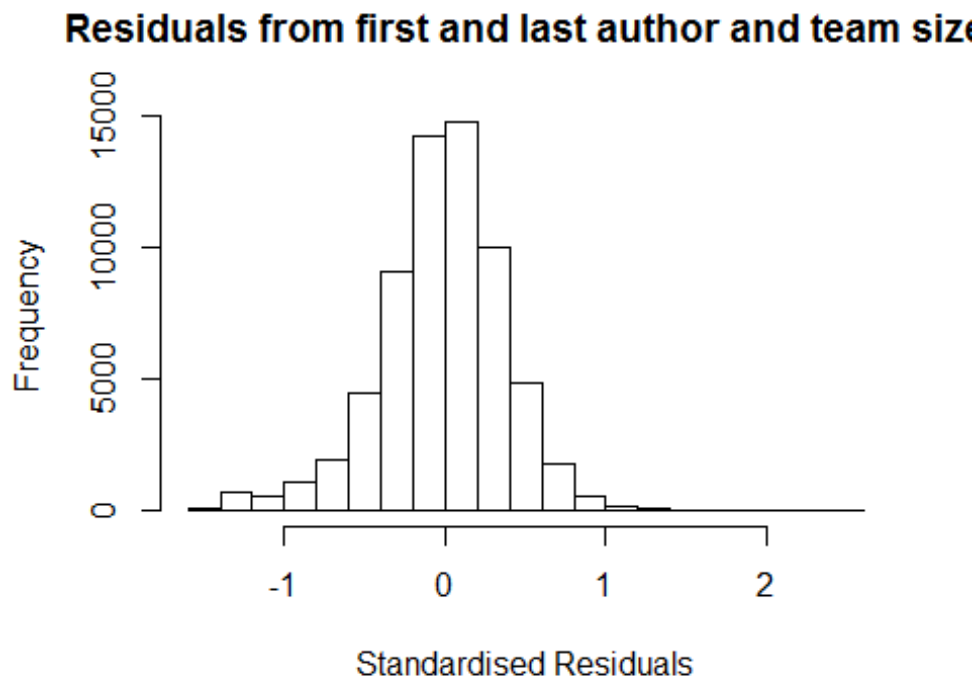
```
## 4553 4496
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 4791 4499 4284 3835 3615 2837 3978 3296 3566 3510 3499 3688 3780 3703 3682
## 2011 2012
## 3796 3708
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
##  Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 5100, df = 16, p-value <2e-16
```




```
##
##  Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 180, df = 1, p-value <2e-16
```

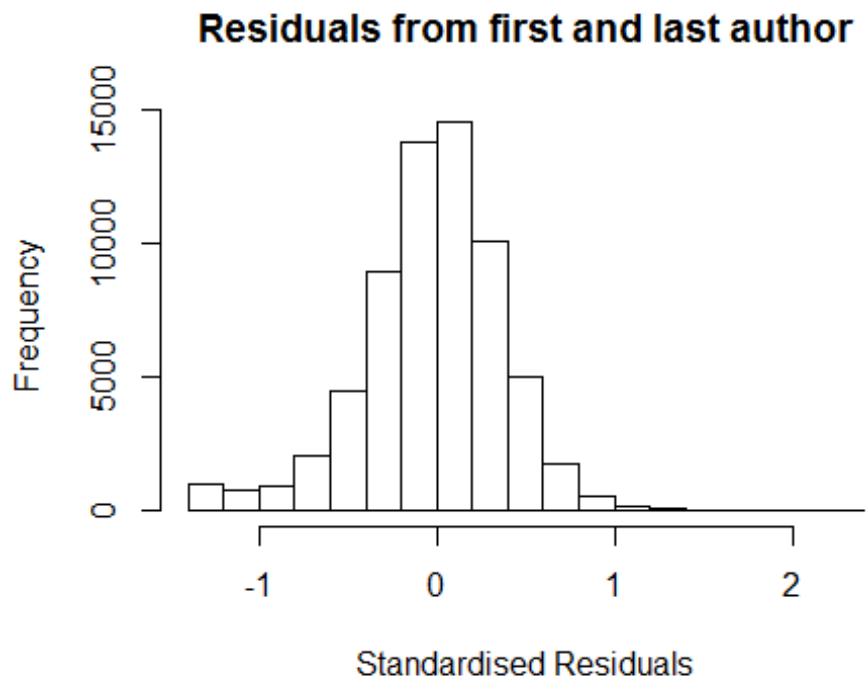



```
## [1] "Female first author team size 2018 geometric mean: 5.15724833186727"
## [1] "Male first author team size 2018 geometric mean: 4.93781452555463"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 8e+05, p-value = 0.3
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 4.64698829074477"
## [1] "Male last author team size 2018 geometric mean: 5.18795674296245"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 580000, p-value = 8e-04
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##                     GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.012  1           1.006
## LastAuthorFemale  1.012  1           1.006
## UniqueAuthors     1.040  4           1.005
## Year              1.048 16           1.001
```



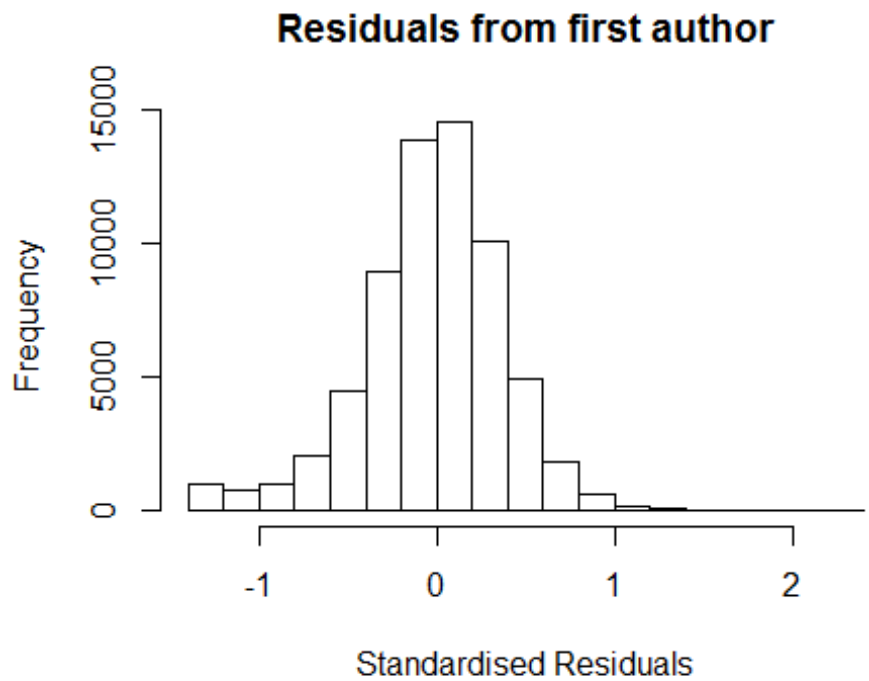

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
UniqueAuthors +
##     Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##     k.max = 1000))
##  \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.44324 -0.22907  0.00148  0.22173  2.47505
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.00947    0.01521   66.39  < 2e-16 ***
## FirstAuthorFemale1 -0.01571    0.00290   -5.42  6.0e-08 ***
## LastAuthorFemale1  -0.02467    0.00344   -7.17  7.4e-13 ***
## UniqueAuthors2      0.27378    0.01348   20.31  < 2e-16 ***
## UniqueAuthors3      0.31506    0.01336   23.59  < 2e-16 ***
## UniqueAuthors4      0.34461    0.01332   25.86  < 2e-16 ***
## UniqueAuthors5      0.43377    0.01307   33.19  < 2e-16 ***
## Year1997           -0.03162    0.01058   -2.99   0.0028 **
## Year1998           -0.09333    0.01002   -9.31  < 2e-16 ***
## Year1999           -0.17293    0.00925  -18.70  < 2e-16 ***
```

```
## Year2000           -0.17507    0.01007  -17.38  < 2e-16 ***
## Year2001           -0.18282    0.00978  -18.70  < 2e-16 ***
## Year2002           -0.18786    0.00920  -20.41  < 2e-16 ***
## Year2003           -0.21421    0.00939  -22.82  < 2e-16 ***
## Year2004           -0.20547    0.00913  -22.51  < 2e-16 ***
## Year2005           -0.21242    0.00929  -22.86  < 2e-16 ***
## Year2006           -0.21480    0.00935  -22.98  < 2e-16 ***
## Year2007           -0.19658    0.00945  -20.81  < 2e-16 ***
## Year2008           -0.19851    0.00932  -21.30  < 2e-16 ***
## Year2009           -0.19419    0.00947  -20.51  < 2e-16 ***
## Year2010           -0.19094    0.00960  -19.90  < 2e-16 ***
## Year2011           -0.19288    0.00951  -20.28  < 2e-16 ***
## Year2012           -0.21331    0.00976  -21.86  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.337
## Multiple R-squared:  0.0925, Adjusted R-squared:  0.0922
## Convergence in 19 IRWLS iterations
##
## Robustness weights:
##  5 observations c(7061,52053,57469,58988,62841)
##   are outliers with |weight| = 0 ( < 1.6e-06);
##  5445 weights are ~= 1. The remaining 58617 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0002  0.8620  0.9510  0.8860  0.9860  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.56e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.008  1           1.004
## LastAuthorFemale  1.012  1           1.006
## Year              1.010 16           1.000
```



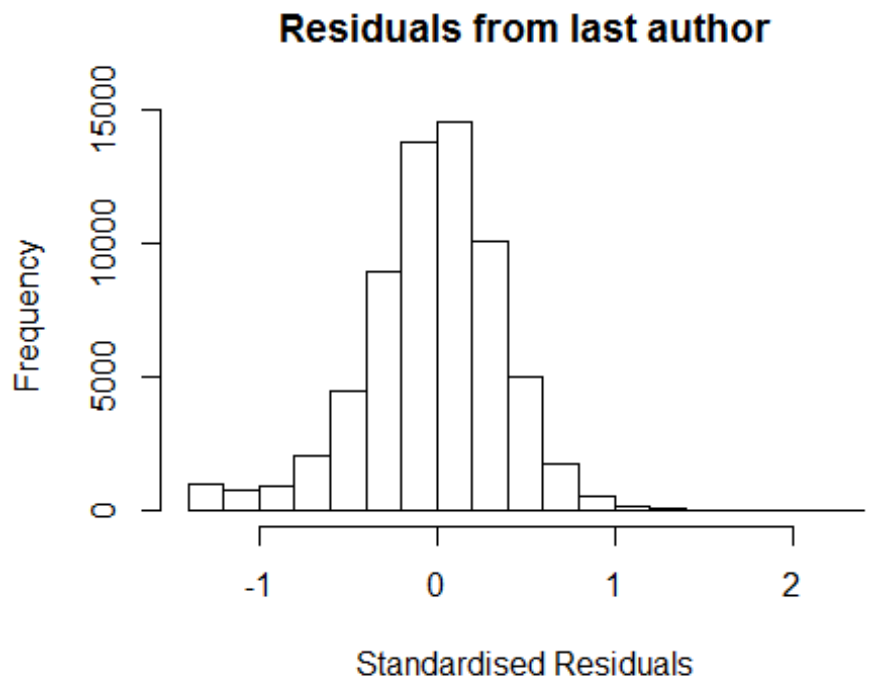

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##       Min        1Q    Median        3Q       Max
## -1.345597 -0.236188  0.000832  0.223418  2.376960
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)          1.34560    0.00784  171.55   <2e-16 ***
## FirstAuthorFemale1 -0.00765    0.00295   -2.59   0.0096 **
## LastAuthorFemale1  -0.02904    0.00349   -8.32   <2e-16 ***
## Year1997           -0.02419    0.01063   -2.27   0.0229 *
## Year1998           -0.09072    0.01013   -8.95   <2e-16 ***
## Year1999           -0.16795    0.00935  -17.96   <2e-16 ***
## Year2000           -0.17032    0.01027  -16.58   <2e-16 ***
## Year2001           -0.16633    0.00996  -16.70   <2e-16 ***
## Year2002           -0.16900    0.00935  -18.08   <2e-16 ***
## Year2003           -0.19393    0.00954  -20.33   <2e-16 ***
## Year2004           -0.18578    0.00928  -20.01   <2e-16 ***
## Year2005           -0.19076    0.00947  -20.15   <2e-16 ***
```

```
## Year2006           -0.19038    0.00945  -20.14   <2e-16 ***
## Year2007           -0.16975    0.00958  -17.72   <2e-16 ***
## Year2008           -0.16894    0.00948  -17.83   <2e-16 ***
## Year2009           -0.16389    0.00965  -16.99   <2e-16 ***
## Year2010           -0.15872    0.00981  -16.17   <2e-16 ***
## Year2011           -0.16117    0.00973  -16.56   <2e-16 ***
## Year2012           -0.17756    0.00995  -17.84   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.344
## Multiple R-squared:  0.0273, Adjusted R-squared:  0.0271
## Convergence in 15 IRWLS iterations
##
## Robustness weights:
##  2 observations c(52053,62841) are outliers with |weight| = 0 ( < 1.6e-
06);
##  5446 weights are ~= 1. The remaining 58619 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0068  0.8610  0.9510  0.8840  0.9860  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.56e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.004  1           1.002
## Year              1.004 16           1.000
```




```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##       Min        1Q    Median        3Q       Max
## -1.341080 -0.237411  0.000765  0.222809  2.383524
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)          1.34108    0.00785  170.80  < 2e-16 ***
## FirstAuthorFemale1 -0.00990    0.00295   -3.36  0.00078 ***
## Year1997           -0.02431    0.01065   -2.28  0.02244 *
## Year1998           -0.09084    0.01015   -8.95  < 2e-16 ***
## Year1999           -0.16873    0.00937  -18.02  < 2e-16 ***
## Year2000           -0.17089    0.01029  -16.60  < 2e-16 ***
## Year2001           -0.16693    0.00998  -16.73  < 2e-16 ***
## Year2002           -0.16959    0.00936  -18.11  < 2e-16 ***
## Year2003           -0.19458    0.00955  -20.37  < 2e-16 ***
## Year2004           -0.18688    0.00930  -20.10  < 2e-16 ***
## Year2005           -0.19156    0.00948  -20.22  < 2e-16 ***
## Year2006           -0.19153    0.00946  -20.24  < 2e-16 ***
```

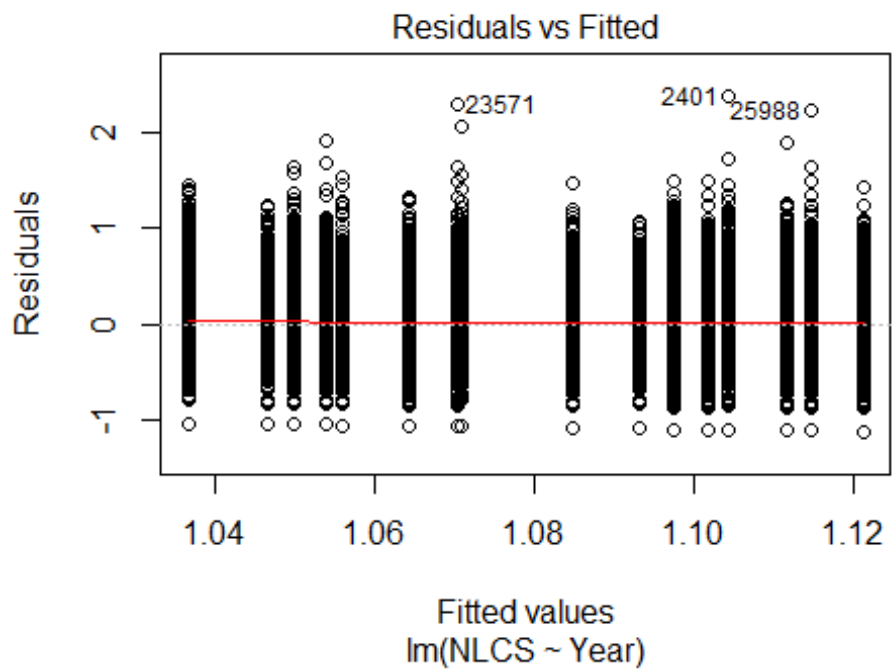
```
## Year2007           -0.17132    0.00958  -17.87  < 2e-16 ***
## Year2008           -0.17051    0.00949  -17.97  < 2e-16 ***
## Year2009           -0.16533    0.00966  -17.11  < 2e-16 ***
## Year2010           -0.16045    0.00982  -16.34  < 2e-16 ***
## Year2011           -0.16325    0.00974  -16.76  < 2e-16 ***
## Year2012           -0.17960    0.00996  -18.03  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.344
## Multiple R-squared:  0.0263, Adjusted R-squared:  0.0261
## Convergence in 15 IRWLS iterations
##
## Robustness weights:
##  2 observations c(52053,62841) are outliers with |weight| = 0 ( < 1.6e-
06);
##  5399 weights are ~= 1. The remaining 58666 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0057  0.8610  0.9510  0.8840  0.9860  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.56e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##                   GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.007  1           1.004
## Year             1.007 16           1.000
```



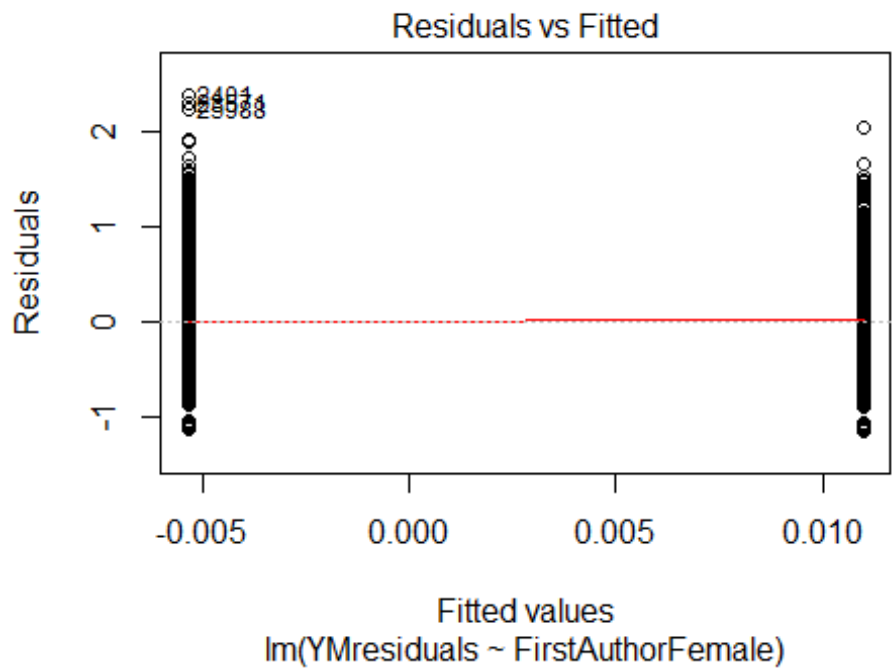

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.34290 -0.23641  0.00117  0.22327  2.38000
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.34290    0.00776  173.12   <2e-16 ***
## LastAuthorFemale1 -0.02987    0.00348   -8.58   <2e-16 ***
## Year1997          -0.02411    0.01064   -2.27    0.023 *
## Year1998          -0.09070    0.01014   -8.95   <2e-16 ***
## Year1999          -0.16800    0.00935  -17.96   <2e-16 ***
## Year2000          -0.17038    0.01028  -16.58   <2e-16 ***
## Year2001          -0.16645    0.00996  -16.71   <2e-16 ***
## Year2002          -0.16899    0.00935  -18.08   <2e-16 ***
## Year2003          -0.19411    0.00954  -20.34   <2e-16 ***
## Year2004          -0.18607    0.00929  -20.04   <2e-16 ***
## Year2005          -0.19097    0.00947  -20.17   <2e-16 ***
## Year2006          -0.19059    0.00946  -20.16   <2e-16 ***
```

```
## Year2007          -0.17002    0.00958  -17.75   <2e-16 ***
## Year2008          -0.16933    0.00948  -17.86   <2e-16 ***
## Year2009          -0.16420    0.00965  -17.02   <2e-16 ***
## Year2010          -0.15917    0.00982  -16.21   <2e-16 ***
## Year2011          -0.16161    0.00973  -16.61   <2e-16 ***
## Year2012          -0.17791    0.00996  -17.87   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.344
## Multiple R-squared:  0.0272, Adjusted R-squared:  0.0269
## Convergence in 15 IRWLS iterations
##
## Robustness weights:
##  2 observations c(52053,62841) are outliers with |weight| = 0 ( < 1.6e-
06);
##  5462 weights are ~= 1. The remaining 58603 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0062  0.8620  0.9510  0.8840  0.9860  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.56e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                  psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Sample size for the above analysis:  64067"
## [1] ""
## [1] ""
## [1] "##################################"
## [1] "Analysis of AJSC 1308"
## [1] "##################################"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 2009 1862 1897 1855 1871 2104 2202 2075 2149 2128 2075 2313 2393 2233 2283
## 2011 2012
## 2333 2095
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
```

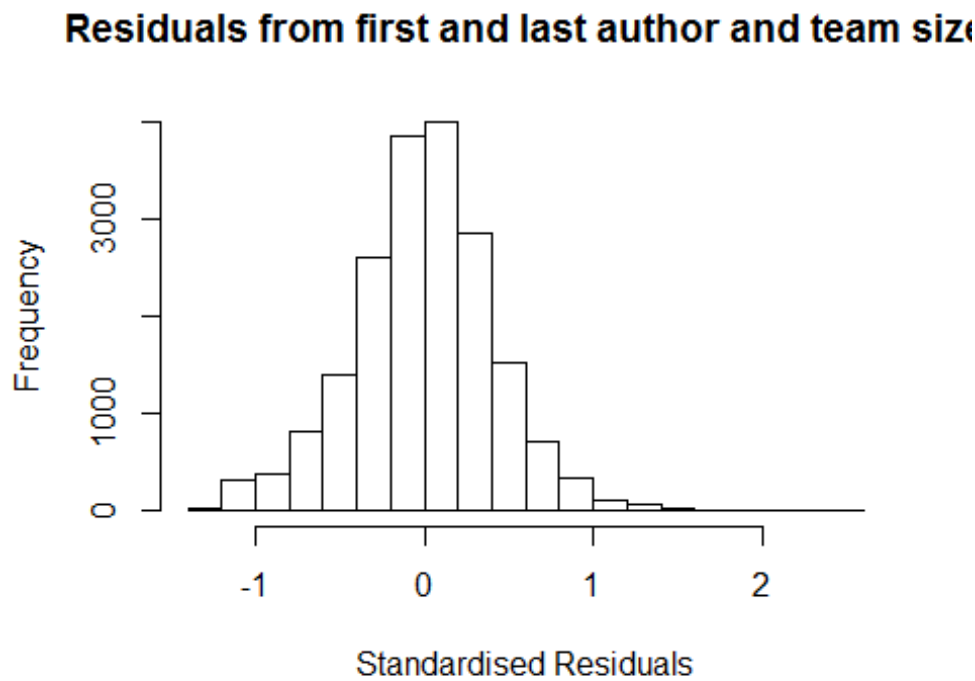
```
## 1327 1186 1219 1201 1020  916 1433 1318 1372 1400 1318 1494 1569 1420 1489
## 2011 2012
## 1515 1400
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 1164 1034 1058 1009  879  780 1241 1115 1146 1164 1104 1217 1320 1160 1201
## 2011 2012
## 1241 1148
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
##  Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 370, df = 16, p-value <2e-16
```




```
##
##  Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 1.1, df = 1, p-value = 0.3
```

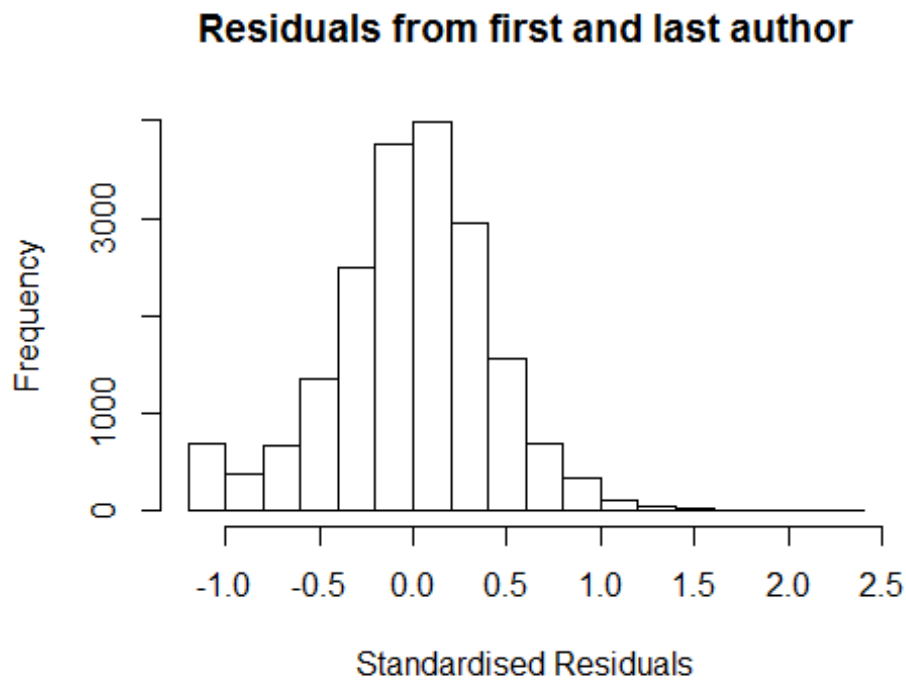



```
## [1] "Female first author team size 2018 geometric mean: 5.12380128923928"
## [1] "Male first author team size 2018 geometric mean: 5.14957084270974"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 64000, p-value = 0.8
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 4.85492199353248"
## [1] "Male last author team size 2018 geometric mean: 5.26792237807183"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 51000, p-value = 0.07
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##                     GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.017  1           1.008
## LastAuthorFemale  1.009  1           1.004
## UniqueAuthors     1.070  4           1.008
## Year              1.078 16           1.002
```



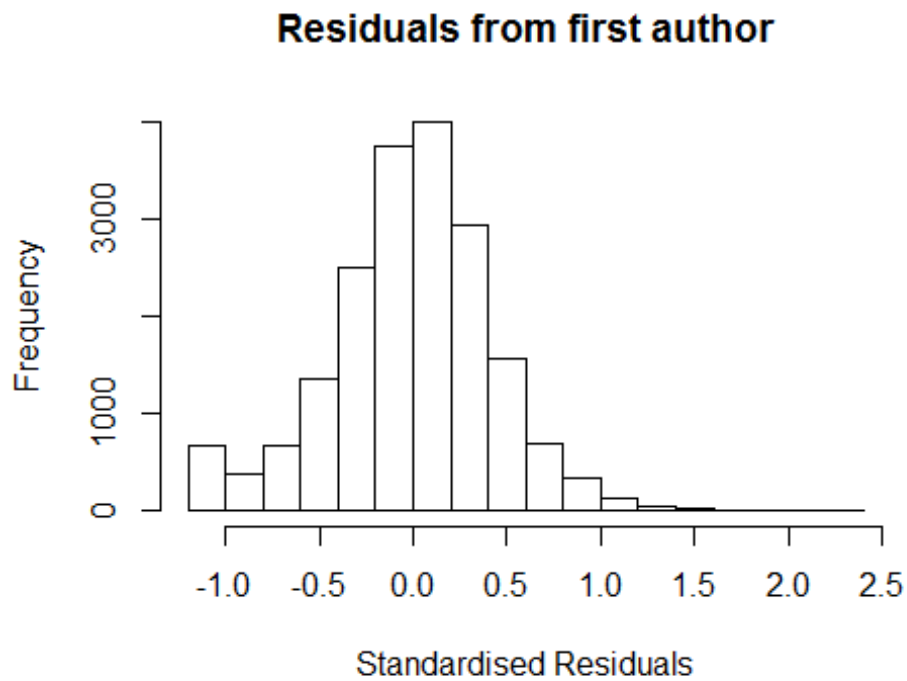

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
UniqueAuthors +
##     Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##     k.max = 1000))
##  \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.24994 -0.24785  0.00447  0.25235  2.45155
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         0.75899    0.02496   30.40   <2e-16 ***
## FirstAuthorFemale1  0.01876    0.00630    2.98   0.0029 **
## LastAuthorFemale1   0.00349    0.00730    0.48   0.6328
## UniqueAuthors2      0.22662    0.02155   10.51   <2e-16 ***
## UniqueAuthors3      0.29414    0.02095   14.04   <2e-16 ***
## UniqueAuthors4      0.33956    0.02081   16.32   <2e-16 ***
## UniqueAuthors5      0.42186    0.01990   21.20   <2e-16 ***
## Year1997            0.04684    0.02297    2.04   0.0414 *
## Year1998            0.04035    0.02129    1.90   0.0581 .
## Year1999            0.00301    0.02166    0.14   0.8895
```

```
## Year2000            0.02243    0.02117    1.06   0.2894
## Year2001            0.01941    0.02104    0.92   0.3563
## Year2002           -0.02918    0.02003   -1.46   0.1451
## Year2003           -0.05534    0.02029   -2.73   0.0064 **
## Year2004           -0.04018    0.01976   -2.03   0.0420 *
## Year2005           -0.05845    0.01989   -2.94   0.0033 **
## Year2006           -0.05798    0.01999   -2.90   0.0037 **
## Year2007            0.01466    0.02058    0.71   0.4763
## Year2008            0.00746    0.01971    0.38   0.7050
## Year2009           -0.00377    0.01983   -0.19   0.8491
## Year2010           -0.03275    0.01981   -1.65   0.0983 .
## Year2011           -0.02251    0.01997   -1.13   0.2595
## Year2012           -0.05164    0.02065   -2.50   0.0124 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.373
## Multiple R-squared:  0.0768, Adjusted R-squared:  0.0758
## Convergence in 16 IRWLS iterations
##
## Robustness weights:
##  9 observations c(1237,7985,8371,11034,11340,11482,12154,13910,18031)
##   are outliers with |weight| = 0 ( < 5.3e-06);
##  1697 weights are ~= 1. The remaining 17275 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0006  0.8570  0.9500  0.8900  0.9860  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          5.27e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.009  1           1.004
## LastAuthorFemale  1.009  1           1.004
## Year              1.014 16           1.000
```



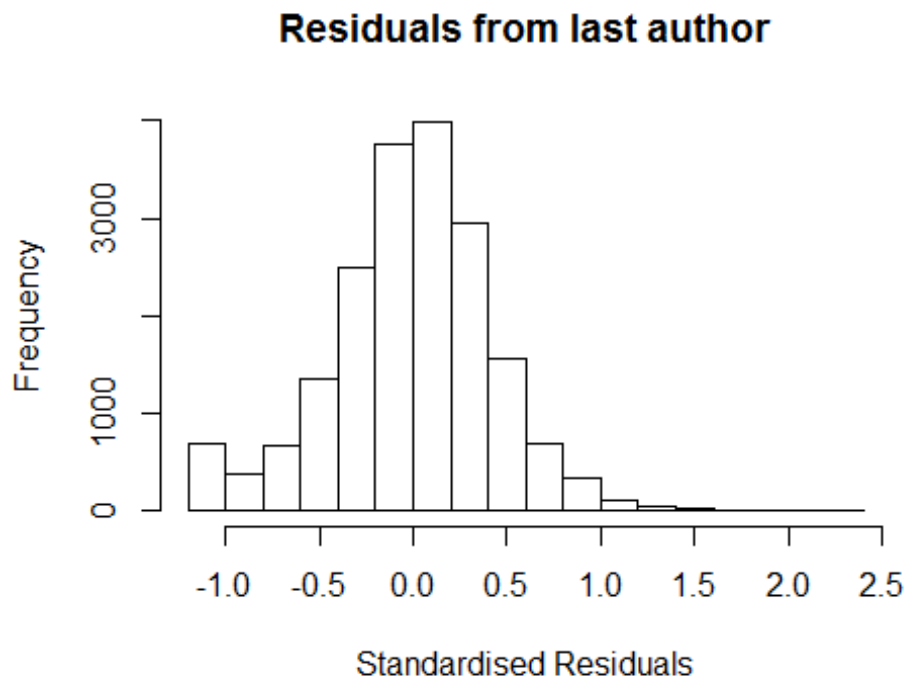

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.15196 -0.25488  0.00718  0.25468  2.36555
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.06839    0.01752   60.99   <2e-16 ***
## FirstAuthorFemale1  0.02278    0.00643    3.54   0.0004 ***
## LastAuthorFemale1  -0.00177    0.00755   -0.23   0.8142
## Year1997            0.05006    0.02371    2.11   0.0348 *
## Year1998            0.05981    0.02227    2.69   0.0073 **
## Year1999            0.01294    0.02244    0.58   0.5643
## Year2000            0.03577    0.02199    1.63   0.1038
## Year2001            0.03792    0.02195    1.73   0.0841 .
## Year2002           -0.00351    0.02070   -0.17   0.8655
## Year2003           -0.02277    0.02105   -1.08   0.2794
## Year2004           -0.00138    0.02061   -0.07   0.9465
## Year2005           -0.01729    0.02072   -0.83   0.4039
```

```
## Year2006           -0.01208    0.02085   -0.58   0.5622
## Year2007            0.06079    0.02119    2.87   0.0041 **
## Year2008            0.04899    0.02050    2.39   0.0169 *
## Year2009            0.05543    0.02048    2.71   0.0068 **
## Year2010            0.02097    0.02049    1.02   0.3061
## Year2011            0.02007    0.02068    0.97   0.3319
## Year2012            0.00121    0.02140    0.06   0.9548
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.38
## Multiple R-squared:  0.00543,    Adjusted R-squared:  0.00449
## Convergence in 16 IRWLS iterations
##
## Robustness weights:
##  6 observations c(1237,8371,11034,12154,13910,18031)
##   are outliers with |weight| = 0 ( < 5.3e-06);
##  1659 weights are ~= 1. The remaining 17316 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0063  0.8590  0.9500  0.8870  0.9860  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          5.27e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.007  1           1.004
## Year              1.007 16           1.000
```




```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.15148 -0.25562  0.00739  0.25493  2.36584
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)          1.06804    0.01745   61.21  < 2e-16 ***
## FirstAuthorFemale1   0.02263    0.00644    3.51  0.00044 ***
## Year1997             0.05012    0.02371    2.11  0.03457 *
## Year1998             0.05983    0.02228    2.69  0.00724 **
## Year1999             0.01294    0.02244    0.58  0.56423
## Year2000             0.03578    0.02199    1.63  0.10368
## Year2001             0.03792    0.02195    1.73  0.08403 .
## Year2002            -0.00350    0.02070   -0.17  0.86573
## Year2003            -0.02269    0.02105   -1.08  0.28104
## Year2004            -0.00134    0.02060   -0.07  0.94812
## Year2005            -0.01726    0.02072   -0.83  0.40473
## Year2006            -0.01206    0.02085   -0.58  0.56309
```

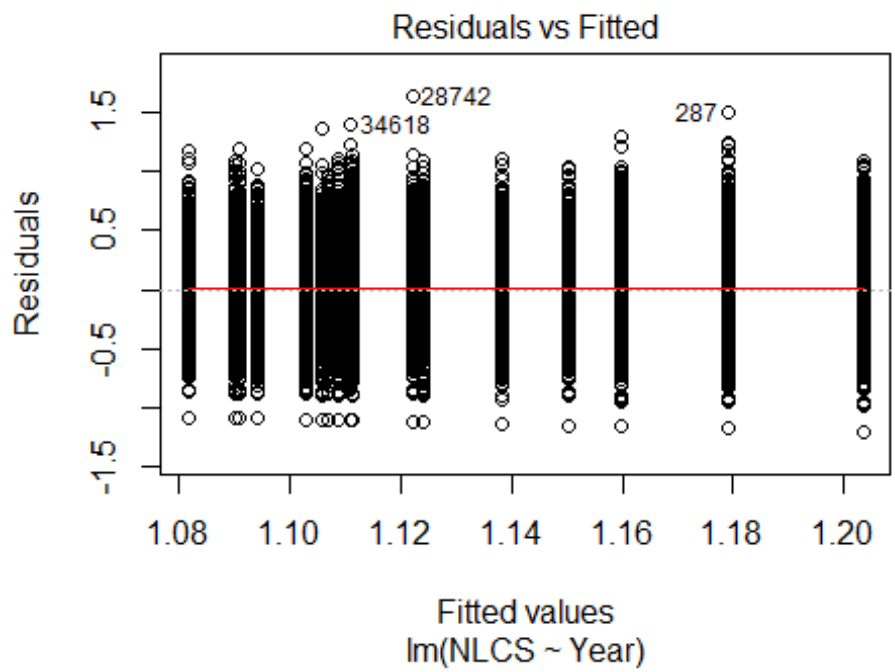
```
## Year2007            0.06081    0.02119    2.87  0.00412 **
## Year2008            0.04895    0.02050    2.39  0.01695 *
## Year2009            0.05545    0.02048    2.71  0.00678 **
## Year2010            0.02100    0.02048    1.03  0.30536
## Year2011            0.02003    0.02068    0.97  0.33293
## Year2012            0.00117    0.02139    0.05  0.95642
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.38
## Multiple R-squared:  0.00543,    Adjusted R-squared:  0.00454
## Convergence in 16 IRWLS iterations
##
## Robustness weights:
##  6 observations c(1237,8371,11034,12154,13910,18031)
##   are outliers with |weight| = 0 ( < 5.3e-06);
##  1658 weights are ~= 1. The remaining 17317 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0063  0.8590  0.9500  0.8870  0.9860  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          5.27e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##                   GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.007  1           1.004
## Year             1.007 16           1.000
```



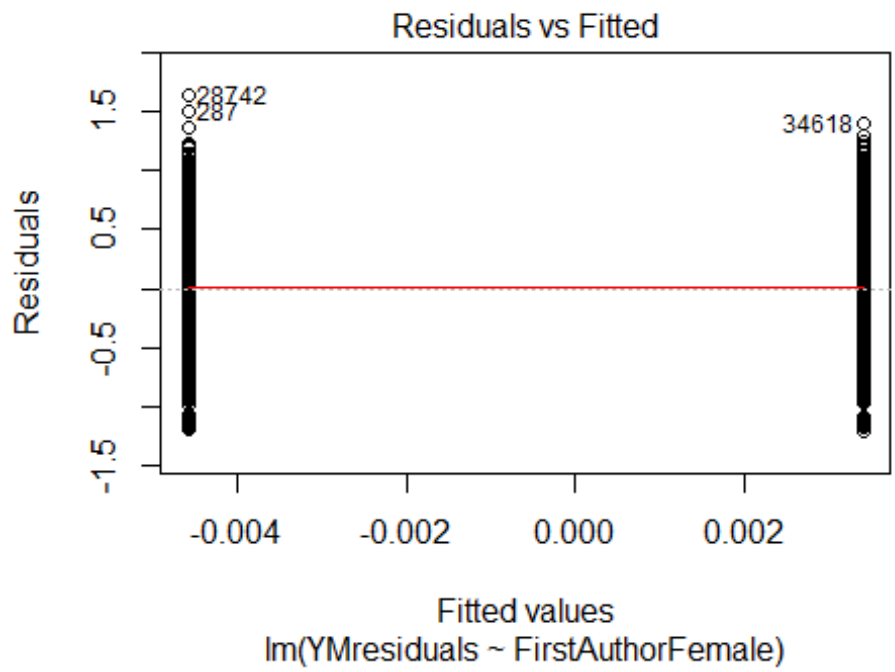

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.13737 -0.25570  0.00863  0.25594  2.35965
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)        1.074231   0.017415   61.68   <2e-16 ***
## LastAuthorFemale1  0.000673   0.007551    0.09   0.9290
## Year1997           0.050115   0.023719    2.11   0.0346 *
## Year1998           0.059826   0.022270    2.69   0.0072 **
## Year1999           0.013648   0.022410    0.61   0.5425
## Year2000           0.036339   0.021979    1.65   0.0983 .
## Year2001           0.038854   0.021948    1.77   0.0767 .
## Year2002          -0.002615   0.020697   -0.13   0.8995
## Year2003          -0.021534   0.021032   -1.02   0.3059
## Year2004          -0.000845   0.020594   -0.04   0.9673
## Year2005          -0.016048   0.020694   -0.78   0.4381
## Year2006          -0.010687   0.020830   -0.51   0.6079
```

```
## Year2007           0.062467   0.021169    2.95   0.0032 **
## Year2008           0.050611   0.020492    2.47   0.0135 *
## Year2009           0.057190   0.020484    2.79   0.0052 **
## Year2010           0.023300   0.020463    1.14   0.2549
## Year2011           0.022189   0.020680    1.07   0.2833
## Year2012           0.003217   0.021387    0.15   0.8804
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.381
## Multiple R-squared:  0.00472,    Adjusted R-squared:  0.00383
## Convergence in 16 IRWLS iterations
##
## Robustness weights:
##  6 observations c(1237,8371,11034,12154,13910,18031)
##   are outliers with |weight| = 0 ( < 5.3e-06);
##  1677 weights are ~= 1. The remaining 17298 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0075  0.8570  0.9500  0.8870  0.9860  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          5.27e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                  psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Sample size for the above analysis:  18981"
## [1] ""
## [1] ""
## [1] "##################################"
## [1] "Analysis of AJSC 1309"
## [1] "##################################"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 2160 1877 1819 1743 1678 1586 1543 1613 1628 1682 1964 1983 2046 1924 1827
## 2011 2012
## 1715 1794
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
```

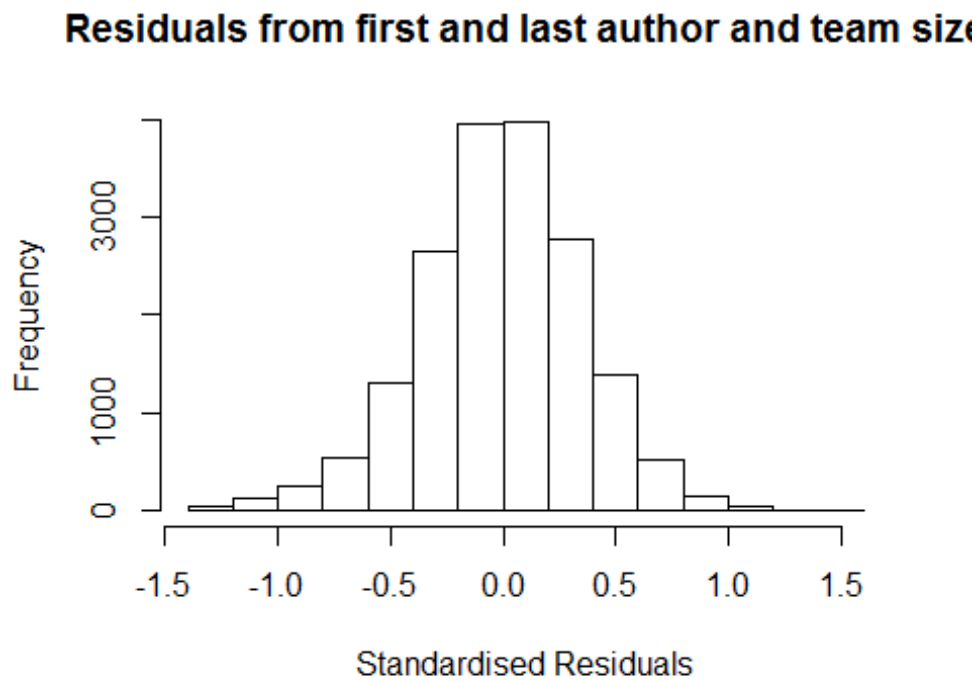
```
## 1444 1228 1238 1192 1027  838 1044 1076 1116 1137 1342 1348 1379 1251 1241
## 2011 2012
## 1140 1211
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 1268 1078 1097 1050  896  726  927  942  981  993 1170 1152 1192 1089 1080
## 2011 2012
##  985 1053
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
##  Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 230, df = 16, p-value <2e-16
```




```
##
##  Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 27, df = 1, p-value = 2e-07
```

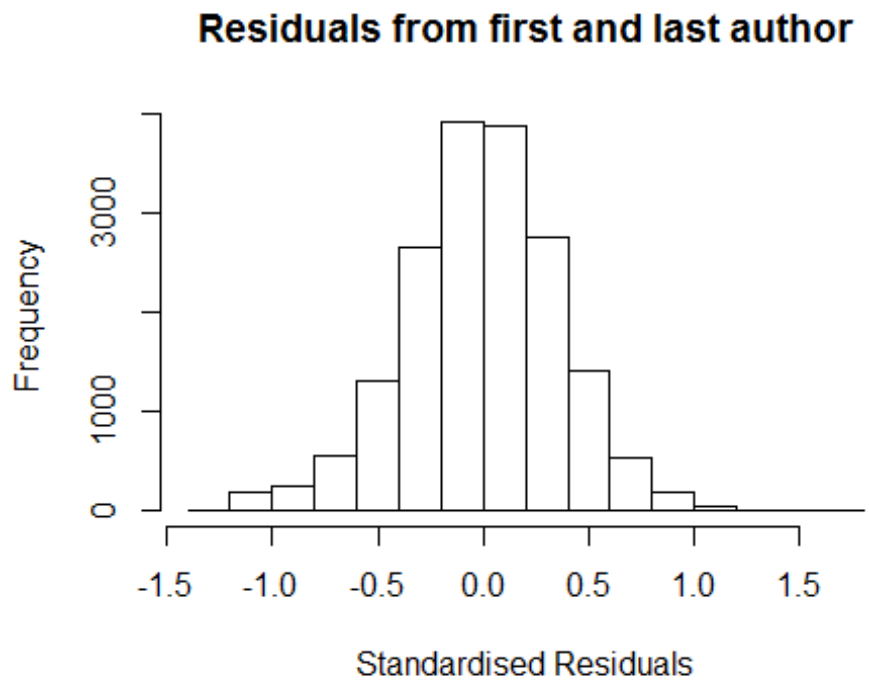



```
## [1] "Female first author team size 2018 geometric mean: 4.8797147752495"
## [1] "Male first author team size 2018 geometric mean: 4.40864544703201"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 95000, p-value = 0.06
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 4.60106698046849"
## [1] "Male last author team size 2018 geometric mean: 4.67302135864264"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 82000, p-value = 0.7
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.028  1           1.014
## LastAuthorFemale  1.022  1           1.011
## UniqueAuthors     1.073  4           1.009
## Year              1.075 16           1.002
```



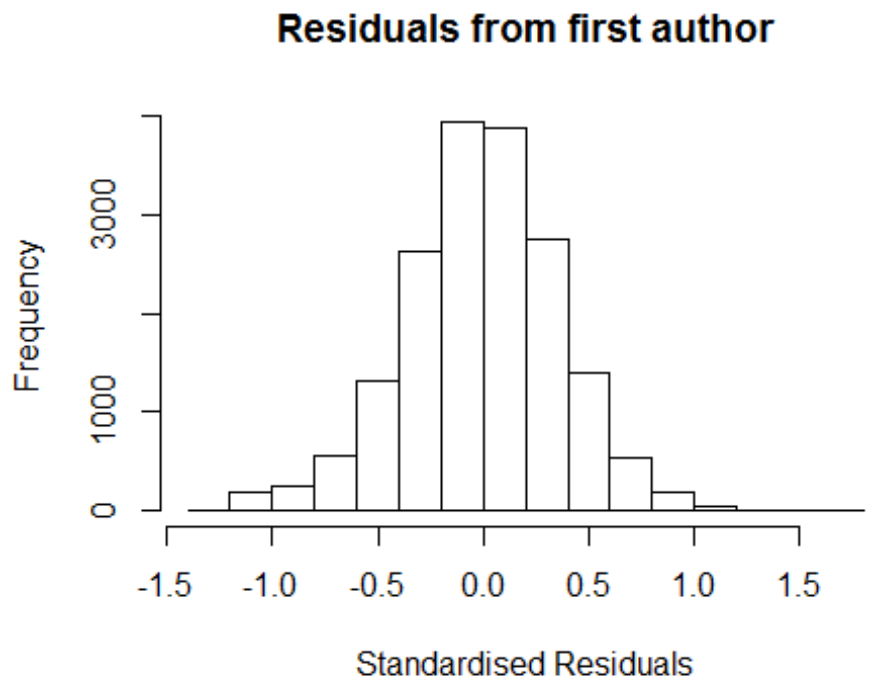

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
UniqueAuthors +
##     Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##     k.max = 1000))
##  \--> method = "MM"
## Residuals:
##       Min        1Q    Median        3Q       Max
## -1.314275 -0.226636  0.000562  0.224622  1.561883
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.06410    0.01773   60.01  < 2e-16 ***
## FirstAuthorFemale1 -0.01318    0.00543   -2.43   0.0152 *
## LastAuthorFemale1  -0.00899    0.00601   -1.50   0.1344
## UniqueAuthors2      0.09687    0.01426    6.79  1.1e-11 ***
## UniqueAuthors3      0.12340    0.01426    8.65  < 2e-16 ***
## UniqueAuthors4      0.14747    0.01454   10.14  < 2e-16 ***
## UniqueAuthors5      0.25018    0.01399   17.88  < 2e-16 ***
## Year1997            0.01709    0.01740    0.98   0.3259
## Year1998           -0.01825    0.01747   -1.04   0.2961
## Year1999           -0.04708    0.01651   -2.85   0.0044 **
```

```
## Year2000           -0.08880    0.01721   -5.16  2.5e-07 ***
## Year2001           -0.12212    0.01769   -6.90  5.3e-12 ***
## Year2002           -0.08476    0.01744   -4.86  1.2e-06 ***
## Year2003           -0.11615    0.01654   -7.02  2.3e-12 ***
## Year2004           -0.10910    0.01664   -6.56  5.6e-11 ***
## Year2005           -0.11900    0.01648   -7.22  5.4e-13 ***
## Year2006           -0.12746    0.01598   -7.98  1.6e-15 ***
## Year2007           -0.10706    0.01580   -6.77  1.3e-11 ***
## Year2008           -0.10061    0.01595   -6.31  2.9e-10 ***
## Year2009           -0.11584    0.01644   -7.05  1.9e-12 ***
## Year2010           -0.10498    0.01685   -6.23  4.8e-10 ***
## Year2011           -0.07124    0.01668   -4.27  2.0e-05 ***
## Year2012           -0.10458    0.01675   -6.24  4.4e-10 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.338
## Multiple R-squared:  0.0534, Adjusted R-squared:  0.0522
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
##  1569 weights are ~= 1. The remaining 16110 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0006  0.8670  0.9500  0.8980  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          5.66e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.021  1           1.011
## LastAuthorFemale  1.020  1           1.010
## Year              1.012 16           1.000
```



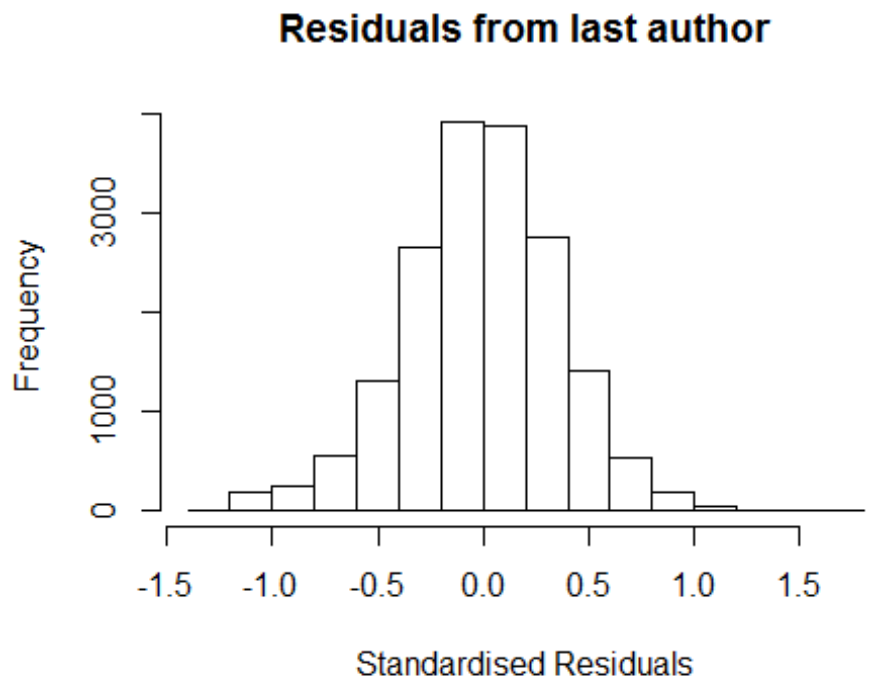

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.21446 -0.23147 -0.00105  0.23105  1.63150
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)          1.19659    0.01296   92.35  < 2e-16 ***
## FirstAuthorFemale1 -0.00762    0.00551   -1.38  0.16684
## LastAuthorFemale1  -0.01458    0.00611   -2.39  0.01706 *
## Year1997            0.01787    0.01767    1.01  0.31203
## Year1998           -0.01450    0.01766   -0.82  0.41154
## Year1999           -0.03801    0.01682   -2.26  0.02382 *
## Year2000           -0.07298    0.01739   -4.20  2.7e-05 ***
## Year2001           -0.10503    0.01782   -5.89  3.8e-09 ***
## Year2002           -0.06763    0.01773   -3.81  0.00014 ***
## Year2003           -0.09610    0.01674   -5.74  9.5e-09 ***
## Year2004           -0.08302    0.01694   -4.90  9.6e-07 ***
## Year2005           -0.09649    0.01674   -5.77  8.3e-09 ***
```

```
## Year2006           -0.10294    0.01630   -6.32  2.8e-10 ***
## Year2007           -0.07899    0.01601   -4.93  8.1e-07 ***
## Year2008           -0.06702    0.01612   -4.16  3.2e-05 ***
## Year2009           -0.07791    0.01666   -4.68  2.9e-06 ***
## Year2010           -0.06247    0.01693   -3.69  0.00023 ***
## Year2011           -0.03000    0.01705   -1.76  0.07846 .
## Year2012           -0.05873    0.01700   -3.46  0.00055 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.344
## Multiple R-squared:  0.0107, Adjusted R-squared:  0.00971
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  observation 14642 is an outlier with |weight| = 0 ( < 5.7e-06);
##  1498 weights are ~= 1. The remaining 16180 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0197  0.8680  0.9500  0.8980  0.9860  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          5.66e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.007  1           1.003
## Year              1.007 16           1.000
```




```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.21143 -0.23081 -0.00186  0.23131  1.63640
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.19429    0.01292   92.41  < 2e-16 ***
## FirstAuthorFemale1 -0.00930    0.00548   -1.70  0.08957 .
## Year1997            0.01714    0.01768    0.97  0.33236
## Year1998           -0.01505    0.01766   -0.85  0.39404
## Year1999           -0.03843    0.01682   -2.28  0.02233 *
## Year2000           -0.07383    0.01740   -4.24  2.2e-05 ***
## Year2001           -0.10609    0.01781   -5.96  2.6e-09 ***
## Year2002           -0.06899    0.01773   -3.89  0.00010 ***
## Year2003           -0.09704    0.01674   -5.80  6.8e-09 ***
## Year2004           -0.08405    0.01693   -4.96  7.0e-07 ***
## Year2005           -0.09718    0.01673   -5.81  6.5e-09 ***
## Year2006           -0.10419    0.01628   -6.40  1.6e-10 ***
```

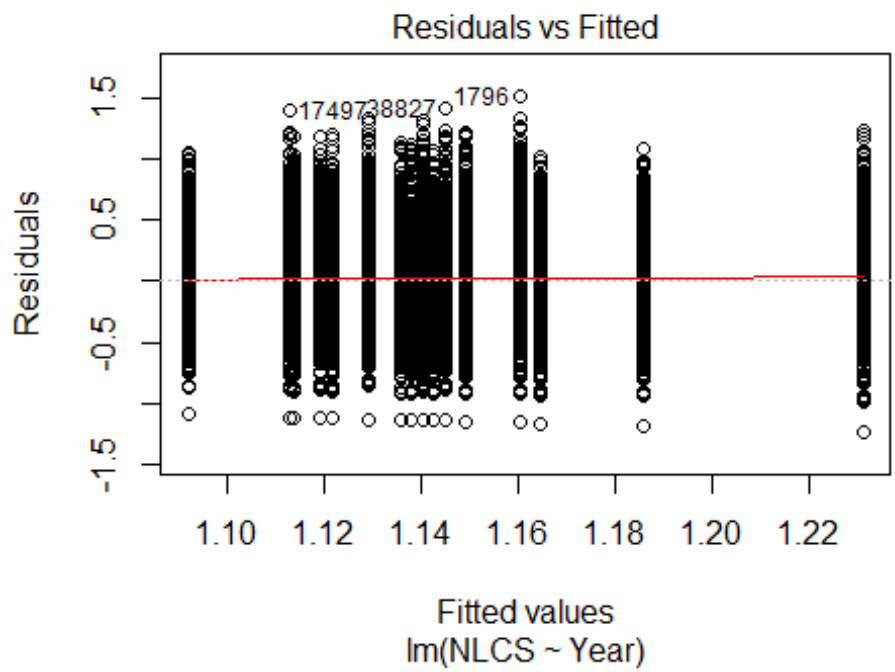
```
## Year2007           -0.07991    0.01601   -4.99  6.0e-07 ***
## Year2008           -0.06868    0.01610   -4.27  2.0e-05 ***
## Year2009           -0.07897    0.01665   -4.74  2.1e-06 ***
## Year2010           -0.06339    0.01693   -3.75  0.00018 ***
## Year2011           -0.03147    0.01703   -1.85  0.06467 .
## Year2012           -0.06013    0.01699   -3.54  0.00040 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.344
## Multiple R-squared:  0.0104, Adjusted R-squared:  0.00944
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  observation 14642 is an outlier with |weight| = 0 ( < 5.7e-06);
##  1528 weights are ~= 1. The remaining 16150 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0184  0.8680  0.9500  0.8980  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          5.66e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##                   GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.006  1           1.003
## Year             1.006 16           1.000
```



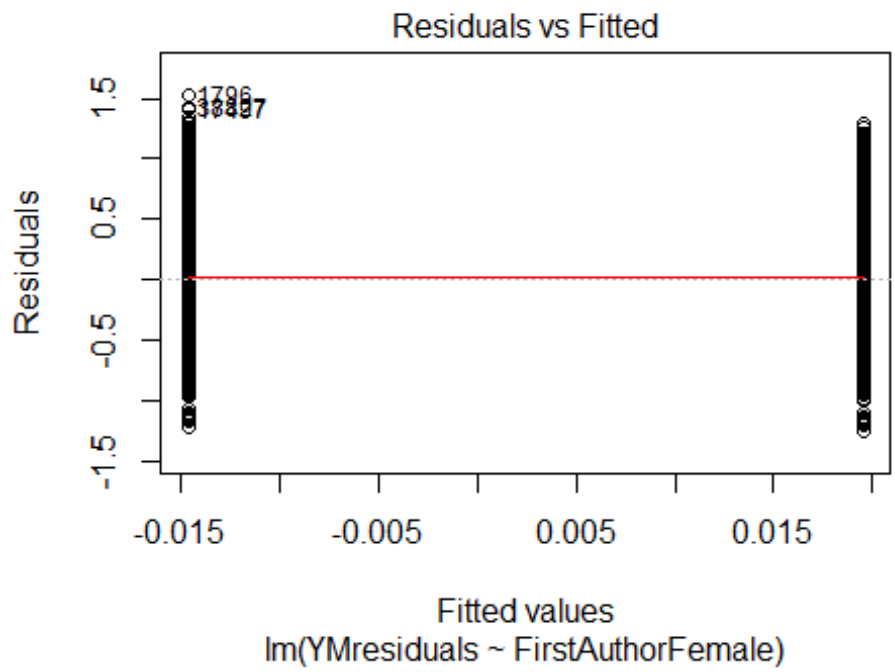

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##       Min        1Q    Median        3Q       Max
## -1.211563 -0.232563 -0.000898  0.231102  1.627102
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.19380    0.01281   93.22  < 2e-16 ***
## LastAuthorFemale1 -0.01566    0.00607   -2.58  0.00986 **
## Year1997            0.01777    0.01767    1.01  0.31452
## Year1998           -0.01455    0.01765   -0.82  0.40978
## Year1999           -0.03812    0.01681   -2.27  0.02339 *
## Year2000           -0.07282    0.01739   -4.19  2.8e-05 ***
## Year2001           -0.10515    0.01781   -5.90  3.6e-09 ***
## Year2002           -0.06785    0.01773   -3.83  0.00013 ***
## Year2003           -0.09649    0.01673   -5.77  8.3e-09 ***
## Year2004           -0.08332    0.01693   -4.92  8.7e-07 ***
## Year2005           -0.09689    0.01673   -5.79  7.1e-09 ***
## Year2006           -0.10291    0.01630   -6.31  2.8e-10 ***
```

```
## Year2007          -0.07931    0.01600   -4.96  7.3e-07 ***
## Year2008          -0.06737    0.01611   -4.18  2.9e-05 ***
## Year2009          -0.07848    0.01664   -4.72  2.4e-06 ***
## Year2010          -0.06290    0.01693   -3.72  0.00020 ***
## Year2011          -0.03039    0.01705   -1.78  0.07469 .
## Year2012          -0.05932    0.01699   -3.49  0.00048 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.344
## Multiple R-squared:  0.0106, Adjusted R-squared:  0.00966
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  observation 14642 is an outlier with |weight| = 0 ( < 5.7e-06);
##  1533 weights are ~= 1. The remaining 16145 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0212  0.8680  0.9500  0.8980  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          5.66e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Sample size for the above analysis:  17679"
## [1] ""
## [1] ""
## [1] "####################################"
## [1] "Analysis of AJSC 1310"
## [1] "####################################"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 2135 2728 2135 2020 2228 2349 2048 1733 1929 2017 2149 2192 2118 1980 2102
## 2011 2012
## 2126 2053
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 1277 1293 1215 1142 1102 1041 1273 1070 1202 1306 1456 1472 1475 1349 1424
```

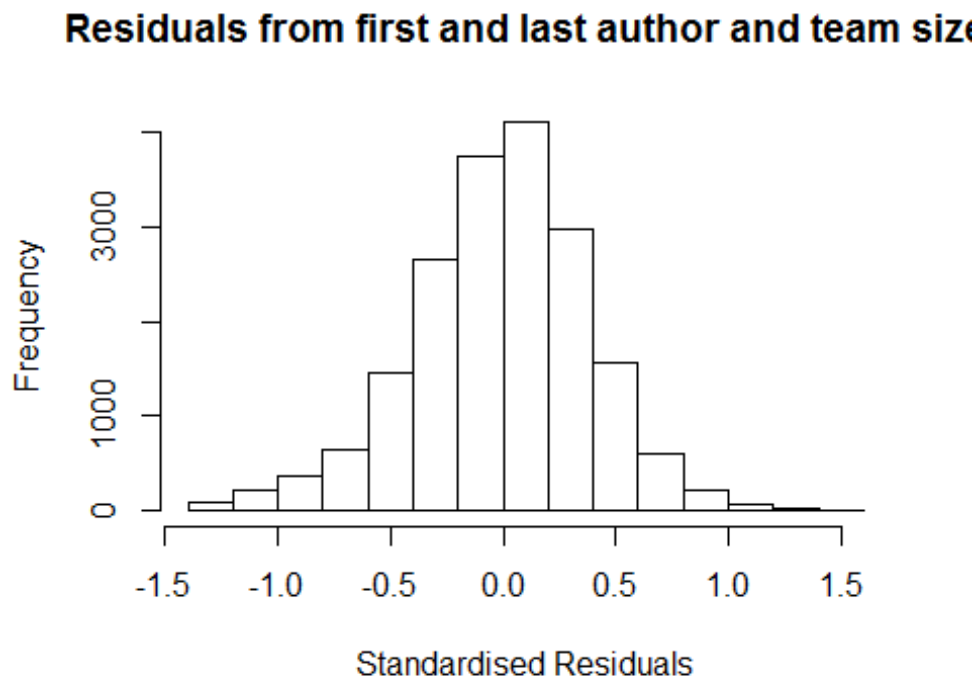
```
## 2011 2012
## 1455 1385
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 1089 1142 1048  959  934  891 1080  907 1029 1086 1224 1238 1274 1164 1195
## 2011 2012
## 1269 1196
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
##  Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 240, df = 16, p-value <2e-16
```




```
##
##  Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 23, df = 1, p-value = 2e-06
```

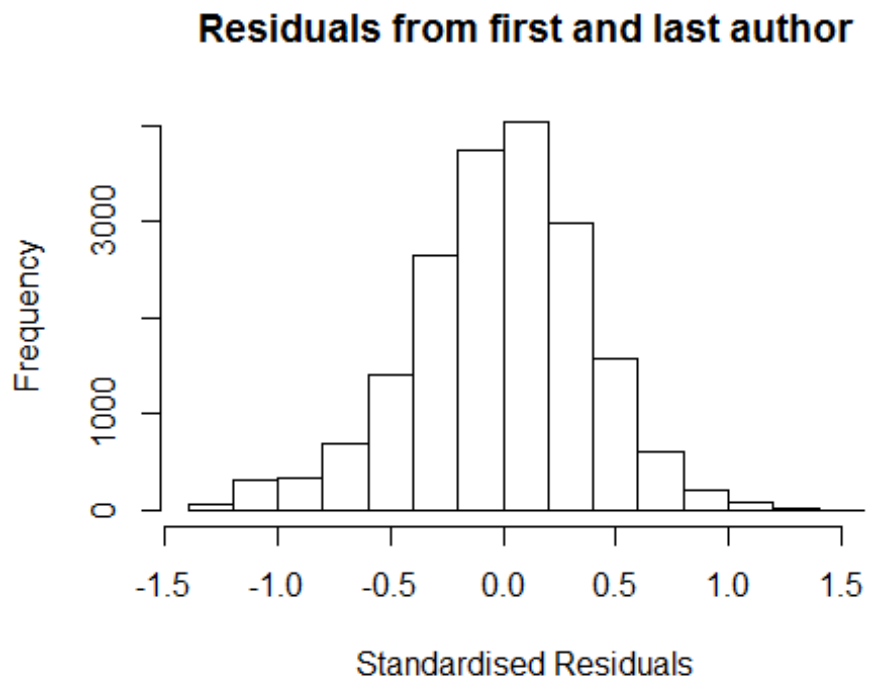



```
## [1] "Female first author team size 2018 geometric mean: 5.13619375292364"
## [1] "Male first author team size 2018 geometric mean: 5.12438491140573"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 110000, p-value = 0.5
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 5.05273781913426"
## [1] "Male last author team size 2018 geometric mean: 5.18661963739662"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 110000, p-value = 0.4
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##                     GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.037  1           1.018
## LastAuthorFemale  1.035  1           1.017
## UniqueAuthors     1.050  4           1.006
## Year              1.082 16           1.002
```



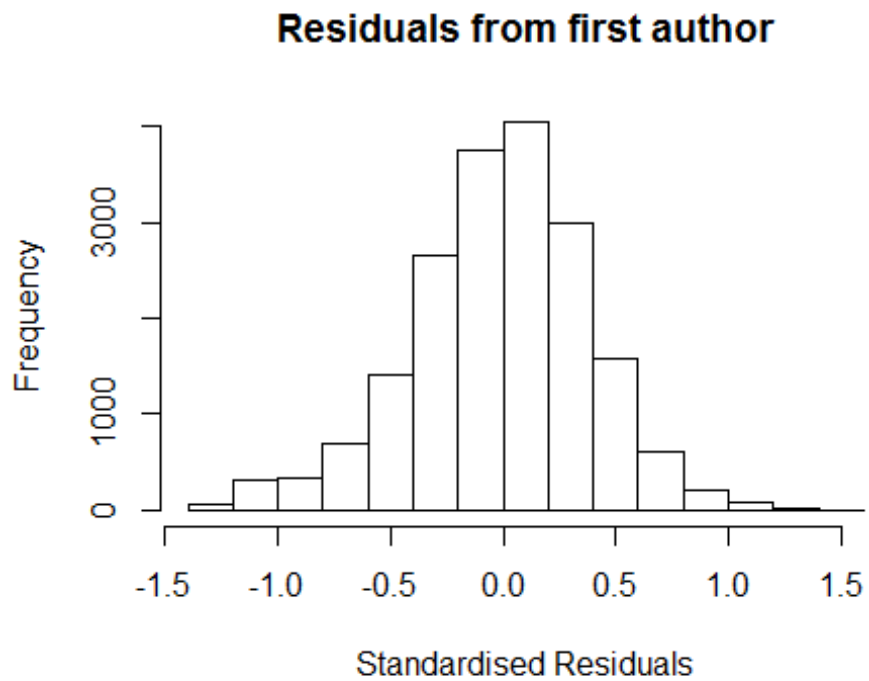

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
UniqueAuthors +
##     Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##     k.max = 1000))
##  \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.39436 -0.24803  0.00887  0.24118  1.51726
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.02343    0.01974   51.84  < 2e-16 ***
## FirstAuthorFemale1  0.02475    0.00572    4.33  1.5e-05 ***
## LastAuthorFemale1   0.00950    0.00632    1.50  0.13263
## UniqueAuthors2      0.08260    0.01660    4.97  6.6e-07 ***
## UniqueAuthors3      0.12971    0.01617    8.02  1.1e-15 ***
## UniqueAuthors4      0.16682    0.01614   10.33  < 2e-16 ***
## UniqueAuthors5      0.25604    0.01537   16.66  < 2e-16 ***
## Year1997            0.09013    0.01935    4.66  3.2e-06 ***
## Year1998           -0.00545    0.01814   -0.30  0.76362
## Year1999           -0.04238    0.01827   -2.32  0.02037 *
```

```
## Year2000           -0.04588    0.01859   -2.47  0.01359 *
## Year2001           -0.04387    0.01816   -2.42  0.01573 *
## Year2002           -0.07229    0.01810   -3.99  6.5e-05 ***
## Year2003           -0.11026    0.01847   -5.97  2.4e-09 ***
## Year2004           -0.09009    0.01768   -5.10  3.5e-07 ***
## Year2005           -0.08408    0.01753   -4.80  1.6e-06 ***
## Year2006           -0.07904    0.01745   -4.53  5.9e-06 ***
## Year2007           -0.07214    0.01709   -4.22  2.4e-05 ***
## Year2008           -0.07042    0.01739   -4.05  5.2e-05 ***
## Year2009           -0.03754    0.01774   -2.12  0.03433 *
## Year2010           -0.08161    0.01739   -4.69  2.7e-06 ***
## Year2011           -0.06629    0.01764   -3.76  0.00017 ***
## Year2012           -0.09443    0.01810   -5.22  1.8e-07 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.362
## Multiple R-squared:  0.0558, Adjusted R-squared:  0.0547
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
##  1579 weights are ~= 1. The remaining 17146 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0399  0.8670  0.9500  0.8950  0.9860  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          5.34e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.036  1           1.018
## LastAuthorFemale  1.034  1           1.017
## Year              1.033 16           1.001
```



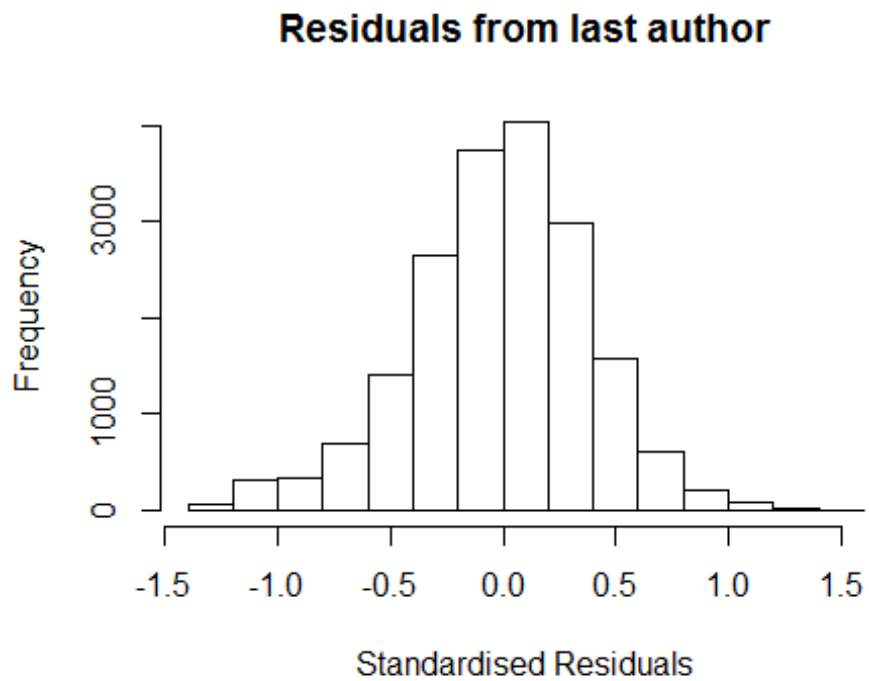

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.3048 -0.2491  0.0067  0.2430  1.4972
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)          1.17381    0.01431   82.01  < 2e-16 ***
## FirstAuthorFemale1   0.03292    0.00582    5.66  1.6e-08 ***
## LastAuthorFemale1    0.00238    0.00640    0.37  0.71004
## Year1997             0.09569    0.01979    4.84  1.3e-06 ***
## Year1998             0.00809    0.01852    0.44  0.66204
## Year1999            -0.02939    0.01872   -1.57  0.11647
## Year2000            -0.03036    0.01911   -1.59  0.11201
## Year2001            -0.02756    0.01855   -1.49  0.13736
## Year2002            -0.05237    0.01848   -2.83  0.00461 **
## Year2003            -0.09026    0.01909   -4.73  2.3e-06 ***
## Year2004            -0.06034    0.01805   -3.34  0.00083 ***
## Year2005            -0.05331    0.01793   -2.97  0.00295 **
```

```
## Year2006           -0.05317    0.01779   -2.99  0.00280 **
## Year2007           -0.03724    0.01748   -2.13  0.03314 *
## Year2008           -0.03370    0.01775   -1.90  0.05762 .
## Year2009            0.00480    0.01791    0.27  0.78863
## Year2010           -0.03874    0.01776   -2.18  0.02913 *
## Year2011           -0.02881    0.01796   -1.60  0.10861
## Year2012           -0.04881    0.01851   -2.64  0.00839 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.368
## Multiple R-squared:  0.0118, Adjusted R-squared:  0.0109
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
##  1544 weights are ~= 1. The remaining 17181 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0598  0.8660  0.9510  0.8940  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          5.34e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##                   GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.02  1           1.010
## Year              1.02 16           1.001
```




```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.30319 -0.24914  0.00679  0.24310  1.49679
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)          1.17421    0.01428   82.24  < 2e-16 ***
## FirstAuthorFemale1   0.03324    0.00577    5.76  8.7e-09 ***
## Year1997             0.09573    0.01978    4.84  1.3e-06 ***
## Year1998             0.00807    0.01851    0.44  0.66286
## Year1999            -0.02931    0.01872   -1.57  0.11729
## Year2000            -0.03040    0.01910   -1.59  0.11152
## Year2001            -0.02753    0.01855   -1.48  0.13780
## Year2002            -0.05223    0.01847   -2.83  0.00469 **
## Year2003            -0.09015    0.01909   -4.72  2.3e-06 ***
## Year2004            -0.06023    0.01804   -3.34  0.00084 ***
## Year2005            -0.05324    0.01792   -2.97  0.00297 **
## Year2006            -0.05303    0.01778   -2.98  0.00287 **
```

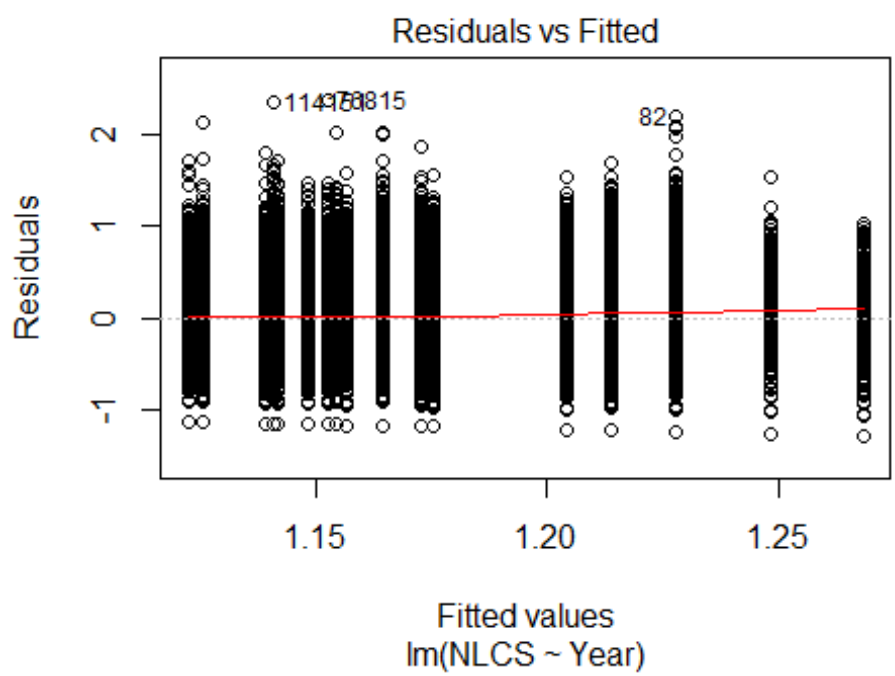
```
## Year2007           -0.03707    0.01747   -2.12  0.03384 *
## Year2008           -0.03355    0.01774   -1.89  0.05861 .
## Year2009            0.00498    0.01790    0.28  0.78062
## Year2010           -0.03851    0.01773   -2.17  0.02992 *
## Year2011           -0.02859    0.01794   -1.59  0.11097
## Year2012           -0.04852    0.01848   -2.63  0.00867 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.368
## Multiple R-squared:  0.0118, Adjusted R-squared:  0.0109
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
##  1541 weights are ~= 1. The remaining 17184 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0601  0.8660  0.9510  0.8940  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          5.34e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##                   GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.017  1           1.009
## Year             1.017 16           1.001
```



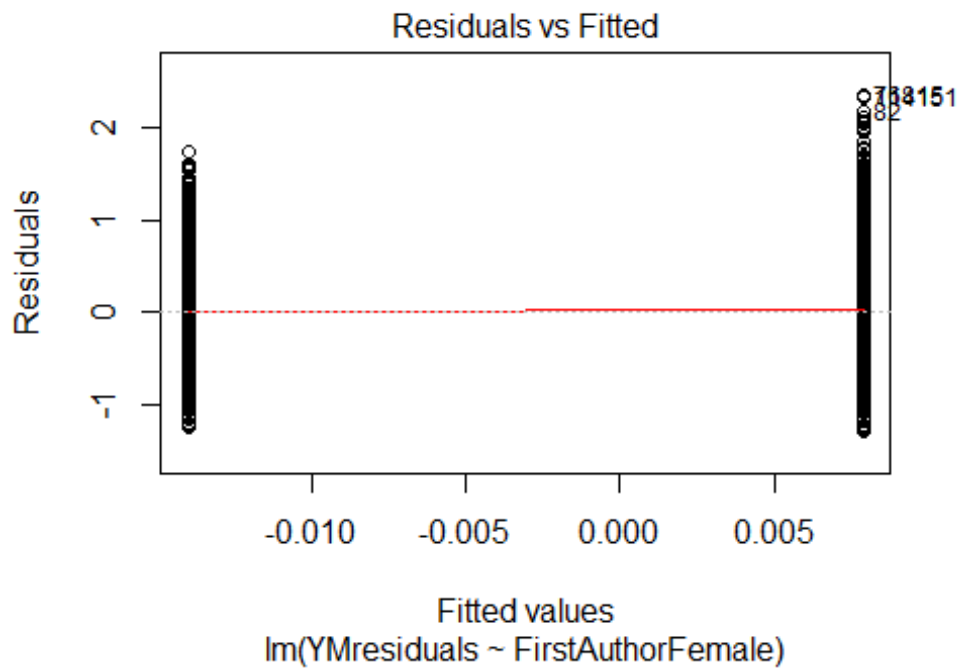

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.28904 -0.25057  0.00566  0.24315  1.48582
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.18518    0.01412   83.91  < 2e-16 ***
## LastAuthorFemale1   0.00775    0.00635    1.22  0.22234
## Year1997            0.09610    0.01977    4.86  1.2e-06 ***
## Year1998            0.00739    0.01850    0.40  0.68946
## Year1999           -0.03021    0.01870   -1.62  0.10613
## Year2000           -0.03042    0.01909   -1.59  0.11094
## Year2001           -0.02789    0.01855   -1.50  0.13285
## Year2002           -0.05112    0.01846   -2.77  0.00562 **
## Year2003           -0.08904    0.01907   -4.67  3.1e-06 ***
## Year2004           -0.05942    0.01804   -3.29  0.00099 ***
## Year2005           -0.05201    0.01793   -2.90  0.00372 **
## Year2006           -0.05141    0.01777   -2.89  0.00382 **
```

```
## Year2007          -0.03449    0.01747   -1.97  0.04835 *
## Year2008          -0.03084    0.01775   -1.74  0.08228 .
## Year2009           0.00855    0.01789    0.48  0.63264
## Year2010          -0.03533    0.01772   -1.99  0.04621 *
## Year2011          -0.02443    0.01794   -1.36  0.17318
## Year2012          -0.04428    0.01848   -2.40  0.01661 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.368
## Multiple R-squared:  0.0101, Adjusted R-squared:  0.00916
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
##  1575 weights are ~= 1. The remaining 17150 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.066   0.865   0.951   0.894   0.985   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          5.34e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Sample size for the above analysis:  18725"
## [1] ""
## [1] ""
## [1] "####################################"
## [1] "Analysis of AJSC 1311"
## [1] "####################################"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 9525 8293  806 5840  700 6624 6352 6027 6706 6556 6643 6578 6260 6704 6275
## 2011 2012
## 6463 6516
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 3934 3717  383 3254  338 2604 3772 3586 3980 3827 3832 3791 3729 3996 3736
## 2011 2012
```

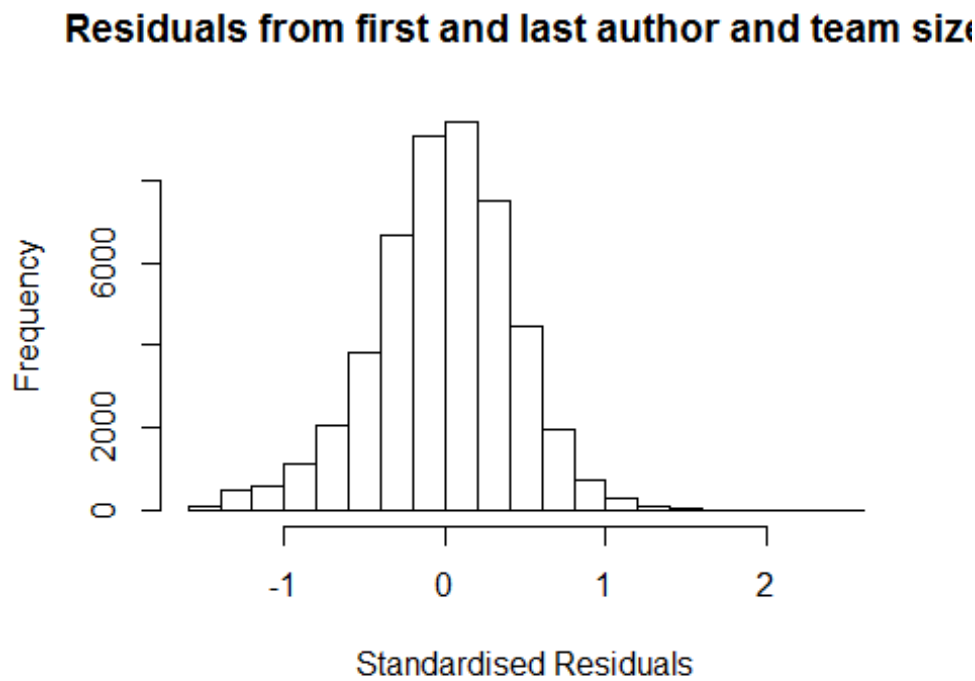
```
## 3934 3950
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 3368 3233  344 2809  299 2239 3264 3088 3407 3272 3281 3259 3193 3395 3210
## 2011 2012
## 3356 3387
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
##  Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 1600, df = 16, p-value <2e-16
```




```
##
##  Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 160, df = 1, p-value <2e-16
```

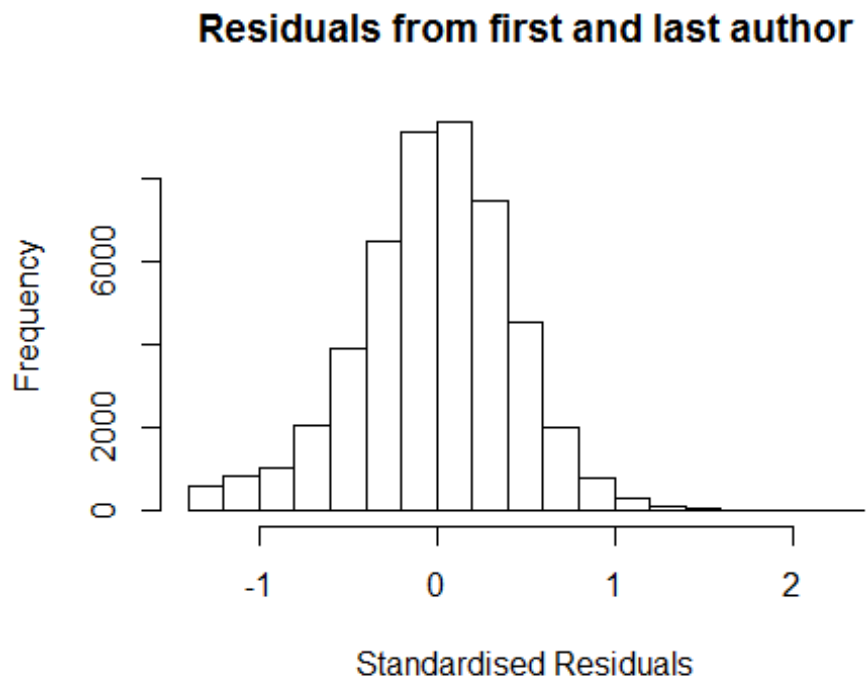



```
## [1] "Female first author team size 2018 geometric mean: 5.05320792379833"
## [1] "Male first author team size 2018 geometric mean: 4.46284818132151"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 1400000, p-value = 5e-08
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 4.66644504765311"
## [1] "Male last author team size 2018 geometric mean: 4.74240521712014"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 1100000, p-value = 0.9
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##                     GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.021  1           1.010
## LastAuthorFemale  1.018  1           1.009
## UniqueAuthors     1.036  4           1.004
## Year              1.041 16           1.001
```



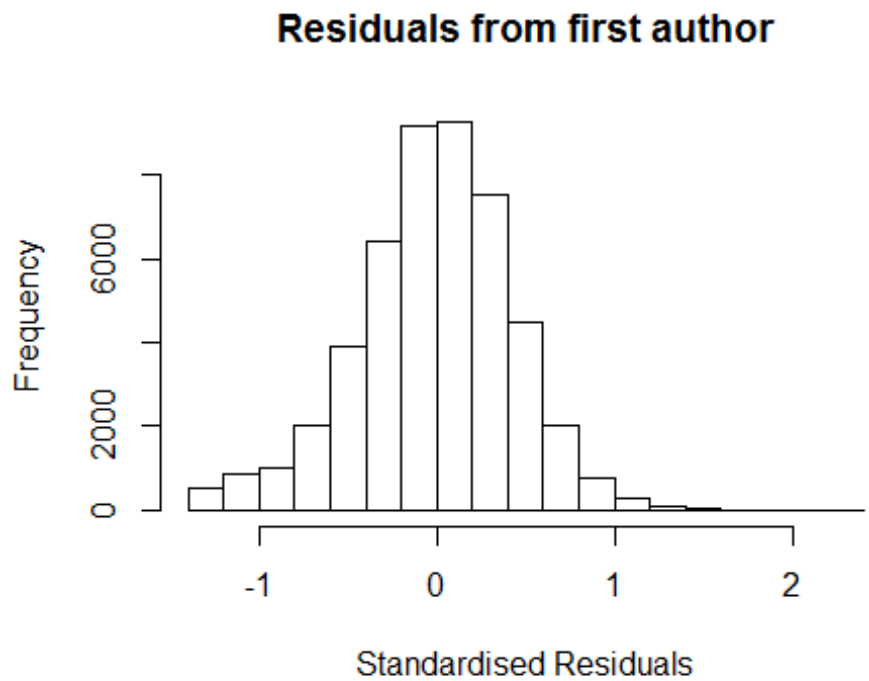

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
UniqueAuthors +
##     Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##     k.max = 1000))
##  \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.44282 -0.26885  0.00734  0.27091  2.40054
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.14244    0.01299   87.98  < 2e-16 ***
## FirstAuthorFemale1 -0.03280    0.00398   -8.24  < 2e-16 ***
## LastAuthorFemale1  -0.02796    0.00456   -6.14  8.4e-10 ***
## UniqueAuthors2      0.15877    0.00976   16.26  < 2e-16 ***
## UniqueAuthors3      0.18351    0.00980   18.72  < 2e-16 ***
## UniqueAuthors4      0.20167    0.01002   20.12  < 2e-16 ***
## UniqueAuthors5      0.29314    0.00934   31.38  < 2e-16 ***
## Year1997           -0.03745    0.01389   -2.70    0.007 **
## Year1998            0.00724    0.02992    0.24    0.809
## Year1999           -0.13414    0.01245  -10.77  < 2e-16 ***
```

```
## Year2000           -0.01542    0.02505   -0.62    0.538
## Year2001           -0.11657    0.01340   -8.70  < 2e-16 ***
## Year2002           -0.14749    0.01232  -11.97  < 2e-16 ***
## Year2003           -0.17886    0.01214  -14.74  < 2e-16 ***
## Year2004           -0.19677    0.01202  -16.37  < 2e-16 ***
## Year2005           -0.18366    0.01202  -15.28  < 2e-16 ***
## Year2006           -0.18011    0.01203  -14.97  < 2e-16 ***
## Year2007           -0.17969    0.01206  -14.89  < 2e-16 ***
## Year2008           -0.15412    0.01232  -12.51  < 2e-16 ***
## Year2009           -0.20226    0.01228  -16.47  < 2e-16 ***
## Year2010           -0.17239    0.01229  -14.03  < 2e-16 ***
## Year2011           -0.17394    0.01232  -14.12  < 2e-16 ***
## Year2012           -0.19578    0.01240  -15.79  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.402
## Multiple R-squared:  0.0535, Adjusted R-squared:  0.0531
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
##  7 observations c(36,29922,37627,39348,43640,44389,45150)
##   are outliers with |weight| = 0 ( < 2.1e-06);
##  4019 weights are ~= 1. The remaining 44378 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0006  0.8660  0.9510  0.8960  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          2.07e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.016  1           1.008
## LastAuthorFemale  1.018  1           1.009
## Year              1.012 16           1.000
```



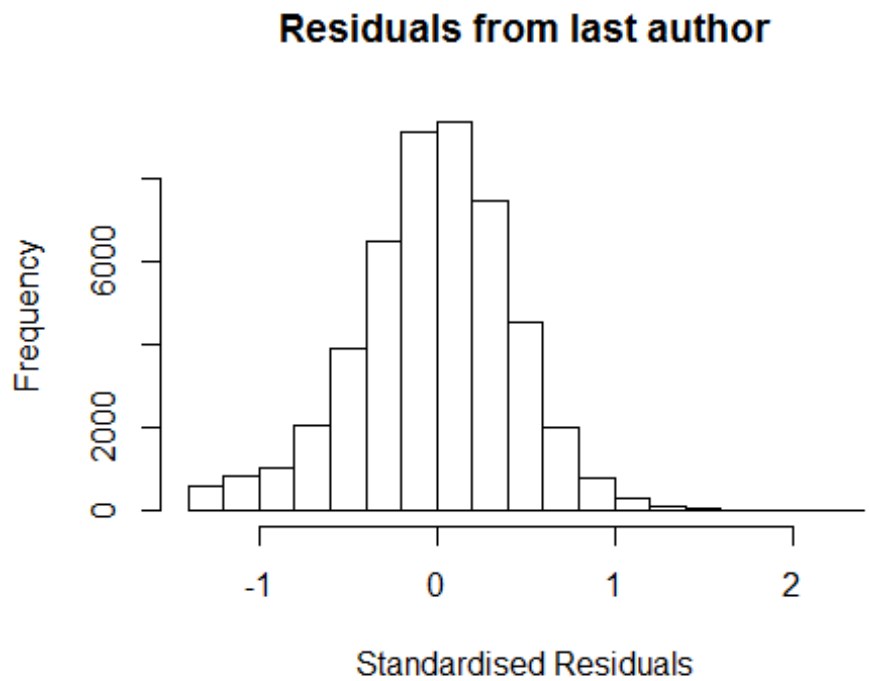

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.32027 -0.27204  0.00592  0.27251  2.34557
##
## Coefficients:
##                      Estimate Std. Error t value Pr(>|t|)
## (Intercept)          1.320270   0.010076  131.04  < 2e-16 ***
## FirstAuthorFemale1 -0.020411   0.004032   -5.06  4.1e-07 ***
## LastAuthorFemale1  -0.031605   0.004612   -6.85  7.3e-12 ***
## Year1997           -0.033815   0.013963   -2.42    0.015 *
## Year1998           -0.000956   0.029739   -0.03    0.974
## Year1999           -0.121189   0.012513   -9.69  < 2e-16 ***
## Year2000           -0.013838   0.025261   -0.55    0.584
## Year2001           -0.089894   0.013382   -6.72  1.9e-11 ***
## Year2002           -0.124965   0.012420  -10.06  < 2e-16 ***
## Year2003           -0.153362   0.012271  -12.50  < 2e-16 ***
## Year2004           -0.171612   0.012139  -14.14  < 2e-16 ***
## Year2005           -0.155130   0.012091  -12.83  < 2e-16 ***
```

```
## Year2006           -0.147823   0.012064  -12.25  < 2e-16 ***
## Year2007           -0.146275   0.012127  -12.06  < 2e-16 ***
## Year2008           -0.117985   0.012411   -9.51  < 2e-16 ***
## Year2009           -0.171591   0.012420  -13.82  < 2e-16 ***
## Year2010           -0.139130   0.012356  -11.26  < 2e-16 ***
## Year2011           -0.137817   0.012432  -11.09  < 2e-16 ***
## Year2012           -0.156230   0.012503  -12.50  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.407
## Multiple R-squared:  0.0154, Adjusted R-squared:  0.015
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
##  9 observations c(36,990,1263,29922,37627,39348,43640,44389,45150)
##   are outliers with |weight| = 0 ( < 2.1e-06);
##  3999 weights are ~= 1. The remaining 44396 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0013  0.8660  0.9510  0.8940  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          2.07e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.005  1           1.003
## Year              1.005 16           1.000
```




```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.31506 -0.27317  0.00615  0.27263  2.34363
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.31506    0.01004  130.93  < 2e-16 ***
## FirstAuthorFemale1 -0.02423    0.00402   -6.03  1.6e-09 ***
## Year1997           -0.03390    0.01398   -2.43    0.015 *
## Year1998           -0.00131    0.02981   -0.04    0.965
## Year1999           -0.12113    0.01251   -9.68  < 2e-16 ***
## Year2000           -0.01466    0.02533   -0.58    0.563
## Year2001           -0.09053    0.01338   -6.77  1.3e-11 ***
## Year2002           -0.12522    0.01242  -10.08  < 2e-16 ***
## Year2003           -0.15358    0.01227  -12.52  < 2e-16 ***
## Year2004           -0.17172    0.01214  -14.15  < 2e-16 ***
## Year2005           -0.15512    0.01210  -12.82  < 2e-16 ***
## Year2006           -0.14835    0.01206  -12.30  < 2e-16 ***
```

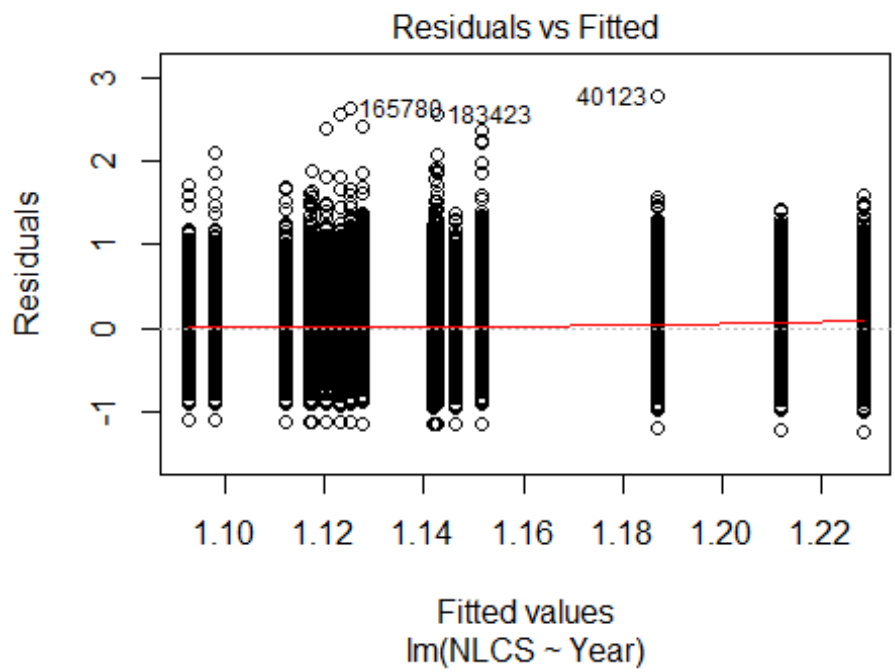
```
## Year2007           -0.14669    0.01213  -12.10  < 2e-16 ***
## Year2008           -0.11911    0.01241   -9.60  < 2e-16 ***
## Year2009           -0.17317    0.01241  -13.95  < 2e-16 ***
## Year2010           -0.14038    0.01236  -11.35  < 2e-16 ***
## Year2011           -0.13957    0.01243  -11.23  < 2e-16 ***
## Year2012           -0.15772    0.01250  -12.61  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.407
## Multiple R-squared:  0.0144, Adjusted R-squared:  0.014
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
##  9 observations c(36,990,1263,29922,37627,39348,43640,44389,45150)
##   are outliers with |weight| = 0 ( < 2.1e-06);
##  4002 weights are ~= 1. The remaining 44393 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0038  0.8660  0.9510  0.8940  0.9860  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          2.07e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##                   GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.007  1           1.004
## Year             1.007 16           1.000
```



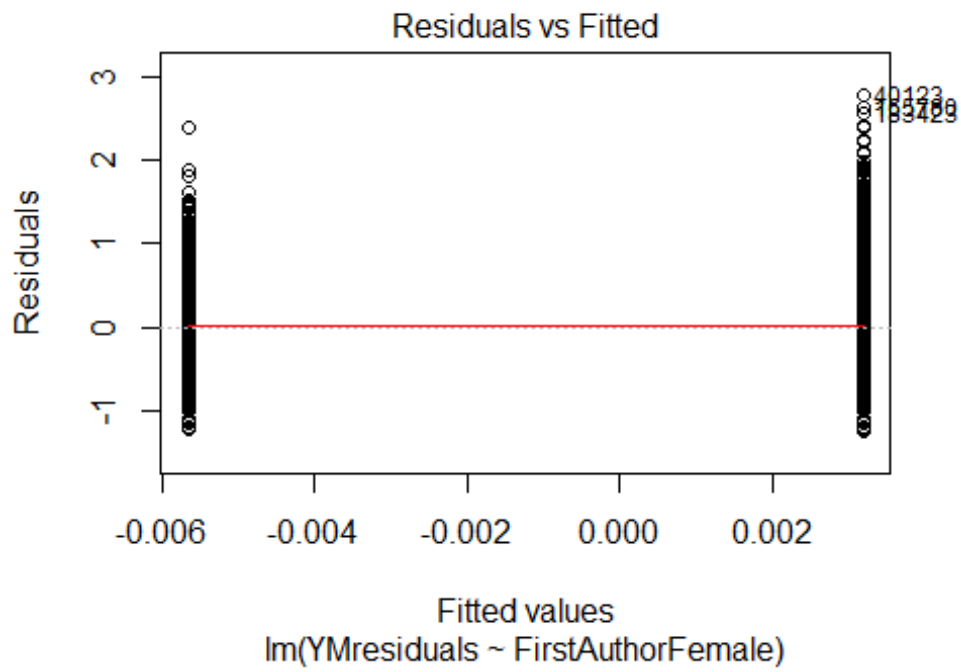

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.31425 -0.27369  0.00604  0.27309  2.35633
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.314250   0.009989  131.58  < 2e-16 ***
## LastAuthorFemale1 -0.034873   0.004595   -7.59  3.3e-14 ***
## Year1997          -0.034181   0.013962   -2.45    0.014 *
## Year1998          -0.000543   0.029745   -0.02    0.985
## Year1999          -0.121293   0.012514   -9.69  < 2e-16 ***
## Year2000          -0.012480   0.025283   -0.49    0.622
## Year2001          -0.090491   0.013384   -6.76  1.4e-11 ***
## Year2002          -0.125323   0.012418  -10.09  < 2e-16 ***
## Year2003          -0.154198   0.012272  -12.57  < 2e-16 ***
## Year2004          -0.172357   0.012137  -14.20  < 2e-16 ***
## Year2005          -0.156158   0.012089  -12.92  < 2e-16 ***
## Year2006          -0.148758   0.012062  -12.33  < 2e-16 ***
```

```
## Year2007          -0.147405   0.012126  -12.16  < 2e-16 ***
## Year2008          -0.119365   0.012405   -9.62  < 2e-16 ***
## Year2009          -0.172991   0.012416  -13.93  < 2e-16 ***
## Year2010          -0.140459   0.012353  -11.37  < 2e-16 ***
## Year2011          -0.139278   0.012430  -11.20  < 2e-16 ***
## Year2012          -0.157709   0.012500  -12.62  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.407
## Multiple R-squared:  0.0148, Adjusted R-squared:  0.0144
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
##  9 observations c(36,990,1263,29922,37627,39348,43640,44389,45150)
##   are outliers with |weight| = 0 ( < 2.1e-06);
##  4002 weights are ~= 1. The remaining 44393 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0007  0.8660  0.9510  0.8940  0.9860  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          2.07e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                  psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Sample size for the above analysis:  48404"
## [1] ""
## [1] ""
## [1] "##################################"
## [1] "Analysis of AJSC 1312"
## [1] "##################################"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
##  1996  1997  1998  1999  2000  2001  2002  2003  2004  2005  2006  2007
## 13387 11903 11108  9424 10142  9628 10152  9184 10457 10243 10731 10487
##  2008  2009  2010  2011  2012
## 10972 11712 11654 11715 10879
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
```

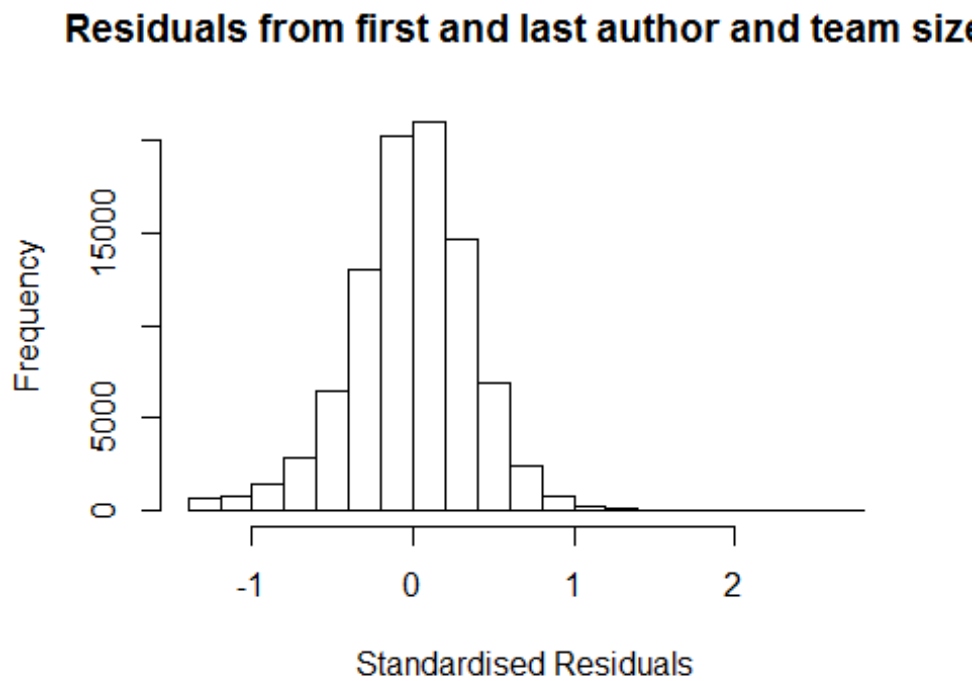
```
## 6986 6571 5911 5909 5175 4533 6416 5773 6585 6315 6453 6386 6822 7273 7200
## 2011 2012
## 7242 6907
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 5991 5613 5122 5028 4423 3882 5465 4922 5557 5257 5385 5374 5741 6086 5960
## 2011 2012
## 6031 5774
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
##  Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 3000, df = 16, p-value <2e-16
```




```
##
##  Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 240, df = 1, p-value <2e-16
```



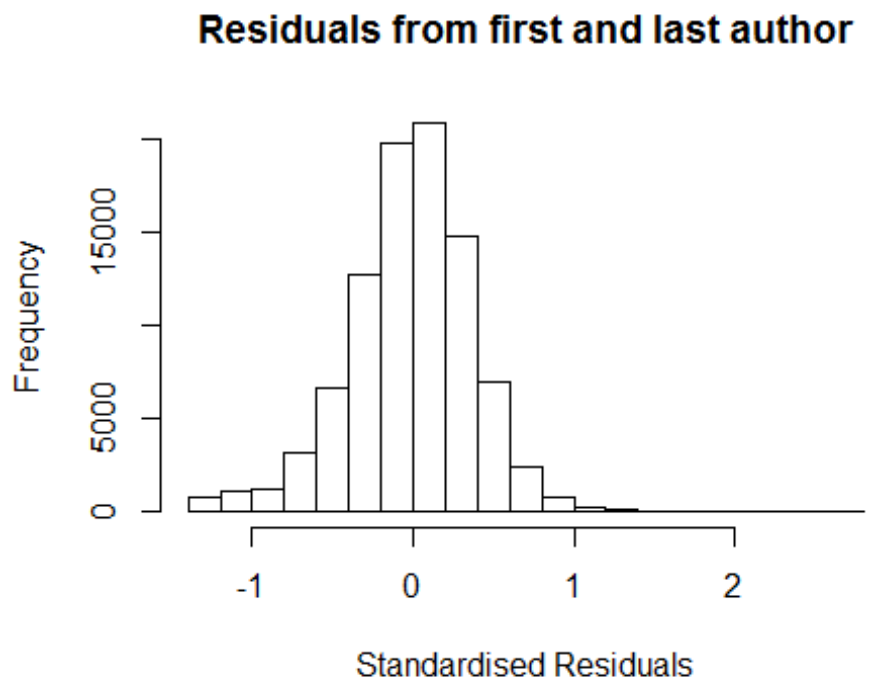

```
## [1] "Female first author team size 2018 geometric mean: 5.08628881306984"
## [1] "Male first author team size 2018 geometric mean: 4.64006035399506"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 2500000, p-value = 2e-05
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 4.83653826029146"
## [1] "Male last author team size 2018 geometric mean: 4.81660430120505"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 1900000, p-value = 0.8
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##                     GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.013  1           1.006
## LastAuthorFemale  1.010  1           1.005
## UniqueAuthors     1.026  4           1.003
## Year              1.029 16           1.001
```




```
## [1] "List of  3 outliers with residuals above  2.5"
##           ScopusId  NLCS Year OneField Fields residuals
## 40123   0031773680 3.959 1998     1303      6     2.769
## 71613   0035710746 3.680 2001     1300      2     2.607
## 165780 65449136284 3.759 2009     1303      7     2.653
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
UniqueAuthors +
##     Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##     k.max = 1000))
##  \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.39271 -0.22707  0.00366  0.22418  2.76882
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.07420    0.01008  106.61  < 2e-16 ***
## FirstAuthorFemale1 -0.01511    0.00243   -6.23  4.7e-10 ***
## LastAuthorFemale1  -0.01589    0.00286   -5.56  2.7e-08 ***
## UniqueAuthors2      0.18242    0.00888   20.55  < 2e-16 ***
## UniqueAuthors3      0.21755    0.00881   24.70  < 2e-16 ***
## UniqueAuthors4      0.24490    0.00882   27.77  < 2e-16 ***
## UniqueAuthors5      0.31851    0.00856   37.19  < 2e-16 ***
## Year1997           -0.02473    0.00837   -2.95   0.0031 **
```

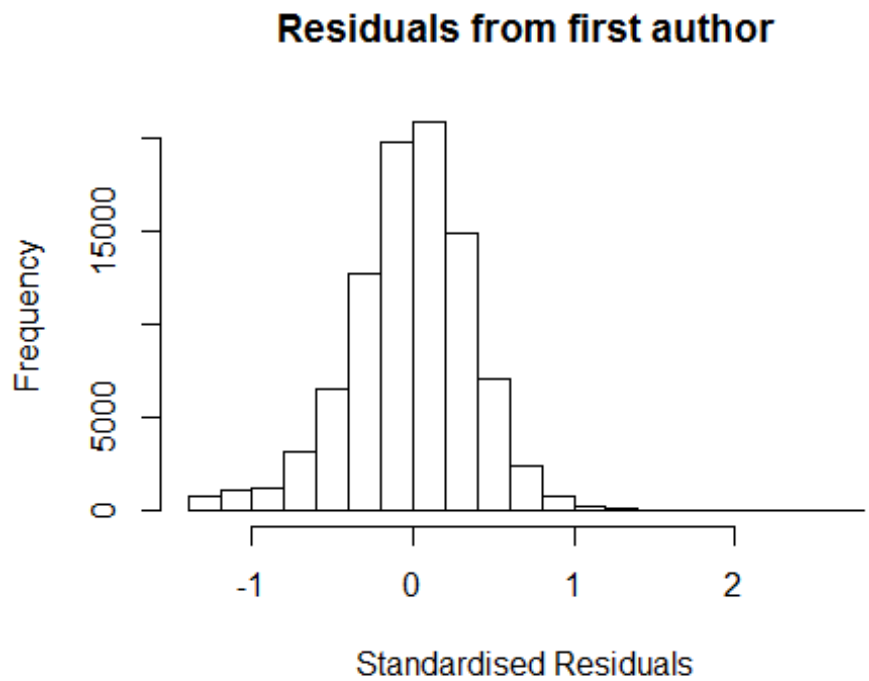
```
## Year1998           -0.06644    0.00813   -8.18  2.9e-16 ***
## Year1999           -0.15245    0.00746  -20.43  < 2e-16 ***
## Year2000           -0.14983    0.00774  -19.35  < 2e-16 ***
## Year2001           -0.18329    0.00789  -23.23  < 2e-16 ***
## Year2002           -0.18054    0.00753  -23.96  < 2e-16 ***
## Year2003           -0.21395    0.00756  -28.30  < 2e-16 ***
## Year2004           -0.20653    0.00739  -27.95  < 2e-16 ***
## Year2005           -0.19294    0.00740  -26.07  < 2e-16 ***
## Year2006           -0.19448    0.00749  -25.96  < 2e-16 ***
## Year2007           -0.18888    0.00751  -25.14  < 2e-16 ***
## Year2008           -0.17381    0.00736  -23.62  < 2e-16 ***
## Year2009           -0.18580    0.00741  -25.09  < 2e-16 ***
## Year2010           -0.17876    0.00750  -23.83  < 2e-16 ***
## Year2011           -0.16777    0.00754  -22.25  < 2e-16 ***
## Year2012           -0.18986    0.00769  -24.68  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.337
## Multiple R-squared:  0.0689, Adjusted R-squared:  0.0687
## Convergence in 15 IRWLS iterations
##
## Robustness weights:
##  32 observations
c(16639,29289,30018,30019,35441,36830,38507,40935,42283,48982,49244,50140,580
01,58849,62029,69423,71098,71339,76218,77492,78552,80748,81922,83113,84497,84
499,84945,89168,89436,89720,89733,90326)
##   are outliers with |weight| = 0 ( < 1.1e-06);
##  7705 weights are ~= 1. The remaining 83874 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0005  0.8660  0.9510  0.8920  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.09e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                    psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##                    GVIF Df GVIF^(1/(2*Df))
```

```
## FirstAuthorFemale 1.009  1           1.004
## LastAuthorFemale  1.009  1           1.004
## Year              1.006 16           1.000
```



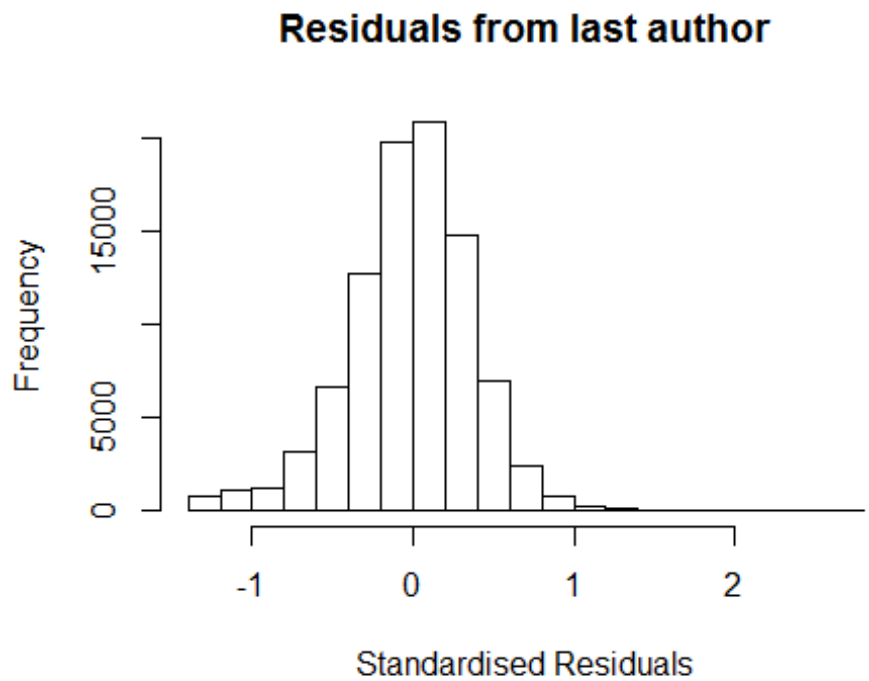

```
## [1] "List of  3 outliers with residuals above  2.5"
##           ScopusId  NLCS Year OneField Fields residuals
## 40123   0031773680 3.959 1998     1303      6     2.717
## 71613   0035710746 3.680 2001     1300      2     2.544
## 165780 65449136284 3.759 2009     1303      7     2.618
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.30269 -0.23101  0.00455  0.22613  2.71713
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)          1.30269    0.00608  214.10  < 2e-16 ***
## FirstAuthorFemale1 -0.00900    0.00247   -3.65  0.00026 ***
## LastAuthorFemale1  -0.01846    0.00290   -6.36  2.1e-10 ***
## Year1997           -0.02177    0.00851   -2.56  0.01049 *
## Year1998           -0.06082    0.00822   -7.40  1.3e-13 ***
## Year1999           -0.14124    0.00759  -18.61  < 2e-16 ***
```

```
## Year2000           -0.13689    0.00786  -17.42  < 2e-16 ***
## Year2001           -0.16654    0.00801  -20.80  < 2e-16 ***
## Year2002           -0.16150    0.00765  -21.11  < 2e-16 ***
## Year2003           -0.19198    0.00772  -24.88  < 2e-16 ***
## Year2004           -0.18546    0.00753  -24.62  < 2e-16 ***
## Year2005           -0.16978    0.00753  -22.55  < 2e-16 ***
## Year2006           -0.16984    0.00758  -22.40  < 2e-16 ***
## Year2007           -0.16317    0.00763  -21.39  < 2e-16 ***
## Year2008           -0.14668    0.00746  -19.66  < 2e-16 ***
## Year2009           -0.16193    0.00755  -21.46  < 2e-16 ***
## Year2010           -0.15006    0.00763  -19.66  < 2e-16 ***
## Year2011           -0.13747    0.00769  -17.88  < 2e-16 ***
## Year2012           -0.15839    0.00783  -20.24  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.341
## Multiple R-squared:  0.0236, Adjusted R-squared:  0.0235
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
##  30 observations
c(16639,29289,30018,30019,35441,36830,40935,42283,48982,50140,55279,58001,588
49,62029,69423,71339,76218,77492,78552,79090,80748,81922,83113,84497,84945,89
168,89436,89720,89733,90326)
##   are outliers with |weight| = 0 ( < 1.1e-06);
##  7624 weights are ~= 1. The remaining 83957 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0004  0.8650  0.9510  0.8900  0.9860  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.09e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.003  1           1.002
## Year              1.003 16           1.000
```




```
## [1] "List of  3 outliers with residuals above  2.5"
##           ScopusId  NLCS Year OneField Fields residuals
## 40123   0031773680 3.959 1998     1303      6     2.717
## 71613   0035710746 3.680 2001     1300      2     2.544
## 165780 65449136284 3.759 2009     1303      7     2.618
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.29983 -0.23105  0.00473  0.22619  2.72018
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.29983    0.00607  213.98  < 2e-16 ***
## FirstAuthorFemale1 -0.01060    0.00246   -4.31  1.7e-05 ***
## Year1997           -0.02190    0.00852   -2.57     0.01 *
## Year1998           -0.06102    0.00822   -7.42  1.2e-13 ***
## Year1999           -0.14144    0.00759  -18.62  < 2e-16 ***
## Year2000           -0.13699    0.00787  -17.41  < 2e-16 ***
## Year2001           -0.16694    0.00802  -20.82  < 2e-16 ***
## Year2002           -0.16183    0.00766  -21.13  < 2e-16 ***
## Year2003           -0.19221    0.00772  -24.89  < 2e-16 ***
## Year2004           -0.18584    0.00754  -24.65  < 2e-16 ***
```

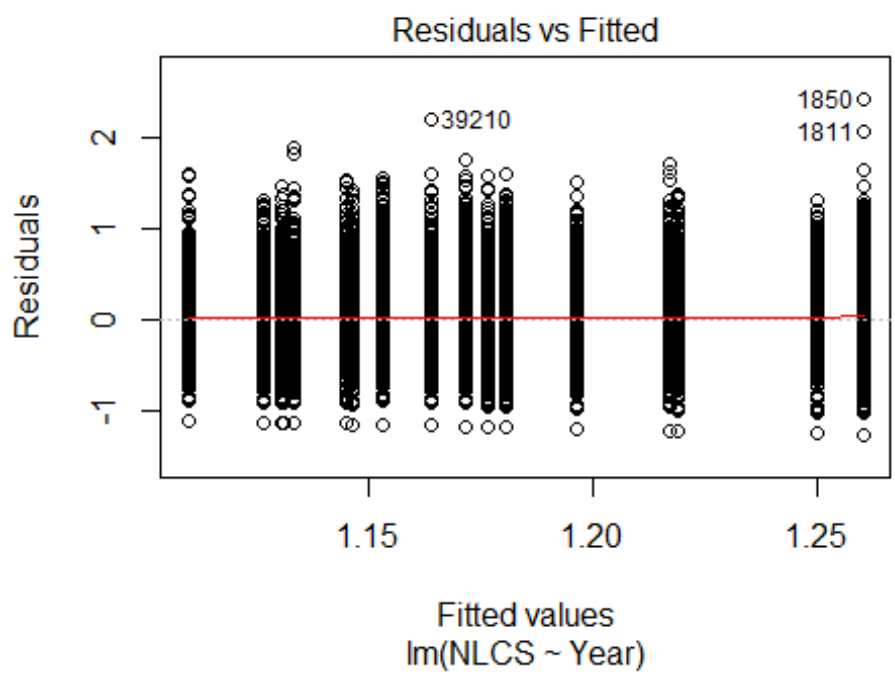
```
## Year2005           -0.17003    0.00753  -22.57  < 2e-16 ***
## Year2006           -0.17049    0.00759  -22.46  < 2e-16 ***
## Year2007           -0.16395    0.00763  -21.48  < 2e-16 ***
## Year2008           -0.14756    0.00746  -19.77  < 2e-16 ***
## Year2009           -0.16268    0.00755  -21.54  < 2e-16 ***
## Year2010           -0.15062    0.00764  -19.72  < 2e-16 ***
## Year2011           -0.13854    0.00769  -18.02  < 2e-16 ***
## Year2012           -0.15918    0.00783  -20.33  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.341
## Multiple R-squared:  0.0232, Adjusted R-squared:  0.023
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
##  29 observations
c(16639,29289,30018,30019,35441,36830,40935,42283,48982,50140,58001,58849,620
29,69423,71339,76218,77492,78552,79090,80748,81922,83113,84497,84945,89168,89
436,89720,89733,90326)
##   are outliers with |weight| = 0 ( < 1.1e-06);
##  7696 weights are ~= 1. The remaining 83886 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0001  0.8650  0.9510  0.8900  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.09e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##                   GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.003  1           1.002
## Year             1.003 16           1.000
```



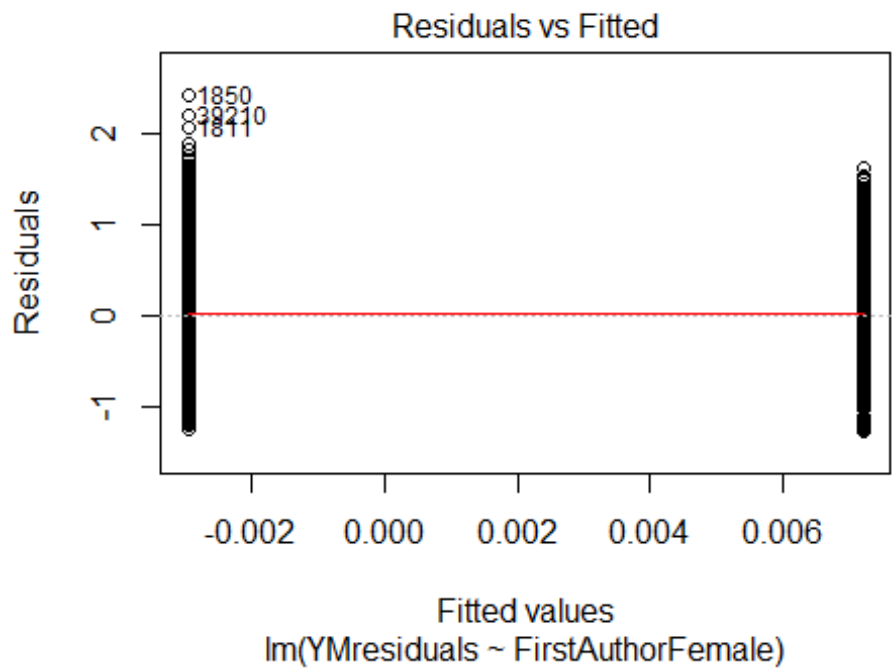

```
## [1] "List of  3 outliers with residuals above  2.5"
##           ScopusId  NLCS Year OneField Fields residuals
## 40123   0031773680 3.959 1998     1303      6     2.717
## 71613   0035710746 3.680 2001     1300      2     2.544
## 165780 65449136284 3.759 2009     1303      7     2.618
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.29979 -0.23089  0.00476  0.22602  2.72001
##
## Coefficients:
##                   Estimate Std. Error t value Pr(>|t|)
## (Intercept)        1.29979    0.00602  215.74  < 2e-16 ***
## LastAuthorFemale1 -0.01955    0.00290   -6.75  1.5e-11 ***
## Year1997          -0.02181    0.00851   -2.56     0.01 *
## Year1998          -0.06080    0.00822   -7.40  1.4e-13 ***
## Year1999          -0.14129    0.00759  -18.62  < 2e-16 ***
## Year2000          -0.13708    0.00786  -17.44  < 2e-16 ***
## Year2001          -0.16654    0.00801  -20.79  < 2e-16 ***
## Year2002          -0.16164    0.00765  -21.13  < 2e-16 ***
## Year2003          -0.19222    0.00772  -24.91  < 2e-16 ***
## Year2004          -0.18569    0.00753  -24.65  < 2e-16 ***
```

```
## Year2005          -0.16990    0.00753  -22.57  < 2e-16 ***
## Year2006          -0.17017    0.00759  -22.44  < 2e-16 ***
## Year2007          -0.16355    0.00763  -21.44  < 2e-16 ***
## Year2008          -0.14708    0.00746  -19.72  < 2e-16 ***
## Year2009          -0.16237    0.00755  -21.52  < 2e-16 ***
## Year2010          -0.15046    0.00763  -19.71  < 2e-16 ***
## Year2011          -0.13793    0.00769  -17.94  < 2e-16 ***
## Year2012          -0.15895    0.00783  -20.31  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.342
## Multiple R-squared:  0.0235, Adjusted R-squared:  0.0233
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
##  29 observations
c(16639,29289,30018,30019,35441,36830,42283,48982,50140,55279,58001,58849,620
29,69423,71339,76218,77492,78552,79090,80748,81922,83113,84497,84945,89168,89
436,89720,89733,90326)
##   are outliers with |weight| = 0 ( < 1.1e-06);
##  7654 weights are ~= 1. The remaining 83928 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.000   0.865   0.951   0.890   0.985   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.09e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Sample size for the above analysis:  91611"
## [1] ""
## [1] ""
## [1] "####################################"
## [1] "Analysis of AJSC 1313"
## [1] "####################################"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
```

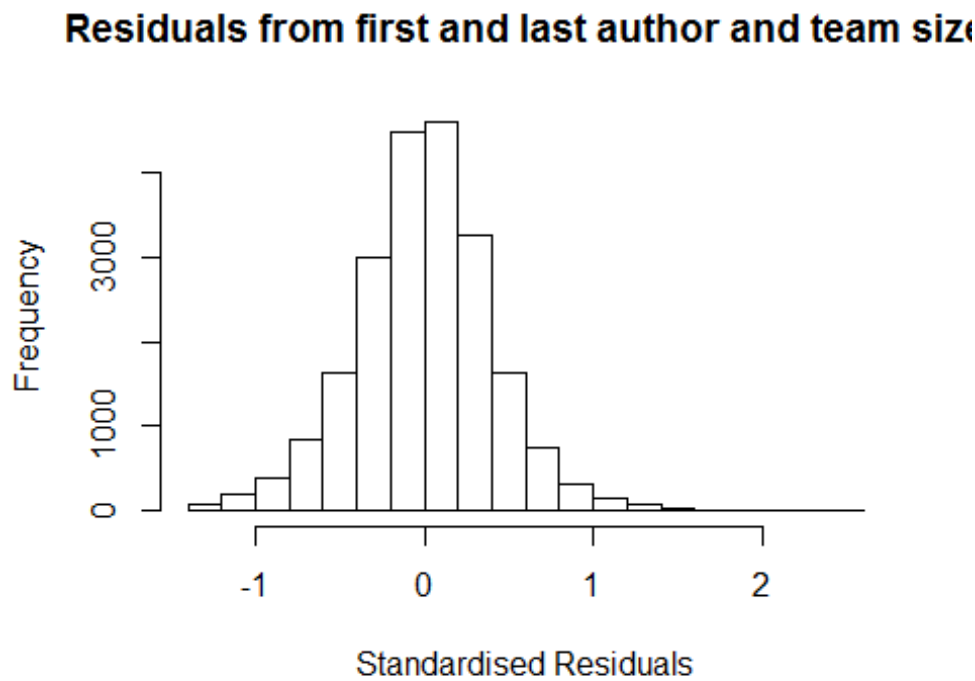
```
## 1863 1710 1693 1783 1788 2145 2143 2217 2412 2459 2733 2736 2914 3339 3496
## 2011 2012
## 3445 3467
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 1213 1086  930 1127  873 1019 1300 1369 1482 1514 1627 1720 1821 2114 2202
## 2011 2012
## 2214 2159
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 1037  935  793  932  750  864 1106 1147 1242 1206 1358 1398 1501 1728 1791
## 2011 2012
## 1803 1777
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
##  Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 420, df = 16, p-value <2e-16
```




```
##
##  Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 3.2, df = 1, p-value = 0.08
```

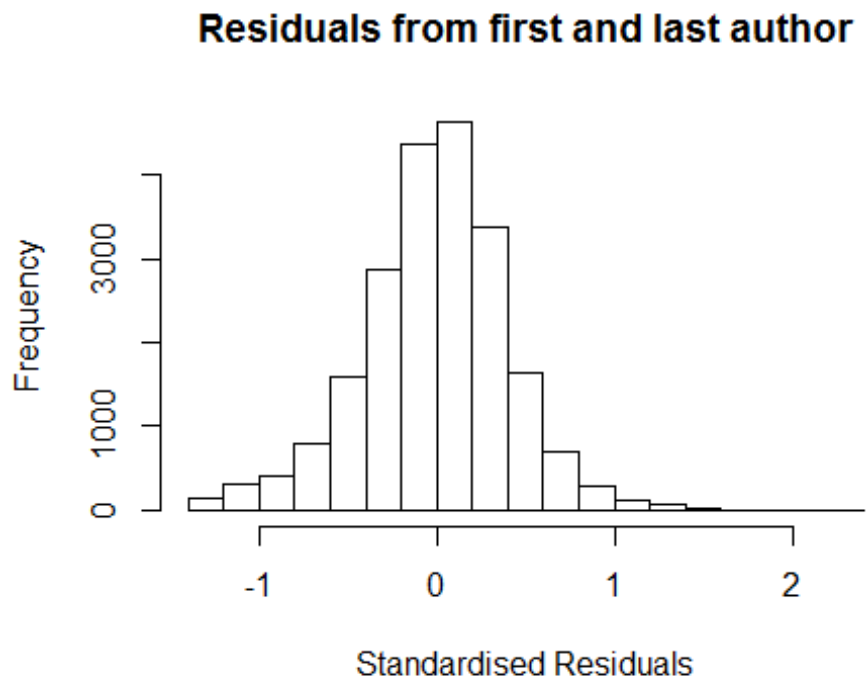



```
## [1] "Female first author team size 2018 geometric mean: 5.83779017507785"
## [1] "Male first author team size 2018 geometric mean: 6.04983570030759"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 220000, p-value = 0.3
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 5.87477264660974"
## [1] "Male last author team size 2018 geometric mean: 5.99297301716131"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 170000, p-value = 0.4
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.013  1           1.006
## LastAuthorFemale  1.006  1           1.003
## UniqueAuthors     1.042  4           1.005
## Year              1.053 16           1.002
```



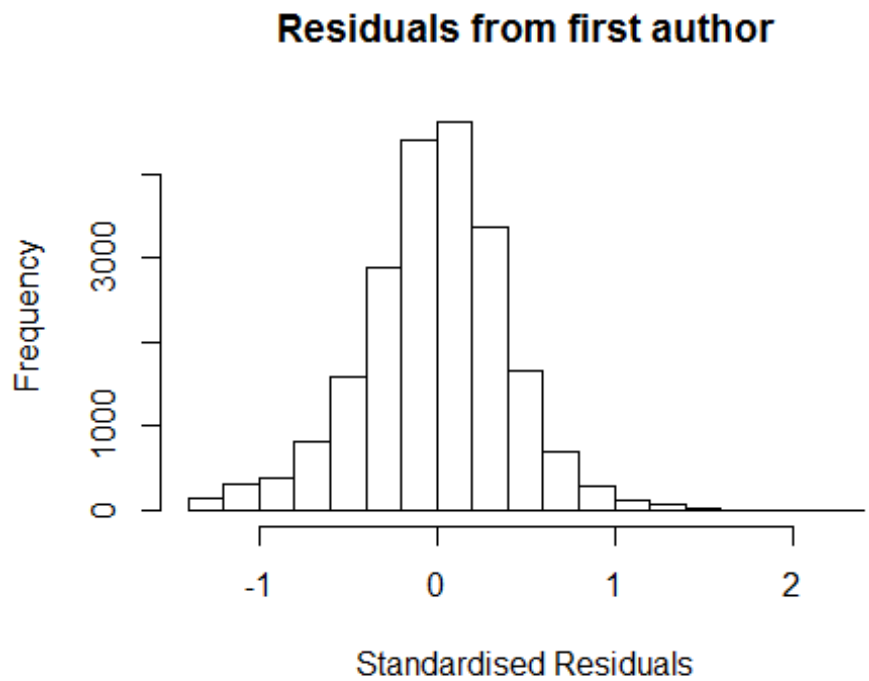

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
UniqueAuthors +
##     Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##     k.max = 1000))
##  \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.35316 -0.24082  0.00562  0.24343  2.47244
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         0.89350    0.02584   34.57  < 2e-16 ***
## FirstAuthorFemale1  0.00920    0.00585    1.57  0.11577
## LastAuthorFemale1  -0.00970    0.00697   -1.39  0.16377
## UniqueAuthors2      0.30906    0.02316   13.35  < 2e-16 ***
## UniqueAuthors3      0.36216    0.02277   15.90  < 2e-16 ***
## UniqueAuthors4      0.39509    0.02247   17.58  < 2e-16 ***
## UniqueAuthors5      0.45966    0.02148   21.40  < 2e-16 ***
## Year1997           -0.02519    0.02297   -1.10  0.27291
## Year1998           -0.00594    0.02155   -0.28  0.78270
## Year1999           -0.07433    0.02034   -3.65  0.00026 ***
```

```
## Year2000           -0.06100    0.02093   -2.91  0.00357 **
## Year2001           -0.14069    0.01997   -7.04  1.9e-12 ***
## Year2002           -0.14353    0.01924   -7.46  8.9e-14 ***
## Year2003           -0.16421    0.01846   -8.90  < 2e-16 ***
## Year2004           -0.18061    0.01830   -9.87  < 2e-16 ***
## Year2005           -0.16721    0.01875   -8.92  < 2e-16 ***
## Year2006           -0.16855    0.01827   -9.23  < 2e-16 ***
## Year2007           -0.15433    0.01830   -8.43  < 2e-16 ***
## Year2008           -0.13078    0.01828   -7.15  8.7e-13 ***
## Year2009           -0.12831    0.01778   -7.22  5.5e-13 ***
## Year2010           -0.15678    0.01783   -8.79  < 2e-16 ***
## Year2011           -0.13345    0.01809   -7.38  1.7e-13 ***
## Year2012           -0.14596    0.01847   -7.90  2.9e-15 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.361
## Multiple R-squared:  0.0852, Adjusted R-squared:  0.0843
## Convergence in 17 IRWLS iterations
##
## Robustness weights:
##  5 observations c(873,900,4893,5287,19317)
##   are outliers with |weight| = 0 ( < 4.7e-06);
##  1849 weights are ~= 1. The remaining 19514 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0023  0.8610  0.9510  0.8910  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          4.68e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.009  1           1.004
## LastAuthorFemale  1.006  1           1.003
## Year              1.015 16           1.000
```



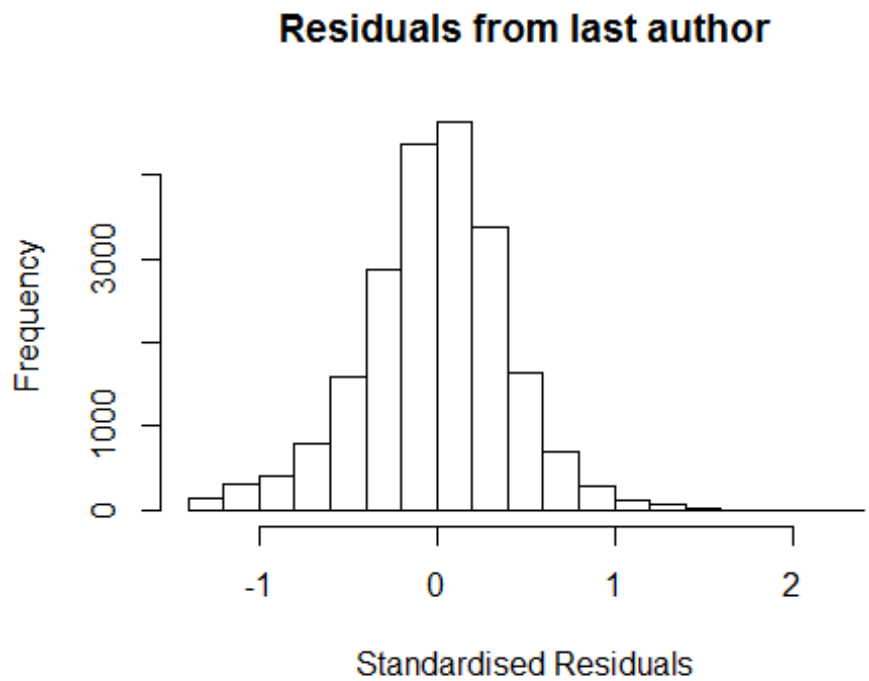

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.29654 -0.24602  0.00761  0.24303  2.39386
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)          1.28114    0.01595   80.31  < 2e-16 ***
## FirstAuthorFemale1  0.01435    0.00592    2.43   0.0153 *
## LastAuthorFemale1  -0.01200    0.00712   -1.69   0.0920 .
## Year1997           -0.01818    0.02410   -0.75   0.4506
## Year1998            0.00106    0.02174    0.05   0.9611
## Year1999           -0.06051    0.02078   -2.91   0.0036 **
## Year2000           -0.05368    0.02107   -2.55   0.0108 *
## Year2001           -0.14723    0.02035   -7.23  4.8e-13 ***
## Year2002           -0.13014    0.01974   -6.59  4.4e-11 ***
## Year2003           -0.13817    0.01887   -7.32  2.5e-13 ***
## Year2004           -0.16222    0.01884   -8.61  < 2e-16 ***
## Year2005           -0.15029    0.01922   -7.82  5.5e-15 ***
```

```
## Year2006           -0.14332    0.01865   -7.68  1.6e-14 ***
## Year2007           -0.12689    0.01882   -6.74  1.6e-11 ***
## Year2008           -0.09869    0.01871   -5.28  1.3e-07 ***
## Year2009           -0.10075    0.01819   -5.54  3.1e-08 ***
## Year2010           -0.12775    0.01828   -6.99  2.9e-12 ***
## Year2011           -0.11199    0.01851   -6.05  1.5e-09 ***
## Year2012           -0.12585    0.01882   -6.69  2.4e-11 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.366
## Multiple R-squared:  0.0145, Adjusted R-squared:  0.0136
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
##  5 observations c(873,900,4893,5287,19317)
##   are outliers with |weight| = 0 ( < 4.7e-06);
##  1777 weights are ~= 1. The remaining 19586 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0025  0.8610  0.9510  0.8870  0.9860  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          4.68e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.009  1           1.004
## Year              1.009 16           1.000
```




```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.2942 -0.2463  0.0074  0.2428  2.3955
##
## Coefficients:
##                      Estimate Std. Error t value Pr(>|t|)
## (Intercept)          1.279507   0.015934   80.30  < 2e-16 ***
## FirstAuthorFemale1   0.013754   0.005925    2.32   0.0203 *
## Year1997            -0.018343   0.024110   -0.76   0.4468
## Year1998             0.000894   0.021751    0.04   0.9672
## Year1999            -0.060493   0.020787   -2.91   0.0036 **
## Year2000            -0.053845   0.021067   -2.56   0.0106 *
## Year2001            -0.147622   0.020361   -7.25  4.3e-13 ***
## Year2002            -0.130252   0.019747   -6.60  4.3e-11 ***
## Year2003            -0.138183   0.018874   -7.32  2.5e-13 ***
## Year2004            -0.162624   0.018847   -8.63  < 2e-16 ***
## Year2005            -0.150221   0.019226   -7.81  5.8e-15 ***
## Year2006            -0.143786   0.018665   -7.70  1.4e-14 ***
```

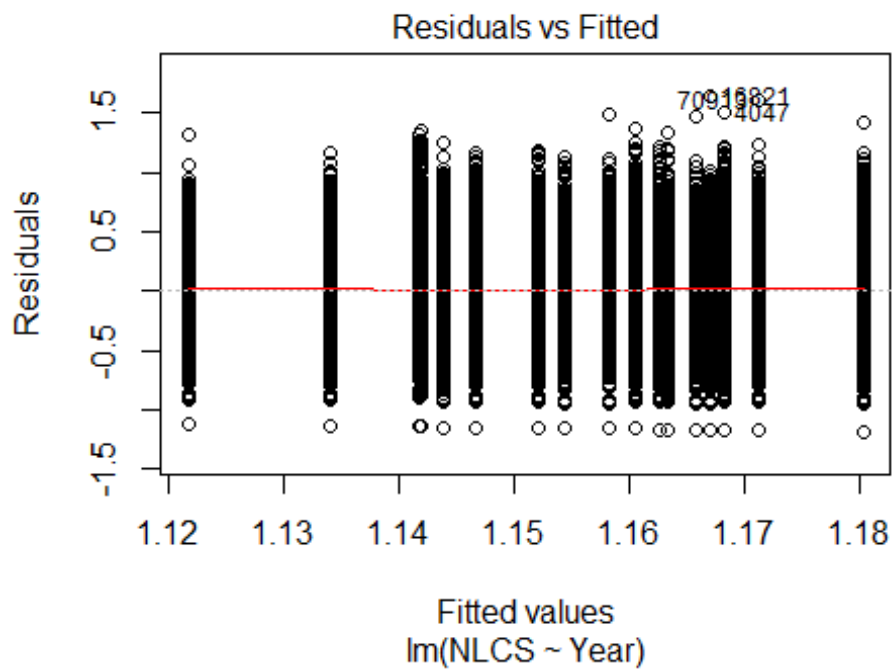
```
## Year2007           -0.127302   0.018822   -6.76  1.4e-11 ***
## Year2008           -0.099121   0.018721   -5.29  1.2e-07 ***
## Year2009           -0.101217   0.018203   -5.56  2.7e-08 ***
## Year2010           -0.128173   0.018294   -7.01  2.5e-12 ***
## Year2011           -0.112630   0.018514   -6.08  1.2e-09 ***
## Year2012           -0.126658   0.018820   -6.73  1.7e-11 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.366
## Multiple R-squared:  0.0143, Adjusted R-squared:  0.0135
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
##  5 observations c(873,900,4893,5287,19317)
##   are outliers with |weight| = 0 ( < 4.7e-06);
##  1810 weights are ~= 1. The remaining 19553 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0023  0.8610  0.9500  0.8870  0.9860  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          4.68e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##                   GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.007  1           1.003
## Year             1.007 16           1.000
```



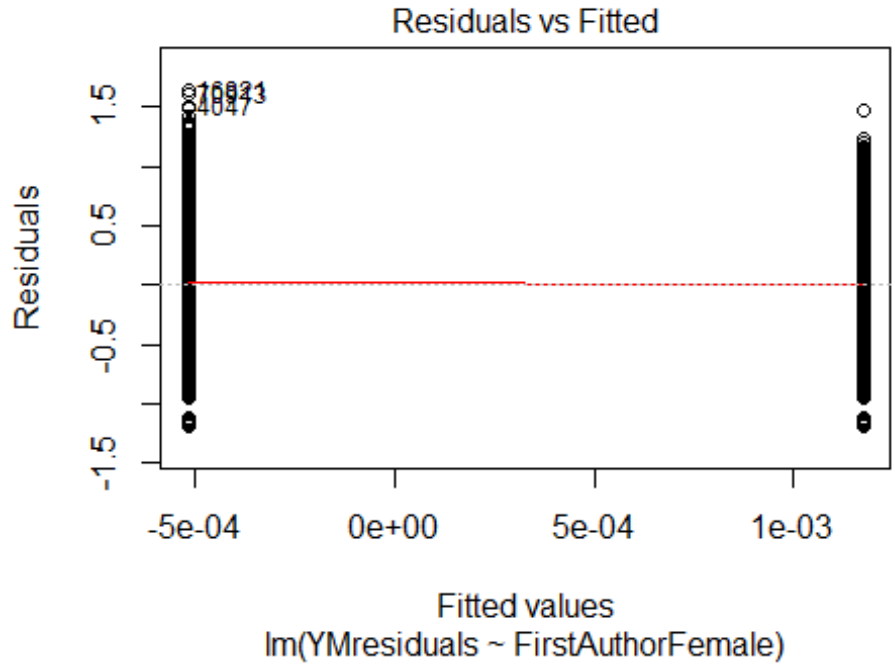

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.2859 -0.2450  0.0072  0.2421  2.3908
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.28425    0.01590   80.76  < 2e-16 ***
## LastAuthorFemale1 -0.01100    0.00714   -1.54   0.1231
## Year1997          -0.01765    0.02411   -0.73   0.4641
## Year1998           0.00162    0.02173    0.07   0.9405
## Year1999          -0.06039    0.02079   -2.91   0.0037 **
## Year2000          -0.05326    0.02107   -2.53   0.0115 *
## Year2001          -0.14647    0.02035   -7.20  6.3e-13 ***
## Year2002          -0.12946    0.01974   -6.56  5.6e-11 ***
## Year2003          -0.13713    0.01886   -7.27  3.7e-13 ***
## Year2004          -0.16178    0.01885   -8.58  < 2e-16 ***
## Year2005          -0.14984    0.01922   -7.80  6.7e-15 ***
## Year2006          -0.14222    0.01865   -7.63  2.5e-14 ***
```

```
## Year2007          -0.12622    0.01882   -6.71  2.0e-11 ***
## Year2008          -0.09716    0.01871   -5.19  2.1e-07 ***
## Year2009          -0.09945    0.01819   -5.47  4.7e-08 ***
## Year2010          -0.12625    0.01827   -6.91  5.0e-12 ***
## Year2011          -0.11035    0.01850   -5.97  2.5e-09 ***
## Year2012          -0.12433    0.01881   -6.61  4.0e-11 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.366
## Multiple R-squared:  0.0142, Adjusted R-squared:  0.0134
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
##  5 observations c(873,900,4893,5287,19317)
##   are outliers with |weight| = 0 ( < 4.7e-06);
##  1804 weights are ~= 1. The remaining 19559 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0031  0.8620  0.9510  0.8870  0.9860  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          4.68e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                  psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Sample size for the above analysis:  21368"
## [1] ""
## [1] ""
## [1] "##################################"
## [1] "Analysis of AJSC 1314"
## [1] "##################################"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 4186 4929 4523 4165 4418 4056 4393 3961 4051 4334 4197 4697 4036 3813 3951
## 2011 2012
## 3927 3914
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
```

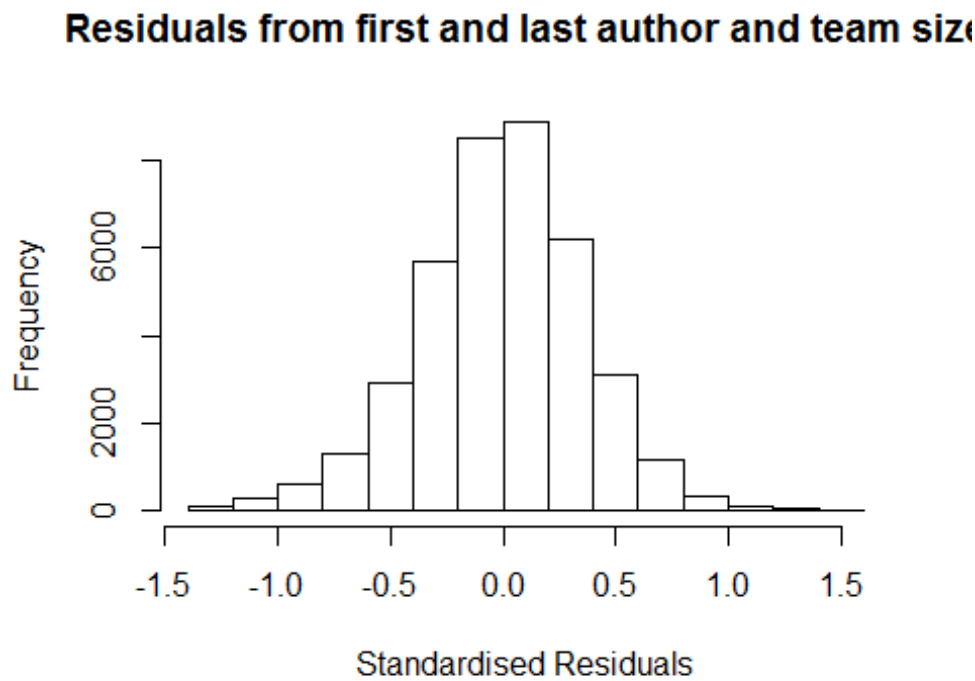
```
## 2519 3087 2836 2566 2449 1998 2796 2548 2656 2822 2796 3078 2698 2521 2637
## 2011 2012
## 2692 2665
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 2172 2692 2514 2230 2098 1732 2403 2174 2318 2409 2438 2663 2326 2160 2274
## 2011 2012
## 2279 2276
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
##  Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 410, df = 16, p-value <2e-16
```




```
##
##  Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 83, df = 1, p-value <2e-16
```

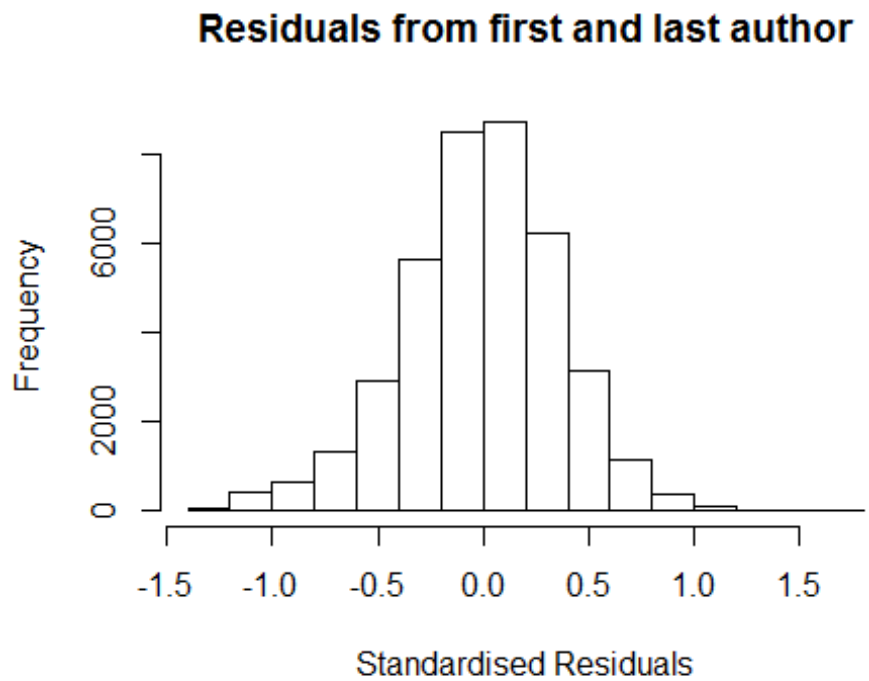



```
## [1] "Female first author team size 2018 geometric mean: 4.52012491979694"
## [1] "Male first author team size 2018 geometric mean: 4.29581153204532"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 520000, p-value = 0.2
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 4.29342377734666"
## [1] "Male last author team size 2018 geometric mean: 4.43647530195472"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 4e+05, p-value = 0.3
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##                     GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.024  1           1.012
## LastAuthorFemale  1.020  1           1.010
## UniqueAuthors     1.047  4           1.006
## Year              1.060 16           1.002
```



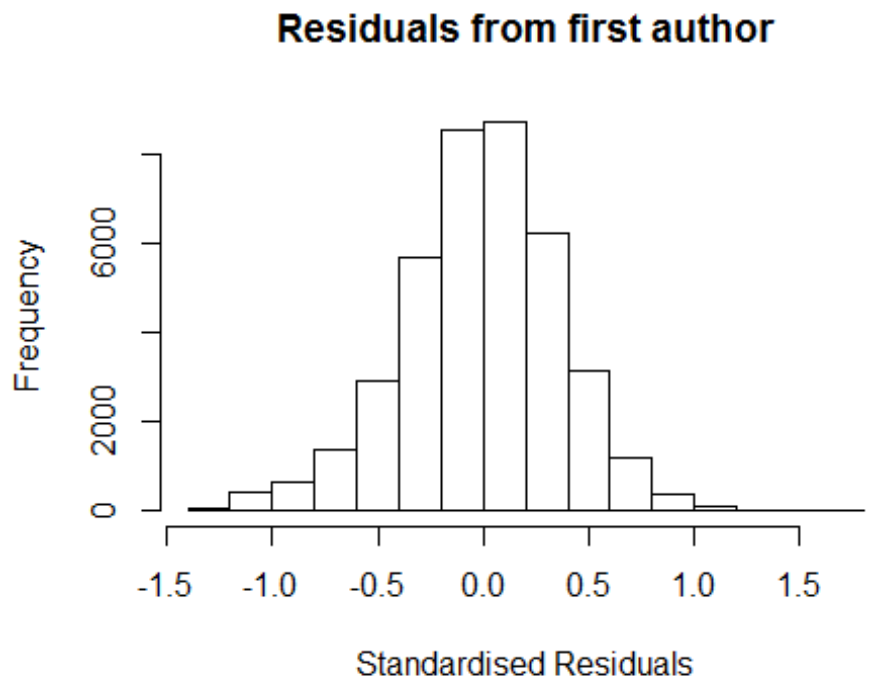

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
UniqueAuthors +
##     Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##     k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.2915 -0.2302  0.0034  0.2276  1.5995
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.04737    0.01314   79.71  < 2e-16 ***
## FirstAuthorFemale1 -0.00610    0.00389   -1.57  0.11688
## LastAuthorFemale1  -0.00750    0.00470   -1.60  0.11028
## UniqueAuthors2      0.11586    0.01079   10.73  < 2e-16 ***
## UniqueAuthors3      0.12769    0.01071   11.93  < 2e-16 ***
## UniqueAuthors4      0.16000    0.01084   14.76  < 2e-16 ***
## UniqueAuthors5      0.23712    0.01045   22.68  < 2e-16 ***
## Year1997           -0.01299    0.01237   -1.05  0.29347
## Year1998            0.00702    0.01208    0.58  0.56140
## Year1999           -0.01938    0.01195   -1.62  0.10471
```

```
## Year2000           -0.03747    0.01182   -3.17  0.00153 **
## Year2001           -0.05358    0.01199   -4.47  8.0e-06 ***
## Year2002           -0.05629    0.01164   -4.84  1.3e-06 ***
## Year2003           -0.07331    0.01170   -6.26  3.8e-10 ***
## Year2004           -0.05816    0.01162   -5.00  5.6e-07 ***
## Year2005           -0.06627    0.01151   -5.76  8.6e-09 ***
## Year2006           -0.04098    0.01164   -3.52  0.00043 ***
## Year2007           -0.06184    0.01134   -5.45  5.0e-08 ***
## Year2008           -0.05031    0.01184   -4.25  2.1e-05 ***
## Year2009           -0.03827    0.01198   -3.19  0.00141 **
## Year2010           -0.03687    0.01204   -3.06  0.00219 **
## Year2011           -0.06938    0.01226   -5.66  1.5e-08 ***
## Year2012           -0.07233    0.01223   -5.92  3.3e-09 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.341
## Multiple R-squared:  0.0357, Adjusted R-squared:  0.0352
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
##  observation 33914 is an outlier with |weight| = 0 ( < 2.6e-06);
##  3381 weights are ~= 1. The remaining 35776 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.005   0.866   0.950   0.897   0.985   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          2.55e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                    psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.022  1           1.011
## LastAuthorFemale  1.018  1           1.009
## Year              1.015 16           1.000
```



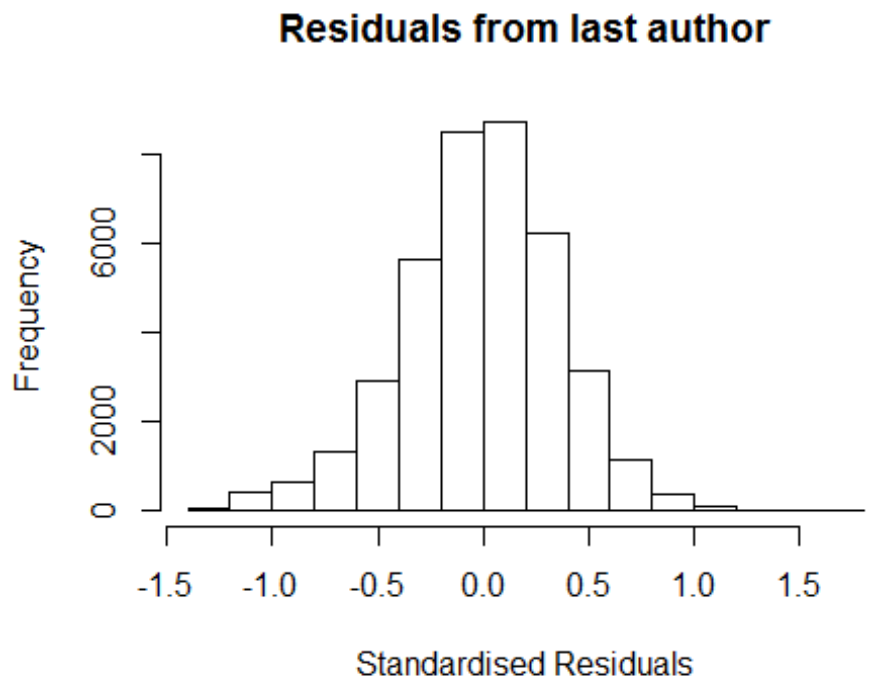

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.20372 -0.23254  0.00307  0.23047  1.62327
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)          1.18502    0.00932  127.08   <2e-16 ***
## FirstAuthorFemale1 -0.00179    0.00394   -0.45   0.6496
## LastAuthorFemale1  -0.01420    0.00477   -2.98   0.0029 **
## Year1997           -0.00356    0.01242   -0.29   0.7744
## Year1998            0.01870    0.01211    1.54   0.1225
## Year1999           -0.00429    0.01197   -0.36   0.7201
## Year2000           -0.01849    0.01185   -1.56   0.1186
## Year2001           -0.02960    0.01202   -2.46   0.0138 *
## Year2002           -0.03228    0.01169   -2.76   0.0058 **
## Year2003           -0.04571    0.01177   -3.88   0.0001 ***
## Year2004           -0.02954    0.01166   -2.53   0.0113 *
## Year2005           -0.03368    0.01155   -2.92   0.0036 **
```

```
## Year2006           -0.01024    0.01168   -0.88   0.3808
## Year2007           -0.02476    0.01138   -2.18   0.0295 *
## Year2008           -0.01380    0.01190   -1.16   0.2463
## Year2009           -0.00164    0.01201   -0.14   0.8914
## Year2010            0.00147    0.01210    0.12   0.9037
## Year2011           -0.03125    0.01227   -2.55   0.0109 *
## Year2012           -0.03248    0.01230   -2.64   0.0083 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.345
## Multiple R-squared:  0.0025, Adjusted R-squared:  0.00204
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
##  observation 8013 is an outlier with |weight| = 0 ( < 2.6e-06);
##  3466 weights are ~= 1. The remaining 35691 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0015  0.8650  0.9500  0.8960  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          2.55e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##                   GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.01  1           1.005
## Year              1.01 16           1.000
```




```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.20194 -0.23253  0.00249  0.23011  1.62522
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.18298    0.00931  127.07   <2e-16 ***
## FirstAuthorFemale1 -0.00319    0.00391   -0.81   0.4154
## Year1997           -0.00347    0.01243   -0.28   0.7803
## Year1998            0.01897    0.01211    1.57   0.1175
## Year1999           -0.00419    0.01198   -0.35   0.7262
## Year2000           -0.01841    0.01185   -1.55   0.1203
## Year2001           -0.02954    0.01202   -2.46   0.0140 *
## Year2002           -0.03245    0.01170   -2.77   0.0055 **
## Year2003           -0.04574    0.01178   -3.88   0.0001 ***
## Year2004           -0.02956    0.01166   -2.53   0.0113 *
## Year2005           -0.03370    0.01156   -2.92   0.0035 **
## Year2006           -0.01050    0.01168   -0.90   0.3686
```

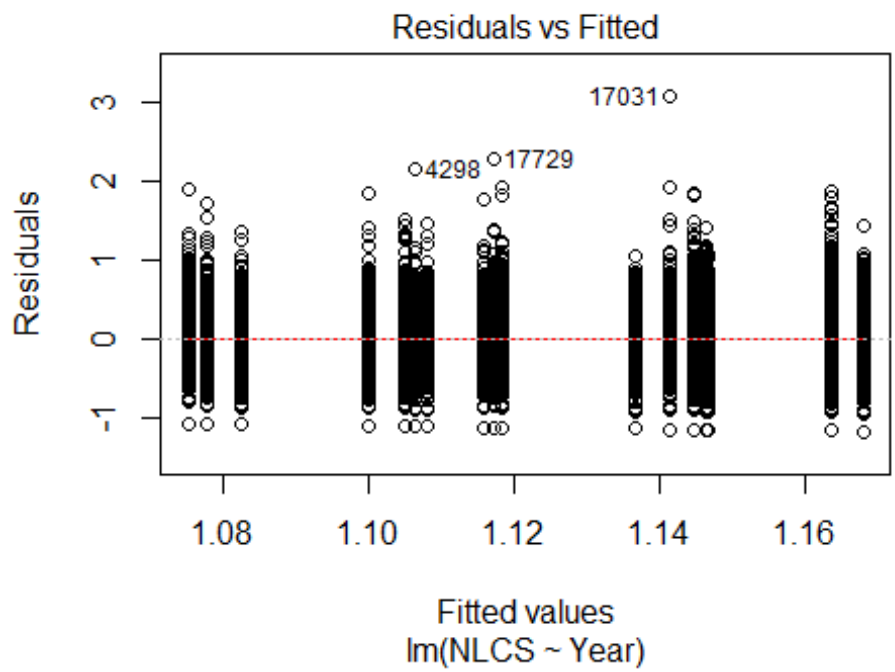
```
## Year2007           -0.02475    0.01138   -2.17   0.0297 *
## Year2008           -0.01408    0.01190   -1.18   0.2368
## Year2009           -0.00224    0.01201   -0.19   0.8520
## Year2010            0.00109    0.01211    0.09   0.9284
## Year2011           -0.03204    0.01226   -2.61   0.0090 **
## Year2012           -0.03272    0.01230   -2.66   0.0078 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.345
## Multiple R-squared:  0.00226,    Adjusted R-squared:  0.00183
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
##  observation 8013 is an outlier with |weight| = 0 ( < 2.6e-06);
##  3414 weights are ~= 1. The remaining 35743 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0013  0.8650  0.9500  0.8960  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          2.55e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##                   GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.006  1           1.003
## Year             1.006 16           1.000
```



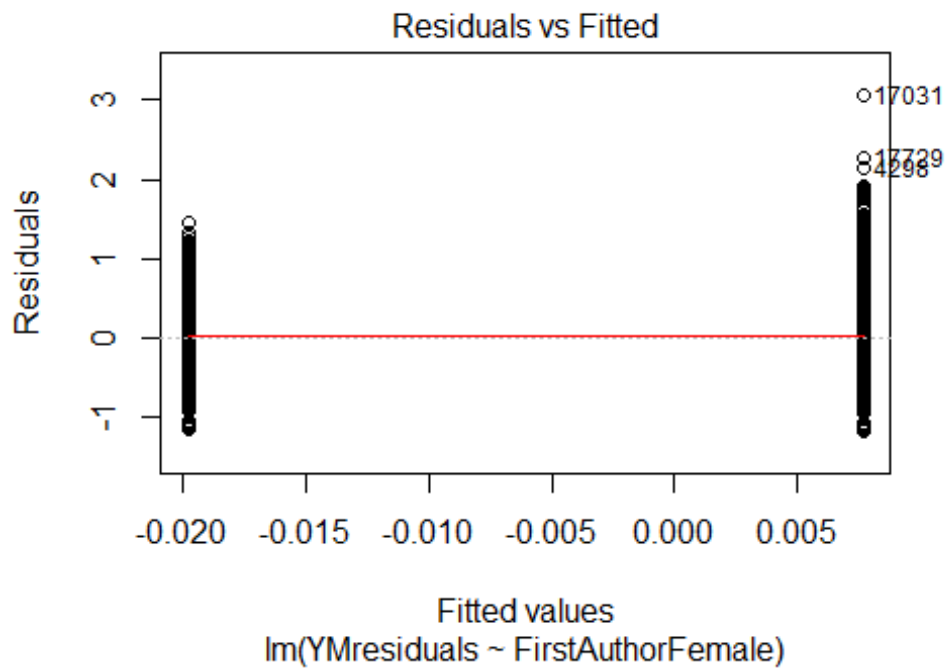

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.20327 -0.23279  0.00277  0.23073  1.62370
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.18461    0.00927  127.74   <2e-16 ***
## LastAuthorFemale1 -0.01445    0.00474   -3.05   0.0023 **
## Year1997          -0.00355    0.01242   -0.29   0.7749
## Year1998           0.01866    0.01211    1.54   0.1233
## Year1999          -0.00431    0.01197   -0.36   0.7186
## Year2000          -0.01857    0.01185   -1.57   0.1171
## Year2001          -0.02967    0.01202   -2.47   0.0135 *
## Year2002          -0.03233    0.01169   -2.77   0.0057 **
## Year2003          -0.04581    0.01177   -3.89    1e-04 ***
## Year2004          -0.02964    0.01165   -2.54   0.0110 *
## Year2005          -0.03382    0.01155   -2.93   0.0034 **
## Year2006          -0.01038    0.01168   -0.89   0.3740
```

```
## Year2007          -0.02493    0.01137   -2.19   0.0284 *
## Year2008          -0.01396    0.01189   -1.17   0.2403
## Year2009          -0.00179    0.01200   -0.15   0.8811
## Year2010           0.00129    0.01210    0.11   0.9150
## Year2011          -0.03147    0.01226   -2.57   0.0102 *
## Year2012          -0.03269    0.01229   -2.66   0.0078 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.345
## Multiple R-squared:  0.00249,    Adjusted R-squared:  0.00206
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
##  observation 8013 is an outlier with |weight| = 0 ( < 2.6e-06);
##  3483 weights are ~= 1. The remaining 35674 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0014  0.8650  0.9500  0.8960  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          2.55e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Sample size for the above analysis:  39158"
## [1] ""
## [1] ""
## [1] "####################################"
## [1] "Analysis of AJSC 1315"
## [1] "####################################"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  873  805  878  904  900  858  832  896 1036 1093 1242 1254 1181 1107 1240
## 2011 2012
## 1194 1008
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  584  528  521  501  528  409  568  578  706  683  775  802  746  713  758
```

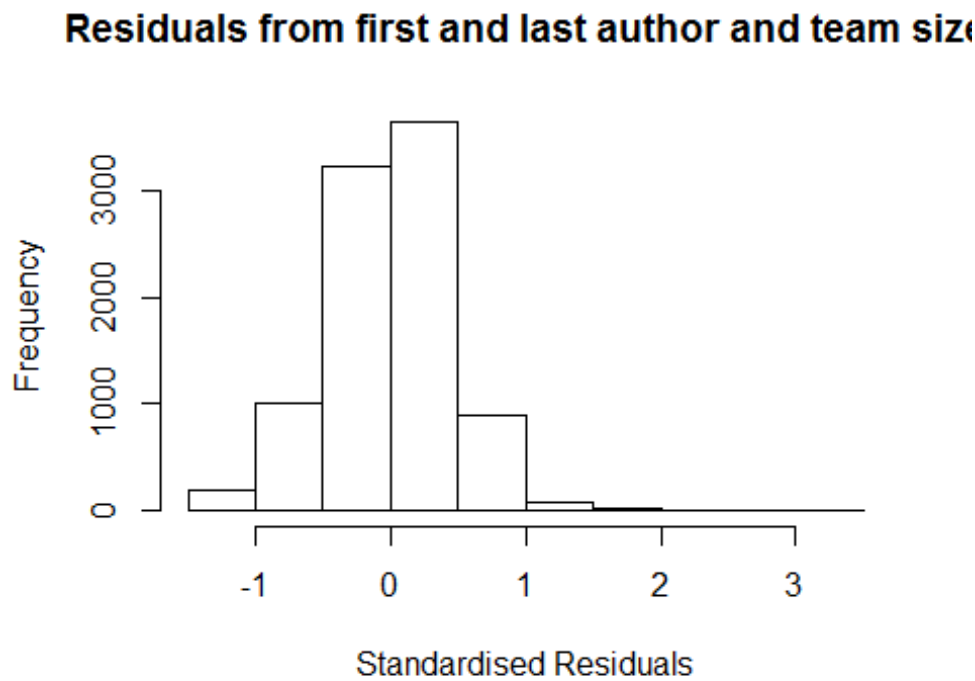
```
## 2011 2012
##  748  638
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  498  453  446  416  440  339  486  491  607  582  658  660  630  596  611
## 2011 2012
##  610  537
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
##  Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 220, df = 16, p-value <2e-16
```




```
##
##  Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 32, df = 1, p-value = 1e-08
```

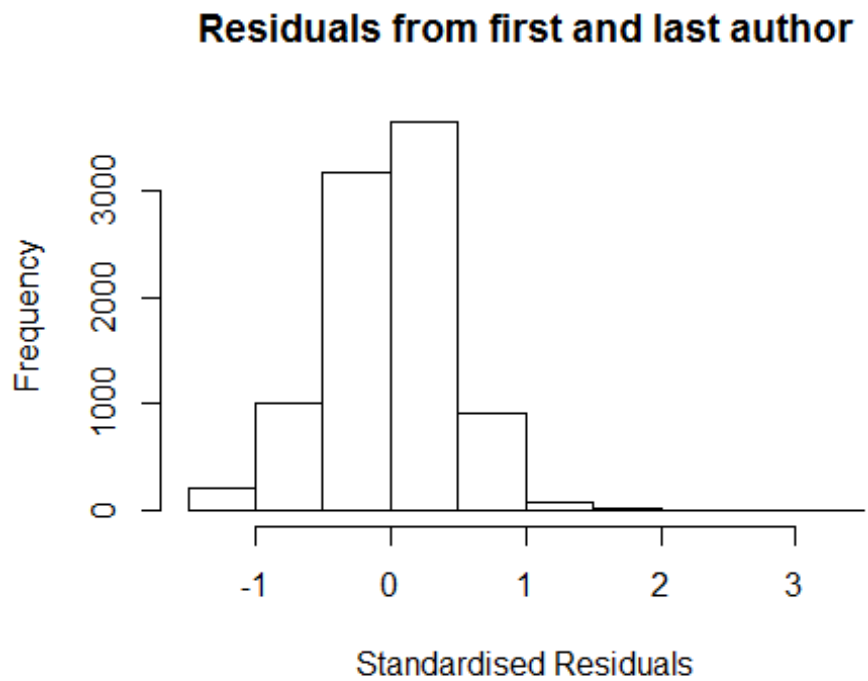



```
## [1] "Female first author team size 2018 geometric mean: 4.08808601852994"
## [1] "Male first author team size 2018 geometric mean: 4.05212568677949"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 22000, p-value = 1
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 4.38576474828248"
## [1] "Male last author team size 2018 geometric mean: 3.98157258116418"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 18000, p-value = 0.08
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##                     GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.034  1           1.017
## LastAuthorFemale  1.034  1           1.017
## UniqueAuthors     1.083  4           1.010
## Year              1.105 16           1.003
```



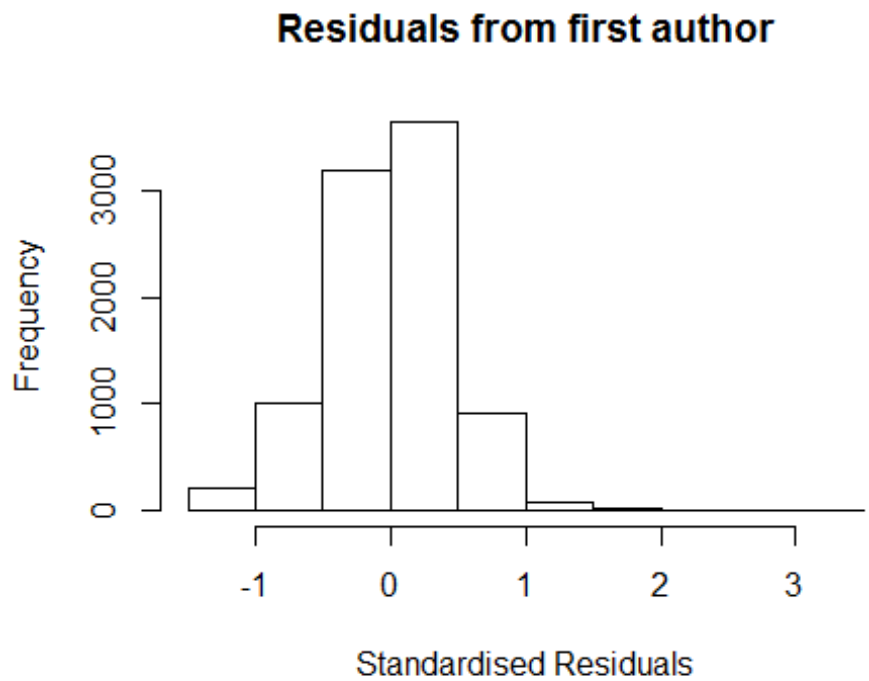

```
## [1] "List of  1 outliers with residuals above  2.5"
##          ScopusId  NLCS Year OneField Fields residuals
## 17031 74549178560 4.208 2010     1315      2     3.004
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
UniqueAuthors +
##     Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##     k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.2549 -0.2821  0.0138  0.2860  3.0035
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)          1.0346     0.0284   36.44  < 2e-16 ***
## FirstAuthorFemale1  -0.0305     0.0101   -3.01  0.00260 **
## LastAuthorFemale1   -0.0161     0.0119   -1.35  0.17829
## UniqueAuthors2       0.1101     0.0231    4.76  2.0e-06 ***
## UniqueAuthors3       0.1255     0.0231    5.43  5.7e-08 ***
## UniqueAuthors4       0.1646     0.0234    7.04  2.1e-12 ***
## UniqueAuthors5       0.2230     0.0224    9.97  < 2e-16 ***
## Year1997             0.0226     0.0305    0.74  0.45795
## Year1998            -0.0027     0.0293   -0.09  0.92663
## Year1999            -0.0554     0.0280   -1.98  0.04816 *
```

```
## Year2000            -0.0493     0.0287   -1.72  0.08536 .
## Year2001            -0.0222     0.0291   -0.76  0.44496
## Year2002            -0.0793     0.0288   -2.75  0.00589 **
## Year2003            -0.0953     0.0289   -3.30  0.00096 ***
## Year2004            -0.0836     0.0275   -3.04  0.00237 **
## Year2005            -0.0793     0.0287   -2.76  0.00573 **
## Year2006            -0.0200     0.0304   -0.66  0.51105
## Year2007            -0.0164     0.0315   -0.52  0.60221
## Year2008            -0.0620     0.0293   -2.12  0.03434 *
## Year2009            -0.0593     0.0275   -2.16  0.03101 *
## Year2010            -0.0532     0.0268   -1.99  0.04705 *
## Year2011            -0.0648     0.0278   -2.33  0.01978 *
## Year2012            -0.1048     0.0295   -3.56  0.00038 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.419
## Multiple R-squared:  0.0245, Adjusted R-squared:  0.0221
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
##  4 observations c(1962,3769,7832,8136)
##   are outliers with |weight| = 0 ( < 1.1e-05);
##  768 weights are ~= 1. The remaining 8288 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0001  0.8680  0.9500  0.9000  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.10e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##                     GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.028  1           1.014
## LastAuthorFemale  1.033  1           1.017
## Year              1.030 16           1.001
```



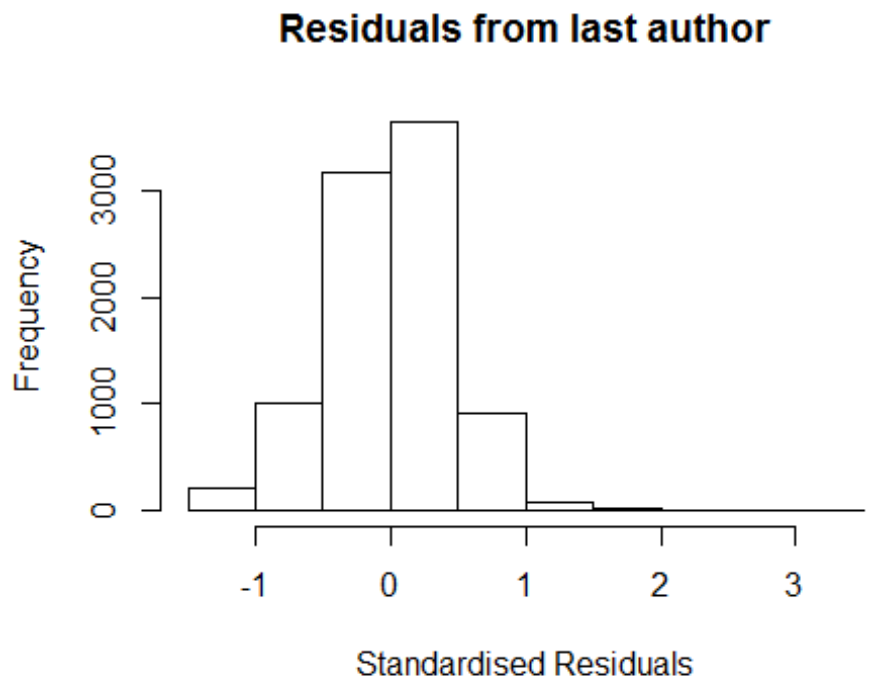

```
## [1] "List of  1 outliers with residuals above  2.5"
##          ScopusId  NLCS Year OneField Fields residuals
## 17031 74549178560 4.208 2010     1315      2     3.069
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.1926 -0.2892  0.0128  0.2875  3.0694
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)          1.16032    0.02214   52.42   <2e-16 ***
## FirstAuthorFemale1 -0.02057    0.01019   -2.02    0.044 *
## LastAuthorFemale1  -0.01371    0.01202   -1.14    0.254
## Year1997            0.03229    0.03068    1.05    0.293
## Year1998            0.00114    0.02944    0.04    0.969
## Year1999           -0.04428    0.02821   -1.57    0.117
## Year2000           -0.03254    0.02896   -1.12    0.261
## Year2001            0.00259    0.02944    0.09    0.930
## Year2002           -0.05216    0.02902   -1.80    0.072 .
## Year2003           -0.06709    0.02910   -2.31    0.021 *
## Year2004           -0.05604    0.02770   -2.02    0.043 *
## Year2005           -0.04797    0.02869   -1.67    0.095 .
```

```
## Year2006            0.01133    0.03033    0.37    0.709
## Year2007            0.00729    0.03178    0.23    0.819
## Year2008           -0.02881    0.02954   -0.98    0.330
## Year2009           -0.02634    0.02760   -0.95    0.340
## Year2010           -0.02169    0.02695   -0.80    0.421
## Year2011           -0.03306    0.02806   -1.18    0.239
## Year2012           -0.07609    0.02980   -2.55    0.011 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.422
## Multiple R-squared:  0.00512,    Adjusted R-squared:  0.00314
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
##  3 observations c(1962,7832,8136)
##   are outliers with |weight| = 0 ( < 1.1e-05);
##  787 weights are ~= 1. The remaining 8270 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0049  0.8670  0.9500  0.8990  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.10e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.014  1           1.007
## Year              1.014 16           1.000
```




```
## [1] "List of  1 outliers with residuals above  2.5"
##          ScopusId  NLCS Year OneField Fields residuals
## 17031 74549178560 4.208 2010     1315      2     3.069
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.1912 -0.2883  0.0137  0.2879  3.0711
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)          1.15850    0.02211   52.40   <2e-16 ***
## FirstAuthorFemale1 -0.02194    0.01012   -2.17    0.030 *
## Year1997             0.03273    0.03069    1.07    0.286
## Year1998             0.00104    0.02946    0.04    0.972
## Year1999            -0.04403    0.02821   -1.56    0.119
## Year2000            -0.03261    0.02895   -1.13    0.260
## Year2001             0.00247    0.02945    0.08    0.933
## Year2002            -0.05229    0.02901   -1.80    0.072 .
## Year2003            -0.06690    0.02911   -2.30    0.022 *
## Year2004            -0.05636    0.02770   -2.03    0.042 *
## Year2005            -0.04821    0.02870   -1.68    0.093 .
## Year2006             0.01093    0.03032    0.36    0.718
```

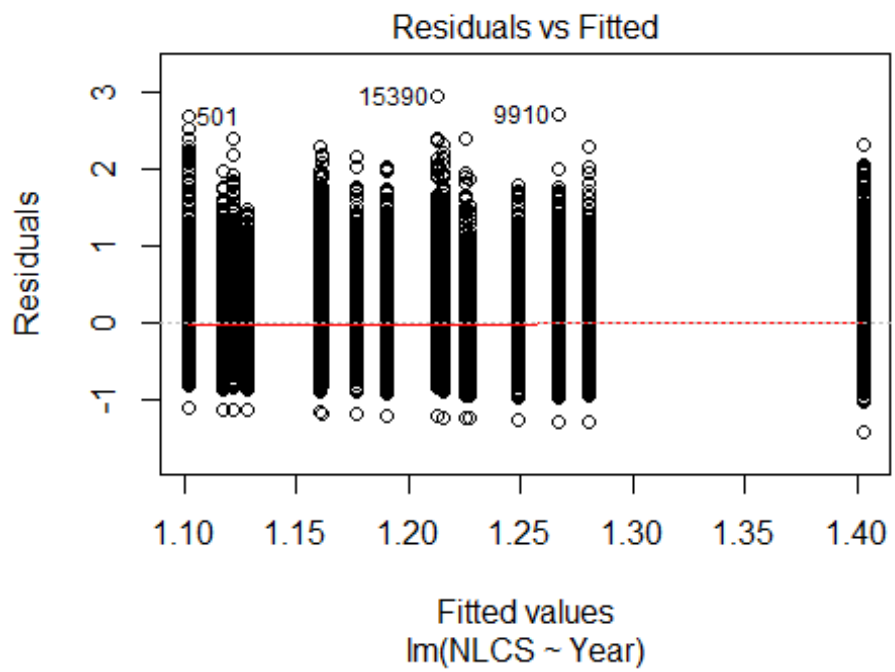
```
## Year2007            0.00665    0.03173    0.21    0.834
## Year2008           -0.02934    0.02953   -0.99    0.320
## Year2009           -0.02696    0.02760   -0.98    0.329
## Year2010           -0.02164    0.02697   -0.80    0.422
## Year2011           -0.03335    0.02807   -1.19    0.235
## Year2012           -0.07660    0.02982   -2.57    0.010 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.422
## Multiple R-squared:  0.00498,    Adjusted R-squared:  0.00311
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
##  3 observations c(1962,7832,8136)
##   are outliers with |weight| = 0 ( < 1.1e-05);
##  778 weights are ~= 1. The remaining 8279 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0046  0.8670  0.9500  0.8990  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.10e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##                   GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.019  1           1.009
## Year             1.019 16           1.001
```



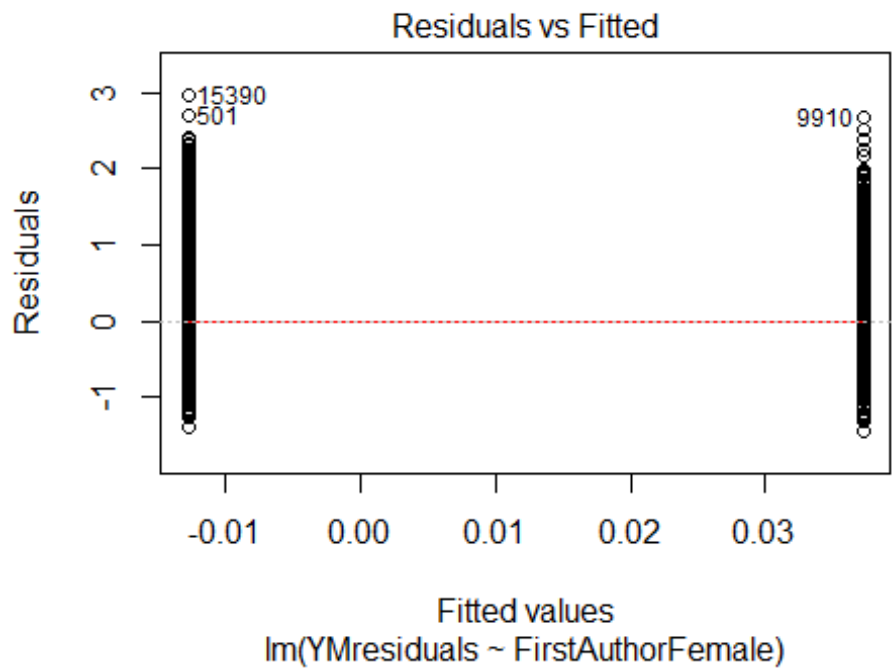

```
## [1] "List of  1 outliers with residuals above  2.5"
##          ScopusId  NLCS Year OneField Fields residuals
## 17031 74549178560 4.208 2010     1315      2     3.069
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.1870 -0.2884  0.0126  0.2878  3.0741
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.15522    0.02195   52.64   <2e-16 ***
## LastAuthorFemale1 -0.01663    0.01194   -1.39    0.164
## Year1997            0.03181    0.03070    1.04    0.300
## Year1998            0.00141    0.02943    0.05    0.962
## Year1999           -0.04432    0.02824   -1.57    0.117
## Year2000           -0.03407    0.02895   -1.18    0.239
## Year2001            0.00167    0.02946    0.06    0.955
## Year2002           -0.05213    0.02902   -1.80    0.072 .
## Year2003           -0.06716    0.02911   -2.31    0.021 *
## Year2004           -0.05606    0.02769   -2.02    0.043 *
## Year2005           -0.04839    0.02870   -1.69    0.092 .
## Year2006            0.01033    0.03032    0.34    0.733
```

```
## Year2007           0.00615    0.03177    0.19    0.846
## Year2008          -0.03019    0.02955   -1.02    0.307
## Year2009          -0.02653    0.02761   -0.96    0.337
## Year2010          -0.02130    0.02695   -0.79    0.429
## Year2011          -0.03298    0.02806   -1.18    0.240
## Year2012          -0.07626    0.02980   -2.56    0.011 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.423
## Multiple R-squared:  0.00465,    Adjusted R-squared:  0.00278
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
##  3 observations c(1962,7832,8136)
##   are outliers with |weight| = 0 ( < 1.1e-05);
##  778 weights are ~= 1. The remaining 8279 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0042  0.8660  0.9490  0.9000  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.10e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Sample size for the above analysis:  9060"
## [1] ""
## [1] ""
## [1] "##################################"
## [1] "Analysis of AJSC 1400"
## [1] "##################################"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  810  748  679  647  791  864  768  558  592  693  843  914 1115 1210 1210
## 2011 2012
## 1068  970
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
```

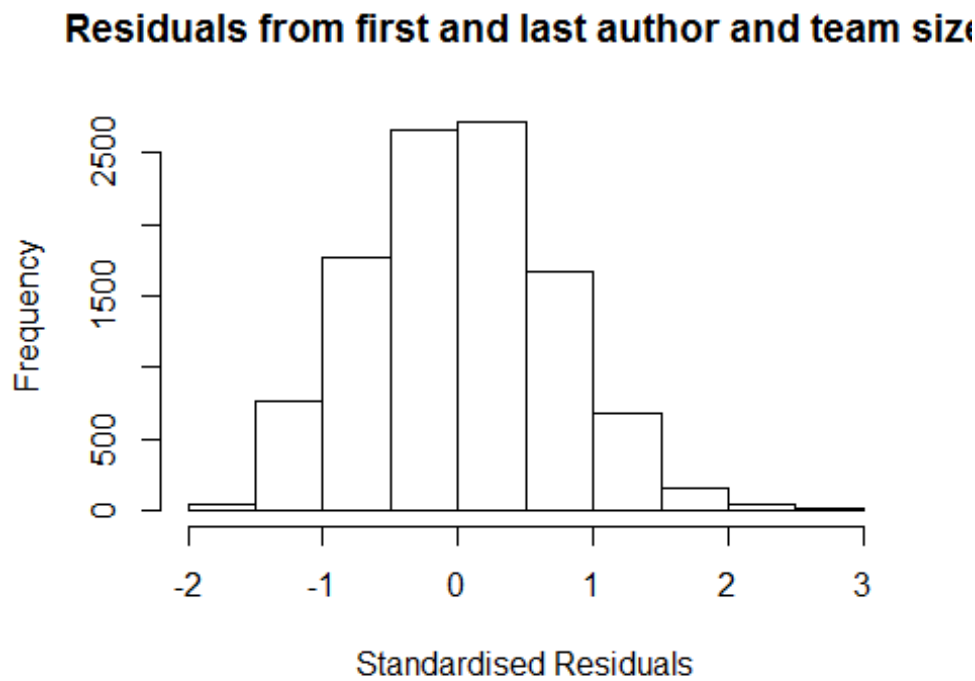
```
##  664  609  497  479  575  638  634  457  489  555  685  760  887  990  956
## 2011 2012
##  845  790
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  629  564  452  447  534  590  579  419  442  506  616  694  797  892  872
## 2011 2012
##  751  705
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
##  Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 190, df = 16, p-value <2e-16
```




```
##
##  Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 2.7, df = 1, p-value = 0.1
```



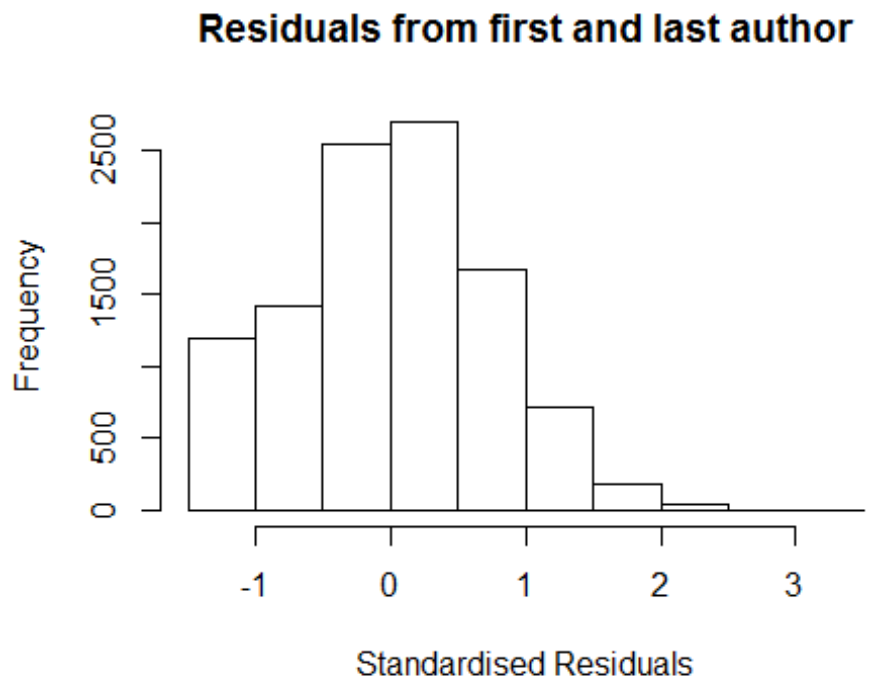

```
## [1] "Female first author team size 2018 geometric mean: 1.81680051053358"
## [1] "Male first author team size 2018 geometric mean: 1.91045774583467"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 36000, p-value = 0.3
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 1.83487901686948"
## [1] "Male last author team size 2018 geometric mean: 1.9029109681069"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 37000, p-value = 0.4
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.402  1           1.184
## LastAuthorFemale  1.396  1           1.181
## UniqueAuthors     1.054  4           1.007
## Year              1.061 16           1.002
```




```
## [1] "List of  7 outliers with residuals above  2.5"
##          ScopusId  NLCS Year OneField Fields residuals
## 501    0030305327 3.789 1996     1400      4     2.607
## 601    0030530698 3.505 1996     1400      3     2.617
## 1556  21744459955 3.522 1997     1400      3     2.600
## 6889   2342620667 3.631 2004     1400      1     2.619
## 9910  34047207566 3.996 2007     1400      4     2.643
## 14955 79955593175 3.584 2011     1400      4     2.610
## 15390 84906331429 4.168 2011     1400      5     2.914
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
UniqueAuthors +
##     Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##     k.max = 1000))
##  \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.68119 -0.48674  0.00344  0.48331  2.91431
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)          0.88840    0.03010   29.52  < 2e-16 ***
## FirstAuthorFemale1   0.03313    0.01919    1.73  0.08435 .
## LastAuthorFemale1    0.01351    0.01935    0.70  0.48505
## UniqueAuthors2       0.27994    0.01641   17.06  < 2e-16 ***
```

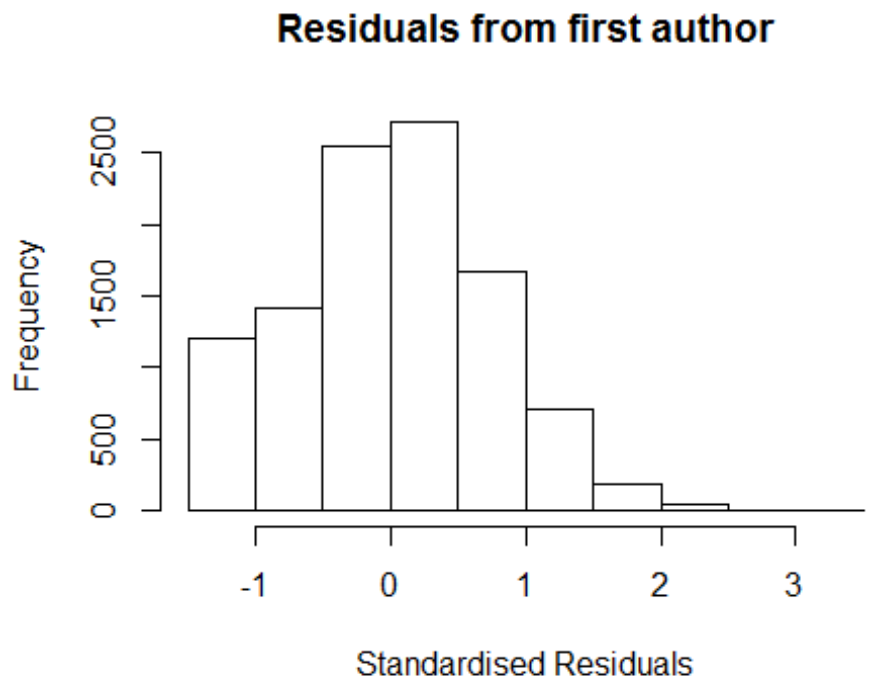
```
## UniqueAuthors3      0.42137    0.02072   20.33  < 2e-16 ***
## UniqueAuthors4      0.48228    0.04070   11.85  < 2e-16 ***
## UniqueAuthors5      0.55756    0.05022   11.10  < 2e-16 ***
## Year1997            0.03400    0.04177    0.81  0.41558
## Year1998            0.06265    0.04286    1.46  0.14385
## Year1999            0.07255    0.04145    1.75  0.08014 .
## Year2000            0.05589    0.03966    1.41  0.15882
## Year2001            0.11258    0.04116    2.74  0.00624 **
## Year2002            0.00113    0.04037    0.03  0.97762
## Year2003            0.13799    0.04225    3.27  0.00109 **
## Year2004            0.12310    0.04675    2.63  0.00848 **
## Year2005            0.12209    0.04225    2.89  0.00387 **
## Year2006            0.04253    0.03980    1.07  0.28531
## Year2007            0.13798    0.04106    3.36  0.00078 ***
## Year2008            0.16215    0.03935    4.12  3.8e-05 ***
## Year2009            0.05299    0.03870    1.37  0.17088
## Year2010            0.09936    0.03955    2.51  0.01201 *
## Year2011            0.08535    0.04086    2.09  0.03674 *
## Year2012            0.27738    0.04579    6.06  1.4e-09 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.718
## Multiple R-squared:  0.0729, Adjusted R-squared:  0.0709
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  922 weights are ~= 1. The remaining 9567 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0622  0.8610  0.9490  0.9100  0.9860  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          9.53e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##                     GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.411  1           1.188
```

```
## LastAuthorFemale  1.407  1           1.186
## Year              1.017 16           1.001
```



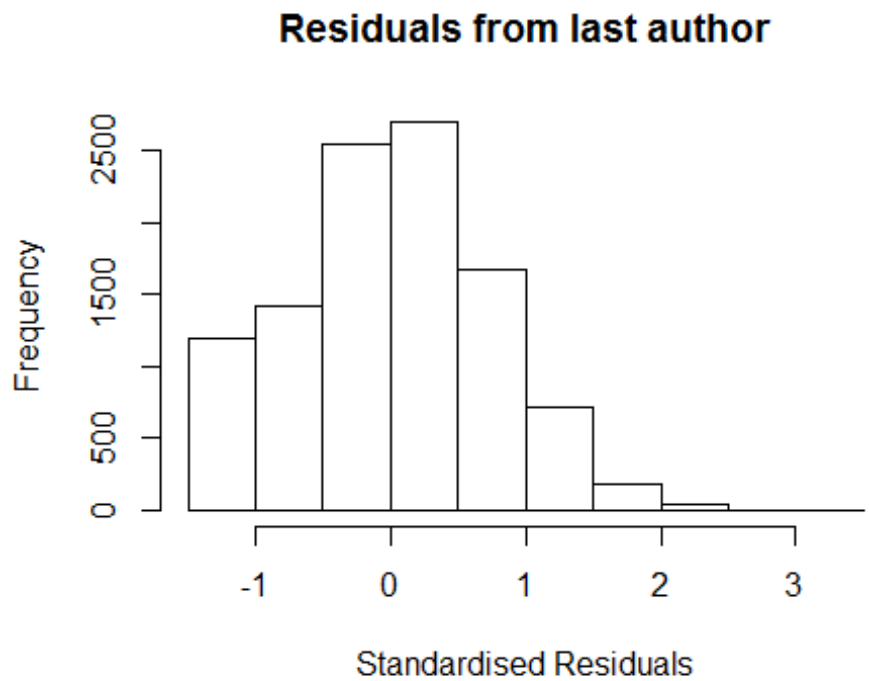

```
## [1] "List of  3 outliers with residuals above  2.5"
##          ScopusId  NLCS Year OneField Fields residuals
## 501    0030305327 3.789 1996     1400      4     2.738
## 9910  34047207566 3.996 2007     1400      4     2.701
## 15390 84906331429 4.168 2011     1400      5     3.015
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.4448 -0.4991  0.0133  0.4991  3.0154
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)           1.0258     0.0305   33.61  < 2e-16 ***
## FirstAuthorFemale1    0.0483     0.0199    2.42  0.01539 *
## LastAuthorFemale1     0.0247     0.0200    1.23  0.21696
## Year1997              0.0482     0.0429    1.12  0.26143
## Year1998              0.0936     0.0445    2.10  0.03556 *
## Year1999              0.1052     0.0425    2.48  0.01325 *
## Year2000              0.0742     0.0408    1.82  0.06918 .
```

```
## Year2001             0.1373     0.0422    3.26  0.00113 **
## Year2002             0.0418     0.0423    0.99  0.32223
## Year2003             0.1674     0.0429    3.90  9.7e-05 ***
## Year2004             0.1786     0.0483    3.70  0.00022 ***
## Year2005             0.1817     0.0430    4.23  2.4e-05 ***
## Year2006             0.0965     0.0411    2.35  0.01891 *
## Year2007             0.1958     0.0422    4.64  3.5e-06 ***
## Year2008             0.2079     0.0408    5.10  3.5e-07 ***
## Year2009             0.0867     0.0399    2.18  0.02960 *
## Year2010             0.1386     0.0413    3.35  0.00080 ***
## Year2011             0.1269     0.0430    2.95  0.00315 **
## Year2012             0.3460     0.0479    7.23  5.1e-13 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.741
## Multiple R-squared:  0.0131, Adjusted R-squared:  0.0114
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  880 weights are ~= 1. The remaining 9609 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.060   0.867   0.951   0.911   0.986   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          9.53e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##                   GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.01  1           1.005
## Year              1.01 16           1.000
```




```
## [1] "List of  3 outliers with residuals above  2.5"
##          ScopusId  NLCS Year OneField Fields residuals
## 501    0030305327 3.789 1996     1400      4     2.738
## 9910  34047207566 3.996 2007     1400      4     2.701
## 15390 84906331429 4.168 2011     1400      5     3.015
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.4363 -0.4991  0.0128  0.4967  3.0128
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)           1.0284     0.0304   33.87  < 2e-16 ***
## FirstAuthorFemale1    0.0619     0.0168    3.68  0.00024 ***
## Year1997              0.0485     0.0429    1.13  0.25891
## Year1998              0.0926     0.0445    2.08  0.03736 *
## Year1999              0.1049     0.0424    2.47  0.01342 *
## Year2000              0.0742     0.0408    1.82  0.06926 .
## Year2001              0.1374     0.0422    3.26  0.00113 **
## Year2002              0.0418     0.0423    0.99  0.32223
## Year2003              0.1669     0.0429    3.89  0.00010 ***
## Year2004              0.1789     0.0483    3.70  0.00021 ***
```

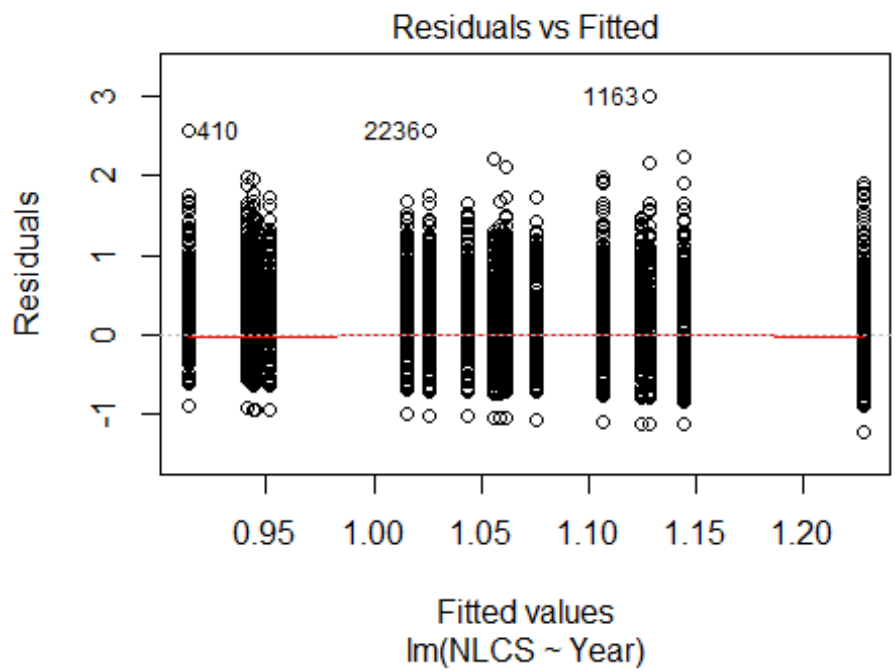
```
## Year2005             0.1816     0.0430    4.22  2.4e-05 ***
## Year2006             0.0972     0.0412    2.36  0.01816 *
## Year2007             0.1964     0.0422    4.66  3.3e-06 ***
## Year2008             0.2085     0.0408    5.11  3.3e-07 ***
## Year2009             0.0871     0.0399    2.18  0.02892 *
## Year2010             0.1392     0.0413    3.37  0.00076 ***
## Year2011             0.1268     0.0430    2.95  0.00316 **
## Year2012             0.3460     0.0479    7.23  5.3e-13 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.741
## Multiple R-squared:  0.0129, Adjusted R-squared:  0.0113
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  884 weights are ~= 1. The remaining 9605 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0606  0.8670  0.9510  0.9110  0.9860  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          9.53e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##                   GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.007  1           1.004
## Year             1.007 16           1.000
```



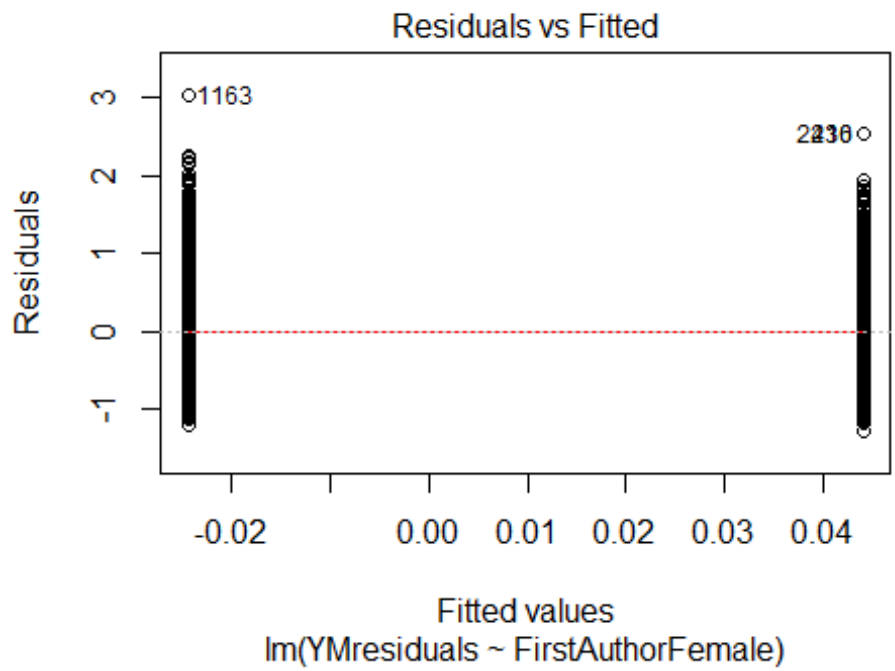

```
## [1] "List of  3 outliers with residuals above  2.5"
##          ScopusId  NLCS Year OneField Fields residuals
## 501    0030305327 3.789 1996     1400      4     2.738
## 9910  34047207566 3.996 2007     1400      4     2.701
## 15390 84906331429 4.168 2011     1400      5     3.015
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.4298 -0.4992  0.0117  0.4985  3.0085
##
## Coefficients:
##                   Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.0310     0.0305   33.83  < 2e-16 ***
## LastAuthorFemale1   0.0515     0.0169    3.05  0.00232 **
## Year1997            0.0492     0.0430    1.14  0.25254
## Year1998            0.0942     0.0446    2.11  0.03447 *
## Year1999            0.1057     0.0424    2.49  0.01281 *
## Year2000            0.0752     0.0408    1.84  0.06551 .
## Year2001            0.1380     0.0422    3.27  0.00109 **
## Year2002            0.0434     0.0423    1.03  0.30506
## Year2003            0.1669     0.0429    3.89  0.00010 ***
## Year2004            0.1793     0.0484    3.71  0.00021 ***
```

```
## Year2005            0.1835     0.0430    4.27  2.0e-05 ***
## Year2006            0.0951     0.0411    2.31  0.02076 *
## Year2007            0.1956     0.0422    4.64  3.6e-06 ***
## Year2008            0.2075     0.0408    5.08  3.8e-07 ***
## Year2009            0.0861     0.0399    2.16  0.03092 *
## Year2010            0.1390     0.0413    3.36  0.00078 ***
## Year2011            0.1285     0.0429    2.99  0.00278 **
## Year2012            0.3473     0.0478    7.26  4.2e-13 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.741
## Multiple R-squared:  0.0126, Adjusted R-squared:  0.0109
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  886 weights are ~= 1. The remaining 9603 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0624  0.8680  0.9500  0.9110  0.9860  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          9.53e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                  psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Sample size for the above analysis:  10489"
## [1] ""
## [1] ""
## [1] "##################################"
## [1] "Analysis of AJSC 1401"
## [1] "##################################"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  393  350  368  336  312  320  340  281  260  273  313  330  350  415  336
## 2011 2012
##  452  446
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
```

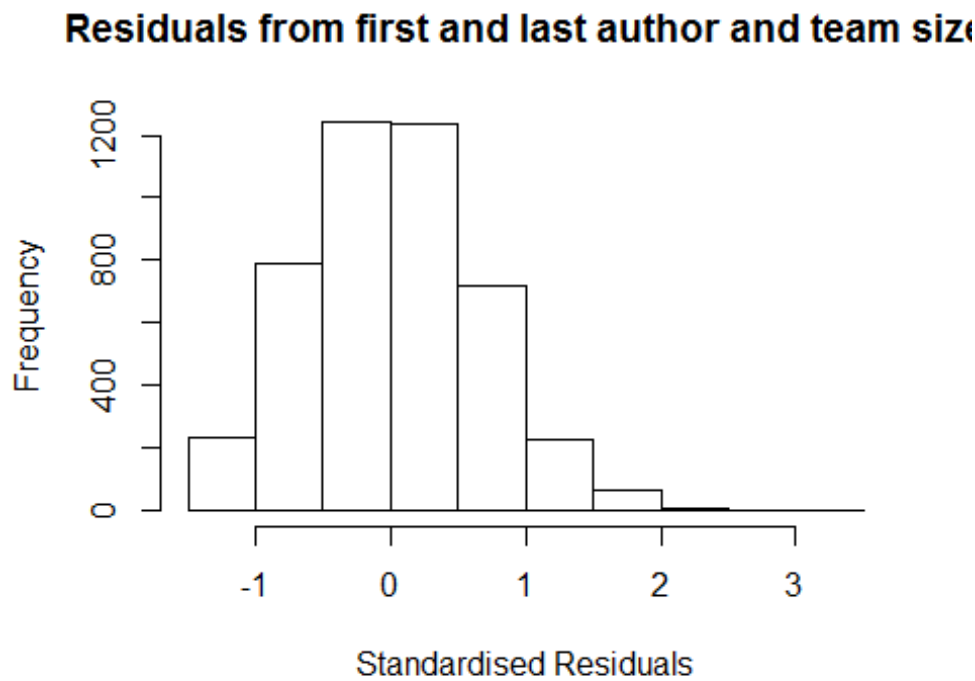
```
##  336  312  322  262  247  258  300  238  220  223  252  278  294  344  271
## 2011 2012
##  363  373
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  309  286  301  238  231  239  279  221  205  208  227  252  267  319  253
## 2011 2012
##  333  342
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
##  Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 35, df = 16, p-value = 0.004
```




```
##
##  Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 0.087, df = 1, p-value = 0.8
```

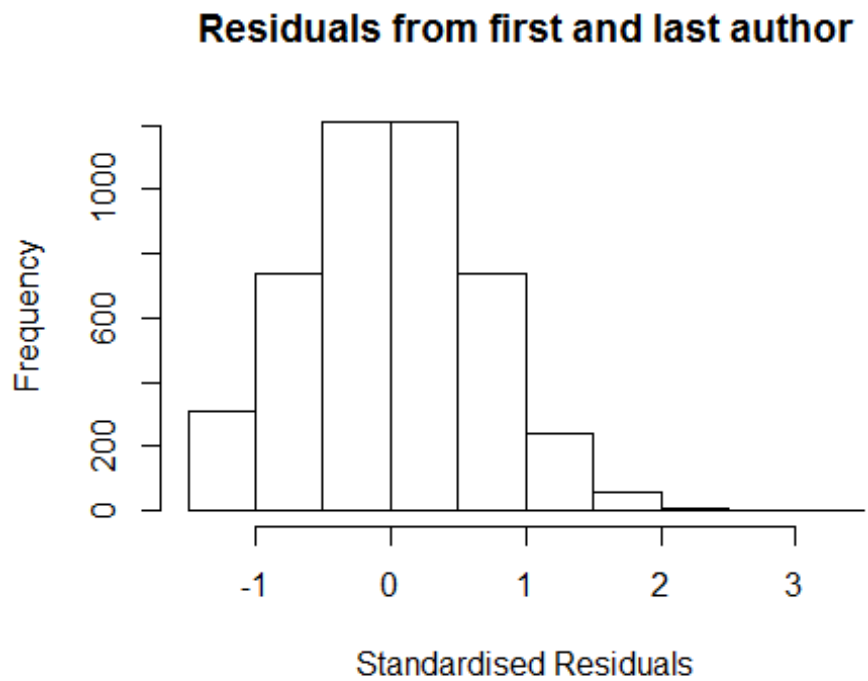



```
## [1] "Female first author team size 2018 geometric mean: 1.88998599815768"
## [1] "Male first author team size 2018 geometric mean: 1.80936349947781"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 7800, p-value = 0.7
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 1.88809316720933"
## [1] "Male last author team size 2018 geometric mean: 1.80558302744141"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 8300, p-value = 0.6
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##                     GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.495  1           1.223
## LastAuthorFemale  1.505  1           1.227
## UniqueAuthors     1.073  4           1.009
## Year              1.086 16           1.003
```



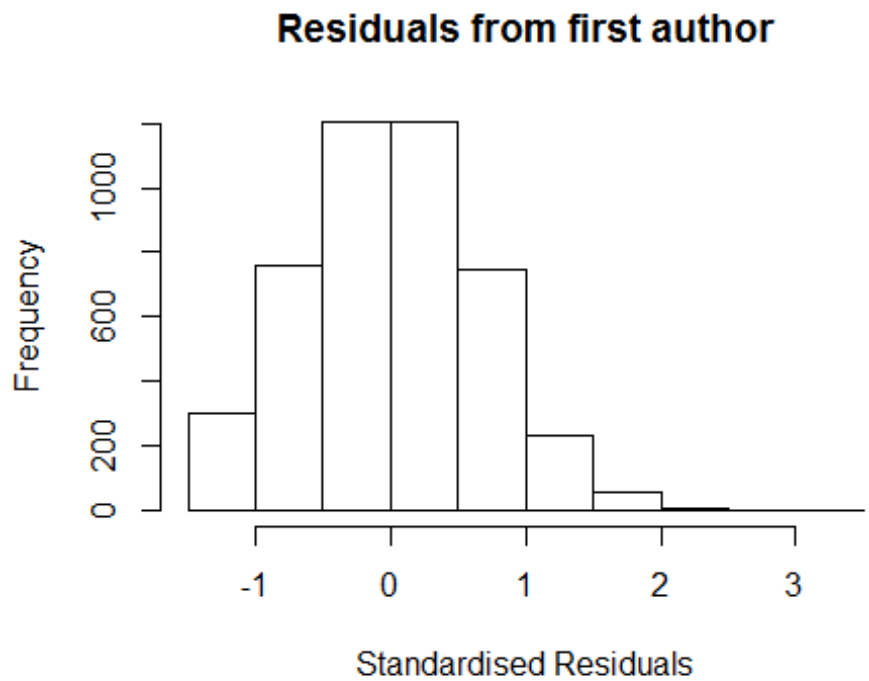

```
## [1] "List of  3 outliers with residuals above  2.5"
##         ScopusId  NLCS Year OneField Fields residuals
## 410   0000171195 3.493 1997     1401      2     2.565
## 1163  0000908020 4.133 1999     1401      2     3.155
## 2236 67549103767 3.604 2002     1401      2     2.511
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
UniqueAuthors +
##     Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##     k.max = 1000))
##  \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.36428 -0.45339 -0.00117  0.45225  3.15501
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)          0.7829     0.0404   19.35  < 2e-16 ***
## FirstAuthorFemale1   0.0190     0.0255    0.74  0.45705
## LastAuthorFemale1    0.0653     0.0258    2.54  0.01126 *
## UniqueAuthors2       0.1852     0.0236    7.85  5.1e-15 ***
## UniqueAuthors3       0.2227     0.0277    8.05  1.1e-15 ***
## UniqueAuthors4       0.2487     0.0463    5.37  8.1e-08 ***
## UniqueAuthors5       0.2584     0.0886    2.92  0.00357 **
## Year1997            -0.0591     0.0540   -1.09  0.27380
```

```
## Year1998             0.0371     0.0518    0.72  0.47339
## Year1999             0.1951     0.0589    3.31  0.00093 ***
## Year2000             0.1630     0.0589    2.77  0.00566 **
## Year2001             0.1918     0.0590    3.25  0.00117 **
## Year2002             0.0686     0.0547    1.25  0.21003
## Year2003            -0.0198     0.0572   -0.35  0.72957
## Year2004             0.1090     0.0598    1.82  0.06822 .
## Year2005             0.2744     0.0664    4.13  3.7e-05 ***
## Year2006             0.1178     0.0574    2.05  0.04021 *
## Year2007             0.0752     0.0575    1.31  0.19096
## Year2008             0.1445     0.0580    2.49  0.01280 *
## Year2009             0.1765     0.0510    3.46  0.00054 ***
## Year2010             0.1379     0.0536    2.57  0.01014 *
## Year2011             0.1106     0.0518    2.14  0.03277 *
## Year2012            -0.0213     0.0540   -0.39  0.69328
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.674
## Multiple R-squared:  0.0443, Adjusted R-squared:  0.0396
## Convergence in 10 IRWLS iterations
##
## Robustness weights:
##  observation 897 is an outlier with |weight| <= 6.4e-08 ( < 2.2e-05);
##  396 weights are ~= 1. The remaining 4113 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.115   0.873   0.951   0.914   0.985   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          2.22e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.506  1           1.227
## LastAuthorFemale  1.512  1           1.230
## Year              1.019 16           1.001
```



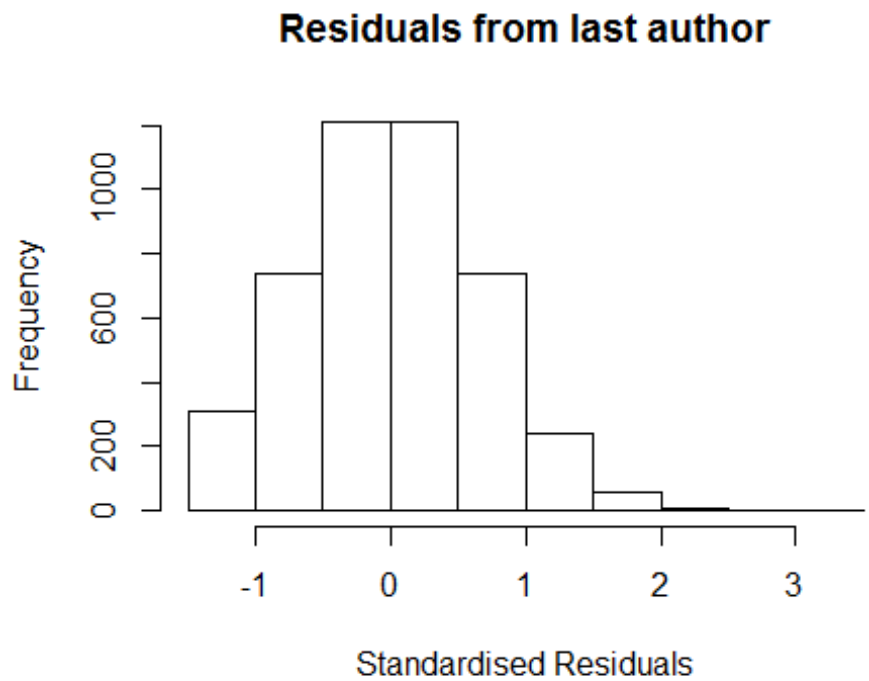

```
## [1] "List of  3 outliers with residuals above  2.5"
##         ScopusId  NLCS Year OneField Fields residuals
## 410   0000171195 3.493 1997     1401      2     2.641
## 1163  0000908020 4.133 1999     1401      2     3.067
## 2236 67549103767 3.604 2002     1401      2     2.623
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.26088 -0.46633 -0.00222  0.45591  3.06722
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)          0.8928     0.0385   23.16  < 2e-16 ***
## FirstAuthorFemale1   0.0258     0.0260    0.99  0.32125
## LastAuthorFemale1    0.0681     0.0262    2.60  0.00946 **
## Year1997            -0.0662     0.0543   -1.22  0.22319
## Year1998             0.0264     0.0514    0.51  0.60702
## Year1999             0.1730     0.0587    2.95  0.00323 **
## Year2000             0.1346     0.0586    2.30  0.02174 *
## Year2001             0.1970     0.0596    3.31  0.00095 ***
## Year2002             0.0624     0.0552    1.13  0.25876
## Year2003            -0.0206     0.0577   -0.36  0.72068
```

```
## Year2004             0.1100     0.0599    1.84  0.06626 .
## Year2005             0.2742     0.0665    4.12  3.8e-05 ***
## Year2006             0.1321     0.0582    2.27  0.02325 *
## Year2007             0.0800     0.0578    1.39  0.16611
## Year2008             0.1410     0.0588    2.40  0.01642 *
## Year2009             0.1862     0.0511    3.64  0.00027 ***
## Year2010             0.1346     0.0538    2.50  0.01234 *
## Year2011             0.1162     0.0530    2.19  0.02827 *
## Year2012            -0.0204     0.0562   -0.36  0.71643
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.683
## Multiple R-squared:  0.0206, Adjusted R-squared:  0.0167
## Convergence in 10 IRWLS iterations
##
## Robustness weights:
##  365 weights are ~= 1. The remaining 4145 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0065  0.8680  0.9510  0.9150  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          2.22e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.008  1           1.004
## Year              1.008 16           1.000
```




```
## [1] "List of  3 outliers with residuals above  2.5"
##         ScopusId  NLCS Year OneField Fields residuals
## 410   0000171195 3.493 1997     1401      2     2.641
## 1163  0000908020 4.133 1999     1401      2     3.067
## 2236 67549103767 3.604 2002     1401      2     2.623
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.24286 -0.46874 -0.00211  0.45534  3.05594
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)          0.9012     0.0386   23.37  < 2e-16 ***
## FirstAuthorFemale1   0.0645     0.0212    3.04  0.00240 **
## Year1997            -0.0615     0.0545   -1.13  0.25940
## Year1998             0.0263     0.0516    0.51  0.60967
## Year1999             0.1758     0.0588    2.99  0.00279 **
## Year2000             0.1350     0.0585    2.31  0.02103 *
## Year2001             0.2009     0.0597    3.37  0.00077 ***
## Year2002             0.0647     0.0553    1.17  0.24181
## Year2003            -0.0185     0.0578   -0.32  0.74903
## Year2004             0.1129     0.0599    1.88  0.05966 .
```

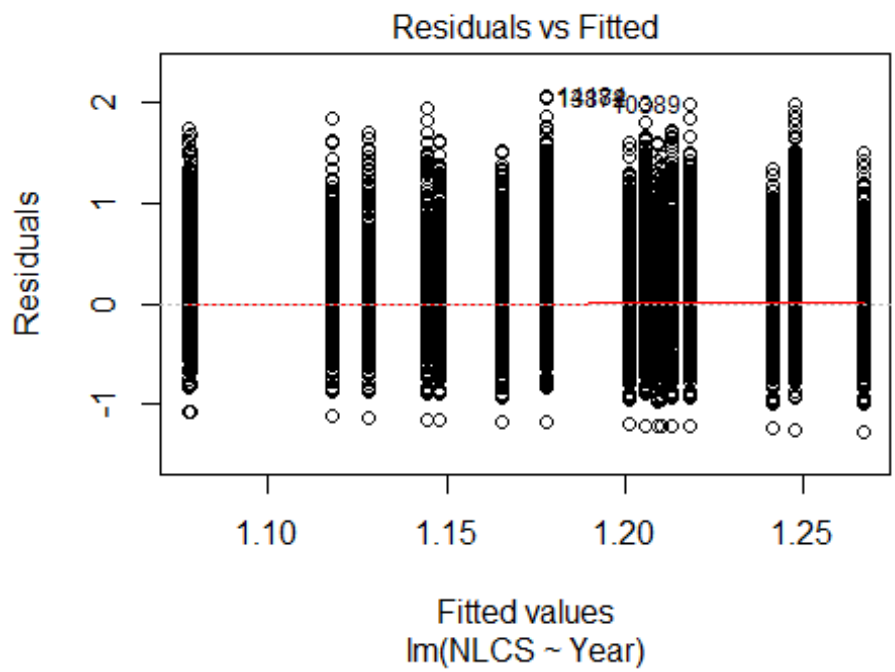
```
## Year2005             0.2772     0.0668    4.15  3.4e-05 ***
## Year2006             0.1348     0.0585    2.30  0.02125 *
## Year2007             0.0839     0.0579    1.45  0.14767
## Year2008             0.1440     0.0589    2.44  0.01457 *
## Year2009             0.1851     0.0511    3.62  0.00030 ***
## Year2010             0.1347     0.0540    2.49  0.01267 *
## Year2011             0.1160     0.0531    2.18  0.02898 *
## Year2012            -0.0210     0.0564   -0.37  0.70960
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.683
## Multiple R-squared:  0.0191, Adjusted R-squared:  0.0153
## Convergence in 10 IRWLS iterations
##
## Robustness weights:
##  378 weights are ~= 1. The remaining 4132 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.008   0.866   0.951   0.915   0.985   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          2.22e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                  psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##                   GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.013  1           1.006
## Year             1.013 16           1.000
```



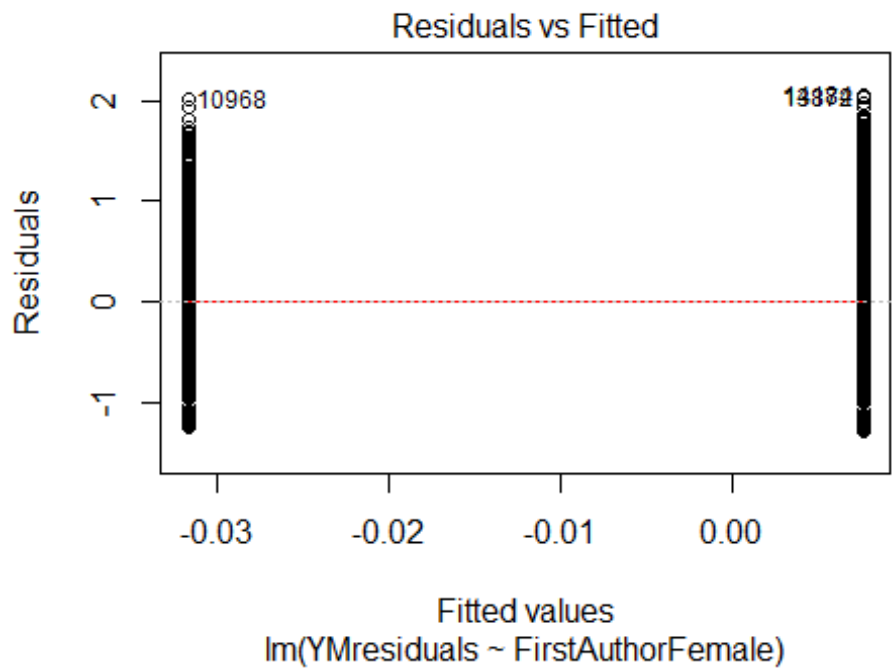

```
## [1] "List of  3 outliers with residuals above  2.5"
##         ScopusId  NLCS Year OneField Fields residuals
## 410   0000171195 3.493 1997     1401      2     2.641
## 1163  0000908020 4.133 1999     1401      2     3.067
## 2236 67549103767 3.604 2002     1401      2     2.623
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.25356 -0.46042 -0.00204  0.45657  3.06354
##
## Coefficients:
##                   Estimate Std. Error t value Pr(>|t|)
## (Intercept)         0.8977     0.0384   23.40  < 2e-16 ***
## LastAuthorFemale1   0.0829     0.0214    3.87  0.00011 ***
## Year1997           -0.0675     0.0543   -1.24  0.21416
## Year1998            0.0262     0.0514    0.51  0.61075
## Year1999            0.1718     0.0587    2.93  0.00345 **
## Year2000            0.1340     0.0587    2.28  0.02251 *
## Year2001            0.1964     0.0596    3.30  0.00099 ***
## Year2002            0.0615     0.0552    1.11  0.26580
## Year2003           -0.0212     0.0577   -0.37  0.71401
## Year2004            0.1082     0.0600    1.80  0.07134 .
```

```
## Year2005            0.2730     0.0664    4.11    4e-05 ***
## Year2006            0.1311     0.0582    2.25  0.02427 *
## Year2007            0.0794     0.0579    1.37  0.17027
## Year2008            0.1398     0.0587    2.38  0.01723 *
## Year2009            0.1854     0.0511    3.63  0.00029 ***
## Year2010            0.1340     0.0537    2.49  0.01268 *
## Year2011            0.1147     0.0530    2.16  0.03055 *
## Year2012           -0.0217     0.0563   -0.38  0.70029
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.683
## Multiple R-squared:  0.0204, Adjusted R-squared:  0.0167
## Convergence in 10 IRWLS iterations
##
## Robustness weights:
##  386 weights are ~= 1. The remaining 4124 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0067  0.8670  0.9510  0.9150  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          2.22e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                  psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Sample size for the above analysis:  4510"
## [1] ""
## [1] ""
## [1] "##################################"
## [1] "Analysis of AJSC 1402"
## [1] "##################################"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  658  562  573  607  675  702  737  516  527  561  720  727  915 1055 1066
## 2011 2012
## 1149 1092
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
```

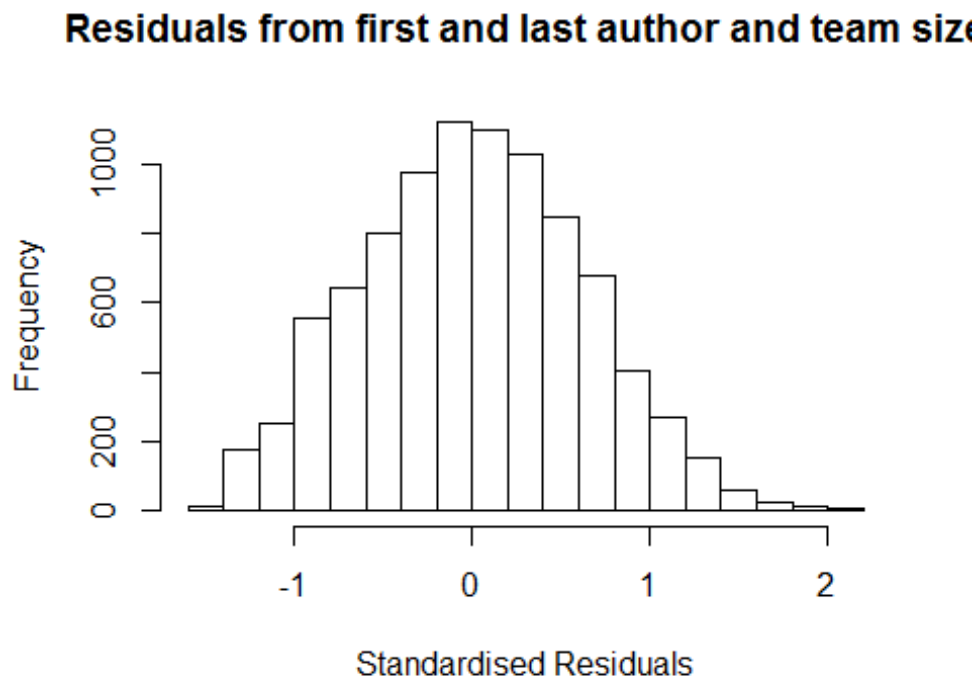
```
##  547  466  487  518  579  589  646  426  430  457  584  577  723  848  854
## 2011 2012
##  919  876
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  481  416  439  479  542  519  563  384  366  388  496  481  616  712  720
## 2011 2012
##  774  734
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
##  Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 130, df = 16, p-value <2e-16
```




```
##
##  Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 1.2, df = 1, p-value = 0.3
```

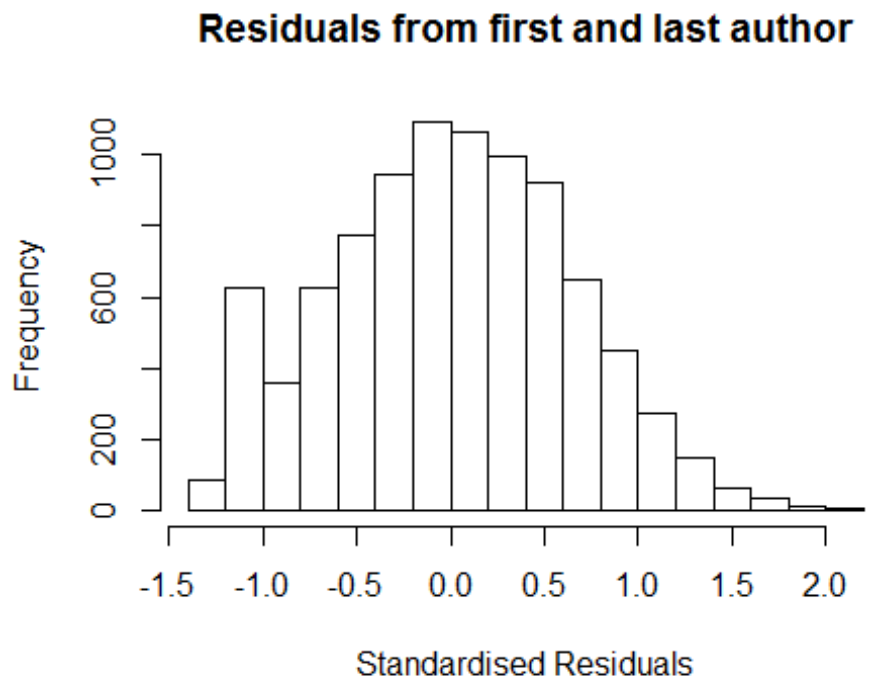



```
## [1] "Female first author team size 2018 geometric mean: 1.86487624539383"
## [1] "Male first author team size 2018 geometric mean: 1.91095091365726"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 31000, p-value = 0.6
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 1.86104104404431"
## [1] "Male last author team size 2018 geometric mean: 1.91163925160065"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 30000, p-value = 0.6
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.206  1           1.098
## LastAuthorFemale  1.214  1           1.102
## UniqueAuthors     1.042  4           1.005
## Year              1.056 16           1.002
```



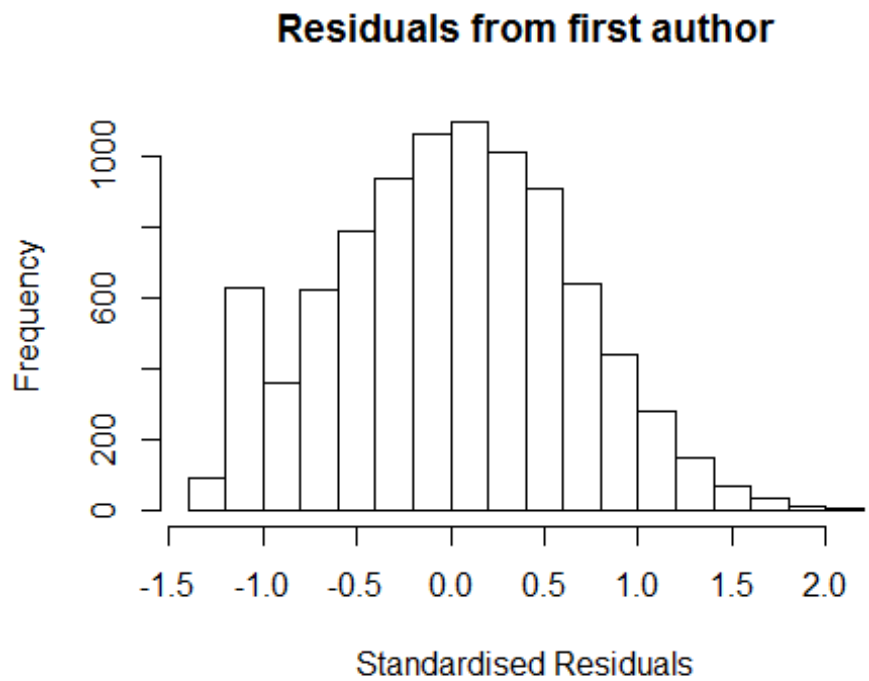

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
UniqueAuthors +
##     Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##     k.max = 1000))
##  \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.47474 -0.43762  0.00321  0.43643  2.10179
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         0.97821    0.02976   32.87   <2e-16 ***
## FirstAuthorFemale1 -0.01698    0.01907   -0.89    0.373
## LastAuthorFemale1  -0.03861    0.01894   -2.04    0.041 *
## UniqueAuthors2      0.22771    0.01599   14.24   <2e-16 ***
## UniqueAuthors3      0.34531    0.01836   18.81   <2e-16 ***
## UniqueAuthors4      0.43308    0.03840   11.28   <2e-16 ***
## UniqueAuthors5     -0.00790    0.09311   -0.08    0.932
## Year1997           -0.03375    0.03931   -0.86    0.391
## Year1998           -0.01462    0.04033   -0.36    0.717
## Year1999           -0.04636    0.04149   -1.12    0.264
```

```
## Year2000           -0.07418    0.04000   -1.85    0.064 .
## Year2001           -0.00238    0.04107   -0.06    0.954
## Year2002            0.01531    0.03894    0.39    0.694
## Year2003            0.06344    0.04040    1.57    0.116
## Year2004            0.08896    0.04115    2.16    0.031 *
## Year2005            0.04328    0.04194    1.03    0.302
## Year2006            0.03851    0.03961    0.97    0.331
## Year2007            0.06846    0.03853    1.78    0.076 .
## Year2008            0.03928    0.04013    0.98    0.328
## Year2009            0.02815    0.03917    0.72    0.472
## Year2010            0.05732    0.03866    1.48    0.138
## Year2011           -0.00464    0.03736   -0.12    0.901
## Year2012           -0.00891    0.03983   -0.22    0.823
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.646
## Multiple R-squared:  0.0561, Adjusted R-squared:  0.0538
## Convergence in 10 IRWLS iterations
##
## Robustness weights:
##  745 weights are ~= 1. The remaining 8365 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.269   0.876   0.950   0.914   0.985   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.10e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.193  1           1.092
## LastAuthorFemale  1.197  1           1.094
## Year              1.018 16           1.001
```



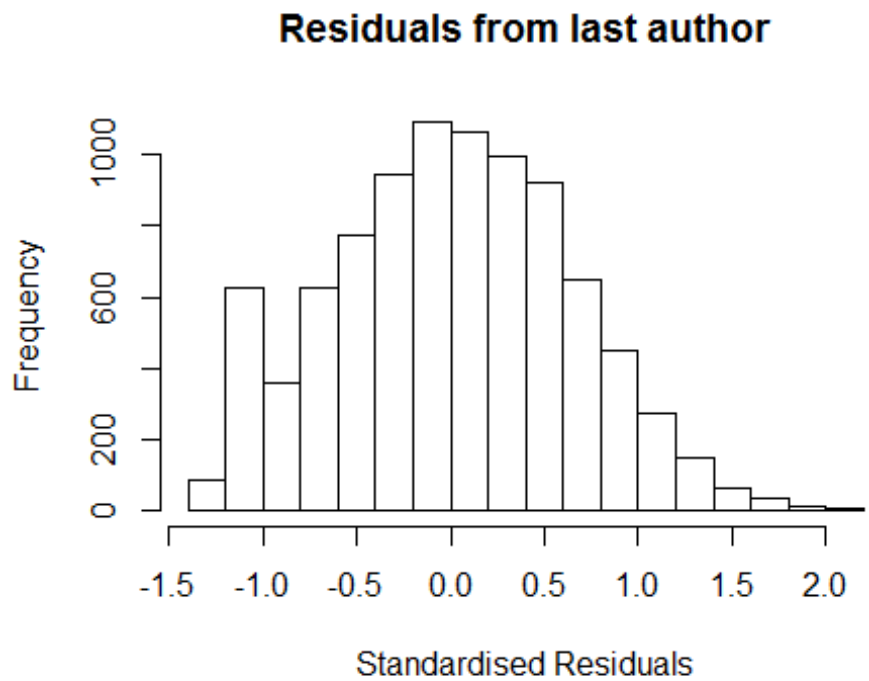

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##    Min     1Q Median     3Q    Max
## -1.247 -0.448  0.010  0.450  2.074
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)          1.1274     0.0284   39.70   <2e-16 ***
## FirstAuthorFemale1  -0.0130     0.0194   -0.67   0.5044
## LastAuthorFemale1   -0.0332     0.0193   -1.72   0.0852 .
## Year1997            -0.0302     0.0396   -0.76   0.4459
## Year1998            -0.0128     0.0408   -0.31   0.7539
## Year1999            -0.0489     0.0420   -1.17   0.2438
## Year2000            -0.0705     0.0403   -1.75   0.0800 .
## Year2001             0.0121     0.0413    0.29   0.7690
## Year2002             0.0321     0.0395    0.81   0.4157
## Year2003             0.0779     0.0406    1.92   0.0548 .
## Year2004             0.1199     0.0409    2.93   0.0034 **
## Year2005             0.0665     0.0422    1.58   0.1146
```

```
## Year2006             0.0593     0.0398    1.49   0.1366
## Year2007             0.1070     0.0389    2.75   0.0059 **
## Year2008             0.0611     0.0409    1.49   0.1353
## Year2009             0.0597     0.0397    1.50   0.1328
## Year2010             0.0923     0.0394    2.34   0.0193 *
## Year2011             0.0537     0.0378    1.42   0.1547
## Year2012             0.0327     0.0407    0.80   0.4222
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.662
## Multiple R-squared:  0.00636,    Adjusted R-squared:  0.00439
## Convergence in 10 IRWLS iterations
##
## Robustness weights:
##  796 weights are ~= 1. The remaining 8314 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.305   0.872   0.950   0.913   0.985   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.10e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.009  1           1.004
## Year              1.009 16           1.000
```




```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.2442 -0.4472  0.0112  0.4493  2.0785
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)           1.1247     0.0284   39.67   <2e-16 ***
## FirstAuthorFemale1   -0.0278     0.0179   -1.55   0.1213
## Year1997             -0.0309     0.0396   -0.78   0.4362
## Year1998             -0.0133     0.0408   -0.33   0.7442
## Year1999             -0.0488     0.0420   -1.16   0.2452
## Year2000             -0.0712     0.0403   -1.77   0.0771 .
## Year2001              0.0119     0.0413    0.29   0.7731
## Year2002              0.0321     0.0395    0.81   0.4157
## Year2003              0.0780     0.0406    1.92   0.0548 .
## Year2004              0.1195     0.0410    2.92   0.0035 **
## Year2005              0.0665     0.0421    1.58   0.1142
## Year2006              0.0578     0.0398    1.45   0.1461
```

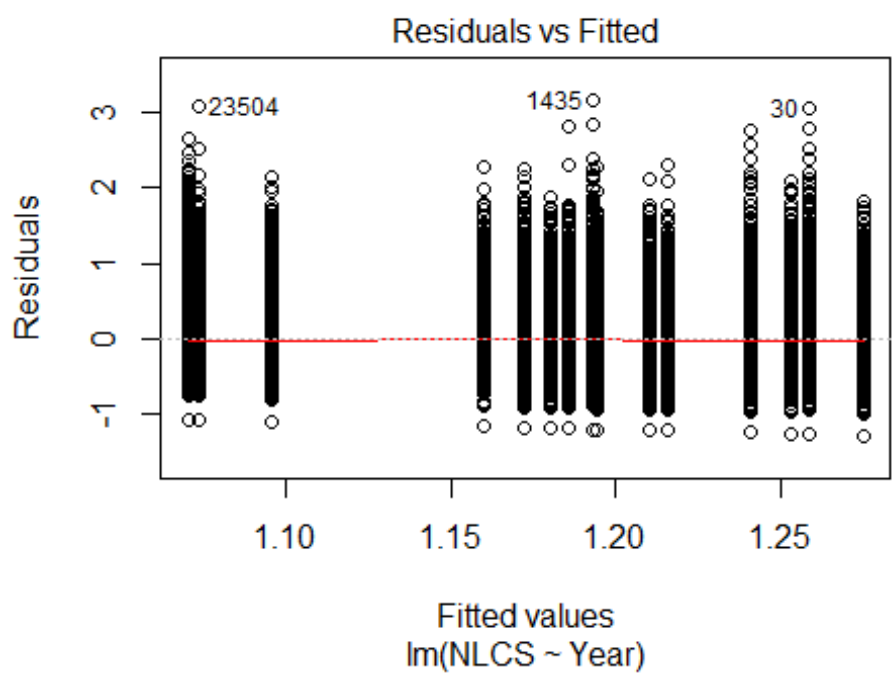
```
## Year2007             0.1054     0.0389    2.71   0.0067 **
## Year2008             0.0592     0.0409    1.45   0.1476
## Year2009             0.0584     0.0397    1.47   0.1419
## Year2010             0.0900     0.0394    2.29   0.0223 *
## Year2011             0.0521     0.0378    1.38   0.1674
## Year2012             0.0308     0.0407    0.76   0.4495
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.662
## Multiple R-squared:  0.00604,    Adjusted R-squared:  0.00418
## Convergence in 10 IRWLS iterations
##
## Robustness weights:
##  810 weights are ~= 1. The remaining 8300 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.303   0.872   0.949   0.913   0.984   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.10e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##                   GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.012  1           1.006
## Year             1.012 16           1.000
```



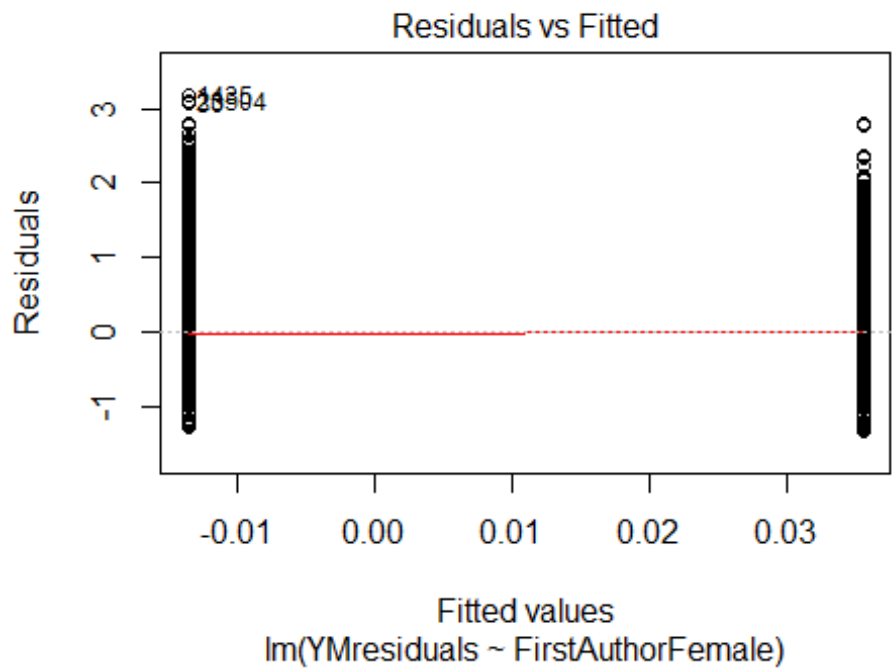

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.24590 -0.44615  0.00991  0.45044  2.07584
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.1264     0.0283   39.75   <2e-16 ***
## LastAuthorFemale1  -0.0390     0.0178   -2.19   0.0284 *
## Year1997           -0.0304     0.0396   -0.77   0.4424
## Year1998           -0.0130     0.0408   -0.32   0.7504
## Year1999           -0.0495     0.0420   -1.18   0.2381
## Year2000           -0.0704     0.0403   -1.75   0.0803 .
## Year2001            0.0115     0.0413    0.28   0.7809
## Year2002            0.0318     0.0395    0.80   0.4212
## Year2003            0.0775     0.0406    1.91   0.0559 .
## Year2004            0.1195     0.0409    2.92   0.0035 **
## Year2005            0.0660     0.0421    1.57   0.1172
## Year2006            0.0587     0.0398    1.47   0.1405
```

```
## Year2007            0.1066     0.0389    2.74   0.0061 **
## Year2008            0.0606     0.0409    1.48   0.1385
## Year2009            0.0587     0.0397    1.48   0.1390
## Year2010            0.0916     0.0394    2.32   0.0201 *
## Year2011            0.0531     0.0378    1.41   0.1597
## Year2012            0.0317     0.0407    0.78   0.4354
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.662
## Multiple R-squared:  0.0063, Adjusted R-squared:  0.00445
## Convergence in 10 IRWLS iterations
##
## Robustness weights:
##  811 weights are ~= 1. The remaining 8299 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.304   0.872   0.950   0.913   0.984   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.10e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Sample size for the above analysis:  9110"
## [1] ""
## [1] ""
## [1] "####################################"
## [1] "Analysis of AJSC 1403"
## [1] "####################################"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  824  922  898  857 1028 1057 1010  894 1046 1169 1263 1375 1632 2071 2111
## 2011 2012
## 1934 1826
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  657  743  703  671  838  807  819  729  847  975 1041 1115 1303 1650 1702
## 2011 2012
```

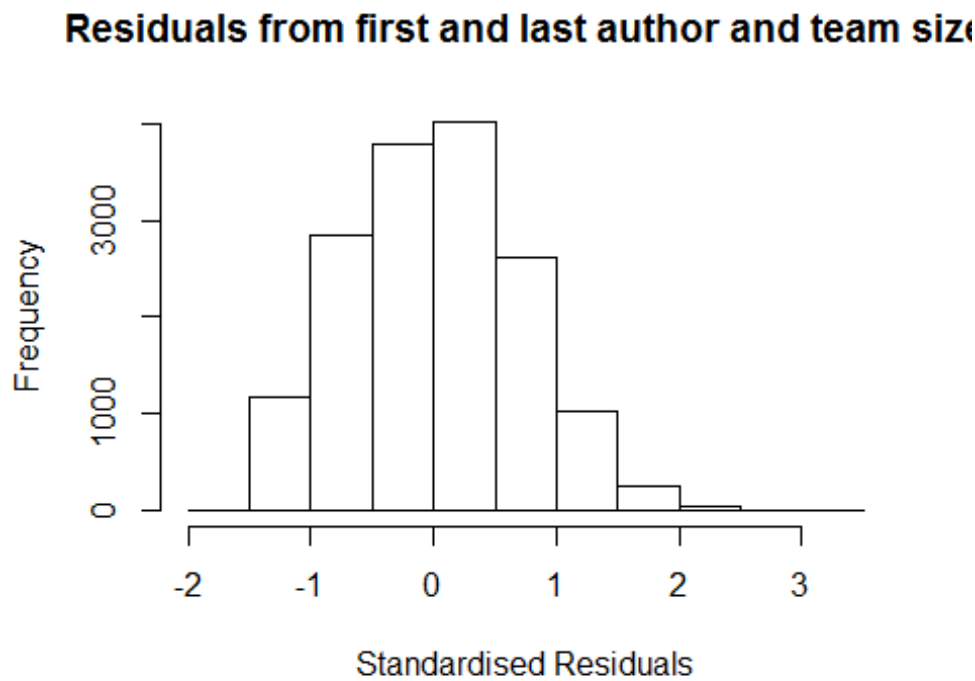
```
## 1521 1447
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  612  677  641  626  783  725  757  672  757  869  919  990 1156 1465 1508
## 2011 2012
## 1336 1273
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
##  Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 120, df = 16, p-value <2e-16
```




```
##
##  Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 0.68, df = 1, p-value = 0.4
```



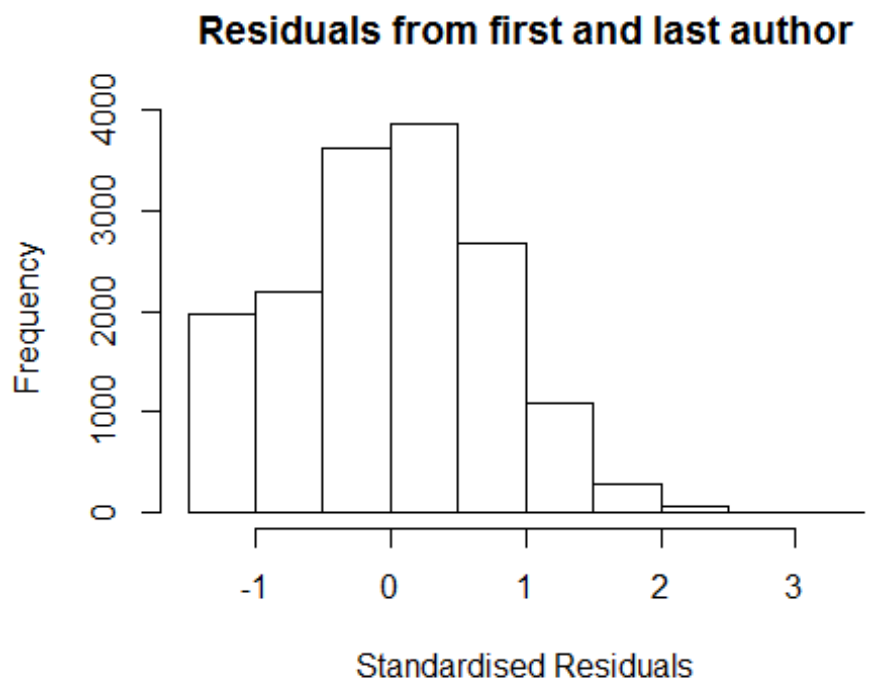

```
## [1] "Female first author team size 2018 geometric mean: 2.0283218044956"
## [1] "Male first author team size 2018 geometric mean: 1.86848755615926"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 150000, p-value = 0.06
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 1.91349366642321"
## [1] "Male last author team size 2018 geometric mean: 1.92992272993353"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 130000, p-value = 0.5
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##                     GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.364  1           1.168
## LastAuthorFemale  1.358  1           1.165
## UniqueAuthors     1.042  4           1.005
## Year              1.039 16           1.001
```




```
## [1] "List of  10 outliers with residuals above  2.5"
##          ScopusId  NLCS Year OneField Fields residuals
## 19     0041751098 4.048 1996     1403      2     2.988
## 30     2942649403 4.310 1996     1403      2     3.250
## 1435   0342775775 4.354 1997     1403      2     3.009
## 2303   0003123930 3.991 1998     1403      1     2.535
## 12680 36248939699 3.481 2007     1403      2     2.531
## 14084 34047207566 3.996 2007     1400      4     2.711
## 22743 79955593175 3.584 2011     1400      4     2.771
## 23504 84906331429 4.168 2011     1400      5     3.045
## 24086 84866640011 3.733 2012     1400      5     2.633
## 25107 84857455219 3.311 2012     1403      2     2.522
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
UniqueAuthors +
##     Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##     k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.5082 -0.5097  0.0104  0.5009  3.2497
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)           1.0603     0.0353   30.02  < 2e-16 ***
```

```
## FirstAuthorFemale1   0.0488     0.0154    3.16   0.0016 **
## LastAuthorFemale1   -0.0255     0.0154   -1.65   0.0989 .
## UniqueAuthors2       0.3107     0.0140   22.20  < 2e-16 ***
## UniqueAuthors3       0.4194     0.0167   25.16  < 2e-16 ***
## UniqueAuthors4       0.4479     0.0284   15.77  < 2e-16 ***
## UniqueAuthors5       0.3174     0.0418    7.59  3.4e-14 ***
## Year1997            -0.1088     0.0476   -2.29   0.0223 *
## Year1998            -0.0241     0.0462   -0.52   0.6028
## Year1999            -0.0441     0.0453   -0.98   0.3295
## Year2000            -0.0955     0.0430   -2.22   0.0262 *
## Year2001            -0.0409     0.0439   -0.93   0.3514
## Year2002            -0.0033     0.0432   -0.08   0.9391
## Year2003            -0.0829     0.0440   -1.88   0.0597 .
## Year2004            -0.0877     0.0431   -2.03   0.0421 *
## Year2005            -0.1058     0.0415   -2.55   0.0107 *
## Year2006            -0.1211     0.0411   -2.95   0.0032 **
## Year2007            -0.1098     0.0412   -2.67   0.0076 **
## Year2008            -0.0909     0.0403   -2.26   0.0240 *
## Year2009            -0.2046     0.0396   -5.16  2.4e-07 ***
## Year2010            -0.2496     0.0401   -6.23  4.8e-10 ***
## Year2011            -0.2478     0.0406   -6.11  1.0e-09 ***
## Year2012            -0.2710     0.0414   -6.54  6.3e-11 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.752
## Multiple R-squared:  0.0678, Adjusted R-squared:  0.0665
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  1296 weights are ~= 1. The remaining 14470 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0223  0.8780  0.9500  0.9170  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          6.34e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
```

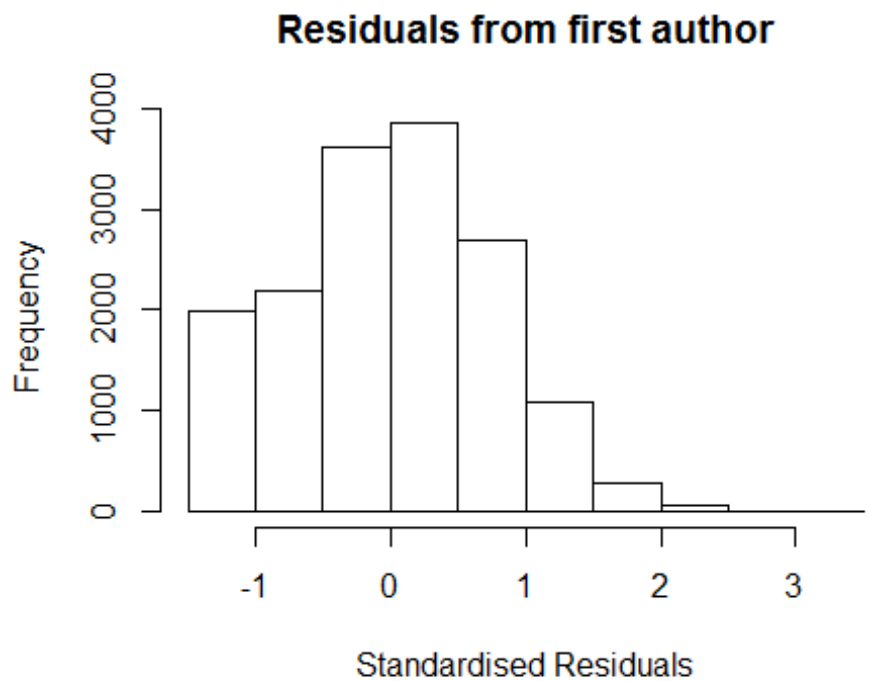
```
factors"
##                     GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.336  1           1.156
## LastAuthorFemale  1.331  1           1.154
## Year              1.011 16           1.000
```




```
## [1] "List of  10 outliers with residuals above  2.5"
##          ScopusId  NLCS Year OneField Fields residuals
## 19     0041751098 4.048 1996     1403      2     2.841
## 30     2942649403 4.310 1996     1403      2     3.103
## 557    0030305327 3.789 1996     1400      4     2.599
## 1435   0342775775 4.354 1997     1403      2     3.257
## 2303   0003123930 3.991 1998     1403      1     2.790
## 14084 34047207566 3.996 2007     1400      4     2.803
## 22743 79955593175 3.584 2011     1400      4     2.559
## 23504 84906331429 4.168 2011     1400      5     3.143
## 24086 84866640011 3.733 2012     1400      5     2.719
## 25051 84860589584 3.526 2012     1403      3     2.529
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.3105 -0.5416  0.0103  0.5261  3.2572
```

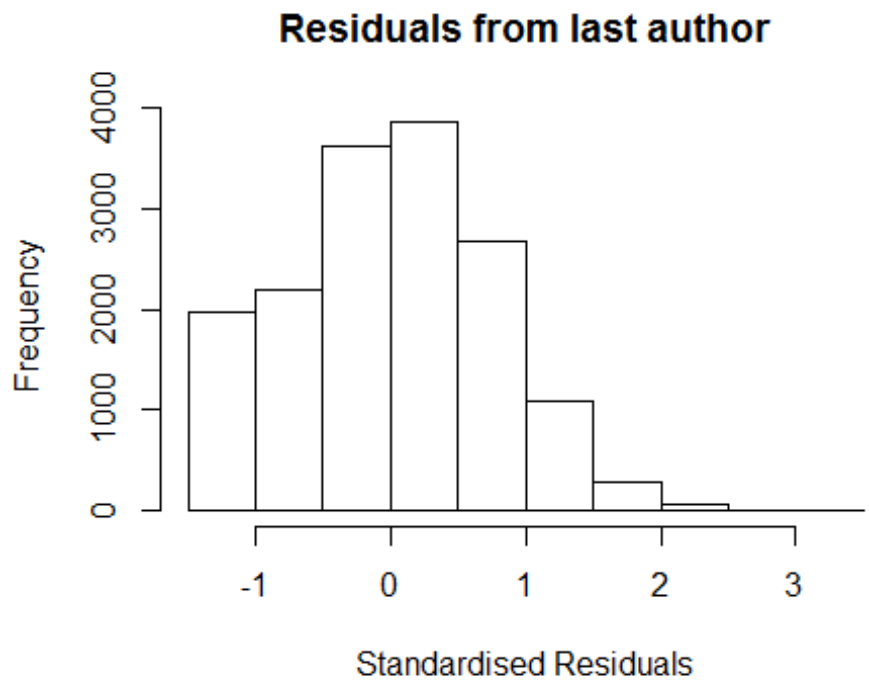
```
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.20657    0.03621   33.32  < 2e-16 ***
## FirstAuthorFemale1  0.06879    0.01580    4.36  1.3e-05 ***
## LastAuthorFemale1  -0.01680    0.01580   -1.06  0.28769
## Year1997           -0.09301    0.04846   -1.92  0.05496 .
## Year1998           -0.00562    0.04797   -0.12  0.90669
## Year1999           -0.02731    0.04704   -0.58  0.56157
## Year2000           -0.07445    0.04464   -1.67  0.09534 .
## Year2001           -0.00550    0.04512   -0.12  0.90301
## Year2002            0.03509    0.04483    0.78  0.43376
## Year2003           -0.05769    0.04546   -1.27  0.20448
## Year2004           -0.03992    0.04507   -0.89  0.37578
## Year2005           -0.06606    0.04306   -1.53  0.12498
## Year2006           -0.07201    0.04294   -1.68  0.09355 .
## Year2007           -0.06516    0.04288   -1.52  0.12861
## Year2008           -0.03246    0.04182   -0.78  0.43770
## Year2009           -0.15866    0.04115   -3.86  0.00012 ***
## Year2010           -0.18543    0.04146   -4.47  7.8e-06 ***
## Year2011           -0.18184    0.04195   -4.33  1.5e-05 ***
## Year2012           -0.19307    0.04286   -4.50  6.7e-06 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.784
## Multiple R-squared:  0.0104, Adjusted R-squared:  0.00927
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  1273 weights are ~= 1. The remaining 14493 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0456  0.8690  0.9500  0.9190  0.9860  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          6.34e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
```

```
##                     GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.009  1           1.005
## Year              1.009 16           1.000
```




```
## [1] "List of  10 outliers with residuals above  2.5"
##          ScopusId  NLCS Year OneField Fields residuals
## 19     0041751098 4.048 1996     1403      2     2.841
## 30     2942649403 4.310 1996     1403      2     3.103
## 557    0030305327 3.789 1996     1400      4     2.599
## 1435   0342775775 4.354 1997     1403      2     3.257
## 2303   0003123930 3.991 1998     1403      1     2.790
## 14084 34047207566 3.996 2007     1400      4     2.803
## 22743 79955593175 3.584 2011     1400      4     2.559
## 23504 84906331429 4.168 2011     1400      5     3.143
## 24086 84866640011 3.733 2012     1400      5     2.719
## 25051 84860589584 3.526 2012     1403      3     2.529
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.2998 -0.5411  0.0108  0.5262  3.2426
##
## Coefficients:
```

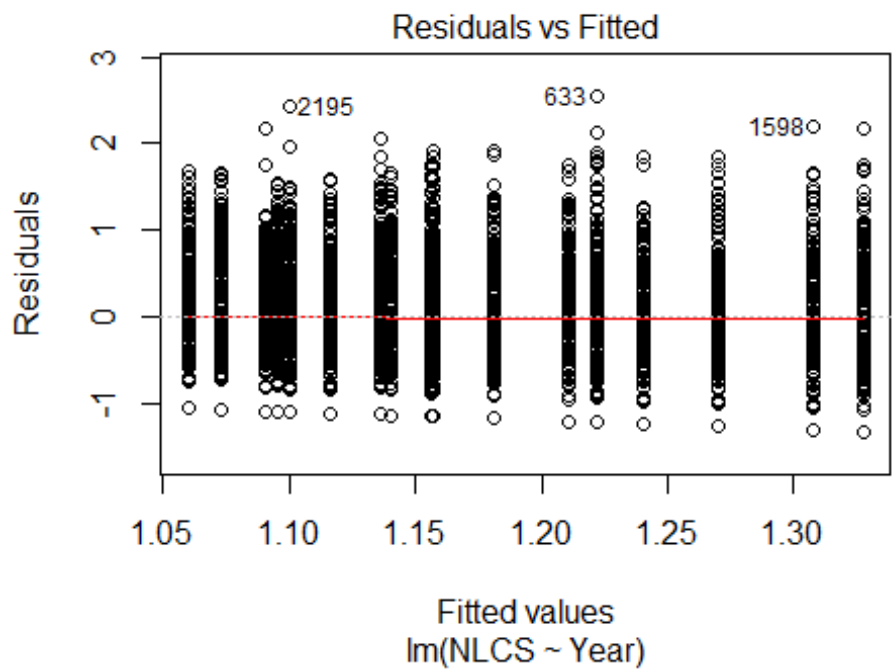
```
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)          1.20447    0.03613   33.34  < 2e-16 ***
## FirstAuthorFemale1  0.06033    0.01375    4.39  1.1e-05 ***
## Year1997           -0.09312    0.04845   -1.92  0.05465 .
## Year1998           -0.00556    0.04796   -0.12  0.90770
## Year1999           -0.02748    0.04707   -0.58  0.55934
## Year2000           -0.07429    0.04464   -1.66  0.09607 .
## Year2001           -0.00563    0.04511   -0.12  0.90074
## Year2002            0.03498    0.04484    0.78  0.43525
## Year2003           -0.05799    0.04545   -1.28  0.20199
## Year2004           -0.03971    0.04506   -0.88  0.37816
## Year2005           -0.06585    0.04305   -1.53  0.12616
## Year2006           -0.07224    0.04294   -1.68  0.09251 .
## Year2007           -0.06526    0.04288   -1.52  0.12800
## Year2008           -0.03276    0.04182   -0.78  0.43347
## Year2009           -0.15865    0.04115   -3.86  0.00012 ***
## Year2010           -0.18567    0.04147   -4.48  7.6e-06 ***
## Year2011           -0.18200    0.04195   -4.34  1.4e-05 ***
## Year2012           -0.19348    0.04285   -4.52  6.4e-06 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.784
## Multiple R-squared:  0.0103, Adjusted R-squared:  0.00926
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  1258 weights are ~= 1. The remaining 14508 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0486  0.8690  0.9500  0.9190  0.9860  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          6.34e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##                   GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.005  1           1.003
## Year             1.005 16           1.000
```




```
## [1] "List of  10 outliers with residuals above  2.5"
##          ScopusId  NLCS Year OneField Fields residuals
## 19     0041751098 4.048 1996     1403      2     2.841
## 30     2942649403 4.310 1996     1403      2     3.103
## 557    0030305327 3.789 1996     1400      4     2.599
## 1435   0342775775 4.354 1997     1403      2     3.257
## 2303   0003123930 3.991 1998     1403      1     2.790
## 14084 34047207566 3.996 2007     1400      4     2.803
## 22743 79955593175 3.584 2011     1400      4     2.559
## 23504 84906331429 4.168 2011     1400      5     3.143
## 24086 84866640011 3.733 2012     1400      5     2.719
## 25051 84860589584 3.526 2012     1403      3     2.529
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.2673 -0.5364  0.0106  0.5255  3.2122
##
## Coefficients:
##                   Estimate Std. Error t value Pr(>|t|)
## (Intercept)        1.21607    0.03615   33.64  < 2e-16 ***
## LastAuthorFemale1  0.01825    0.01368    1.33   0.1822
## Year1997          -0.09253    0.04850   -1.91   0.0564 .
```

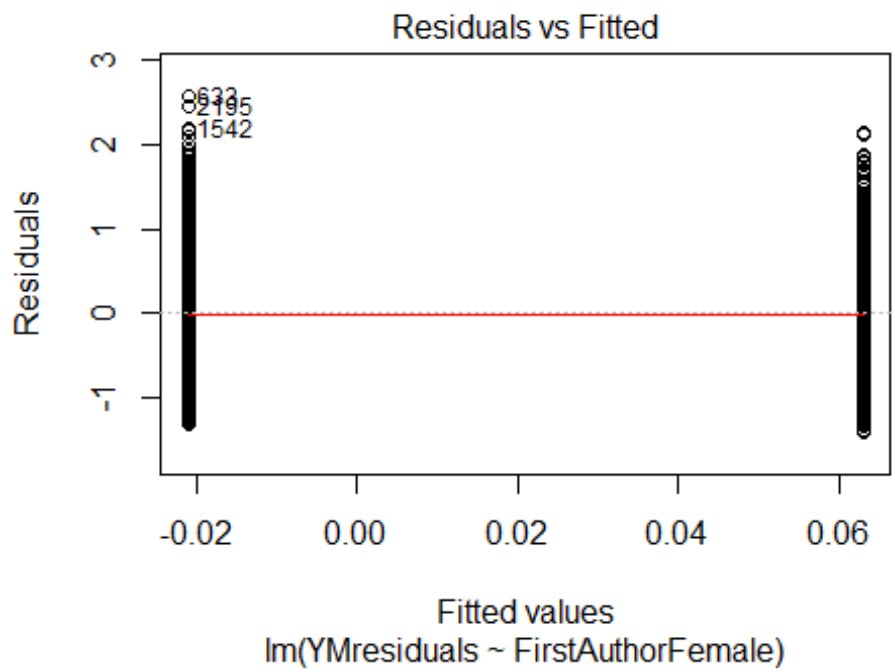
```
## Year1998          -0.00641    0.04796   -0.13   0.8936
## Year1999          -0.02671    0.04705   -0.57   0.5702
## Year2000          -0.07264    0.04463   -1.63   0.1036
## Year2001          -0.00504    0.04514   -0.11   0.9112
## Year2002           0.03297    0.04485    0.74   0.4622
## Year2003          -0.05757    0.04544   -1.27   0.2052
## Year2004          -0.04188    0.04508   -0.93   0.3529
## Year2005          -0.06582    0.04306   -1.53   0.1264
## Year2006          -0.07320    0.04296   -1.70   0.0884 .
## Year2007          -0.06477    0.04289   -1.51   0.1310
## Year2008          -0.03200    0.04182   -0.77   0.4442
## Year2009          -0.15997    0.04116   -3.89   0.0001 ***
## Year2010          -0.18327    0.04148   -4.42  1.0e-05 ***
## Year2011          -0.18213    0.04195   -4.34  1.4e-05 ***
## Year2012          -0.19268    0.04287   -4.49  7.0e-06 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.785
## Multiple R-squared:  0.00915,    Adjusted R-squared:  0.00808
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  1273 weights are ~= 1. The remaining 14493 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0564  0.8690  0.9500  0.9190  0.9860  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          6.34e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Sample size for the above analysis:  15766"
## [1] ""
## [1] ""
## [1] "####################################"
## [1] "Analysis of AJSC 1404"
## [1] "####################################"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
```

```
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  282  258  286  280  299  254  253  283  244  324  434  390  445  443  407
## 2011 2012
##  418  366
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  220  173  198  189  203  159  176  223  196  234  297  275  290  296  262
## 2011 2012
##  282  246
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  200  151  182  161  180  140  154  194  159  204  247  227  236  255  222
## 2011 2012
##  231  194
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
##  Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 69, df = 16, p-value = 2e-08
```



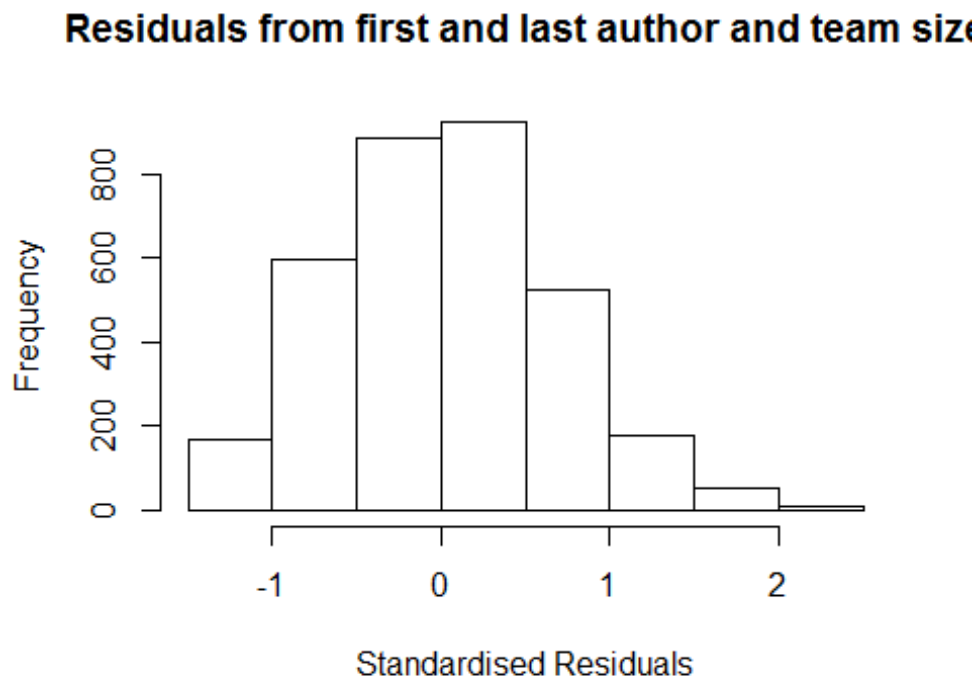

```
##
##  Bartlett test of homogeneity of variances
##
```

```
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 4.3, df = 1, p-value = 0.04
```



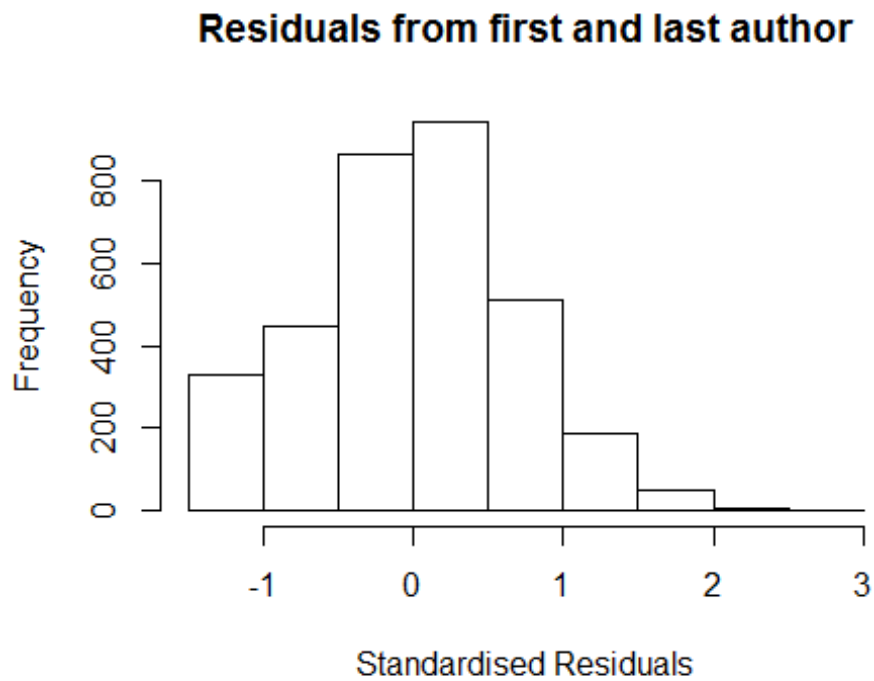

```
## [1] "Female first author team size 2018 geometric mean: 2.15413634178072"
## [1] "Male first author team size 2018 geometric mean: 2.13309749068613"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 3300, p-value = 0.9
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 2.04265848368542"
## [1] "Male last author team size 2018 geometric mean: 2.16801877057392"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 2700, p-value = 0.5
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.144  1           1.070
## LastAuthorFemale  1.126  1           1.061
```

```
## UniqueAuthors     1.100  4           1.012
## Year              1.123 16           1.004
```



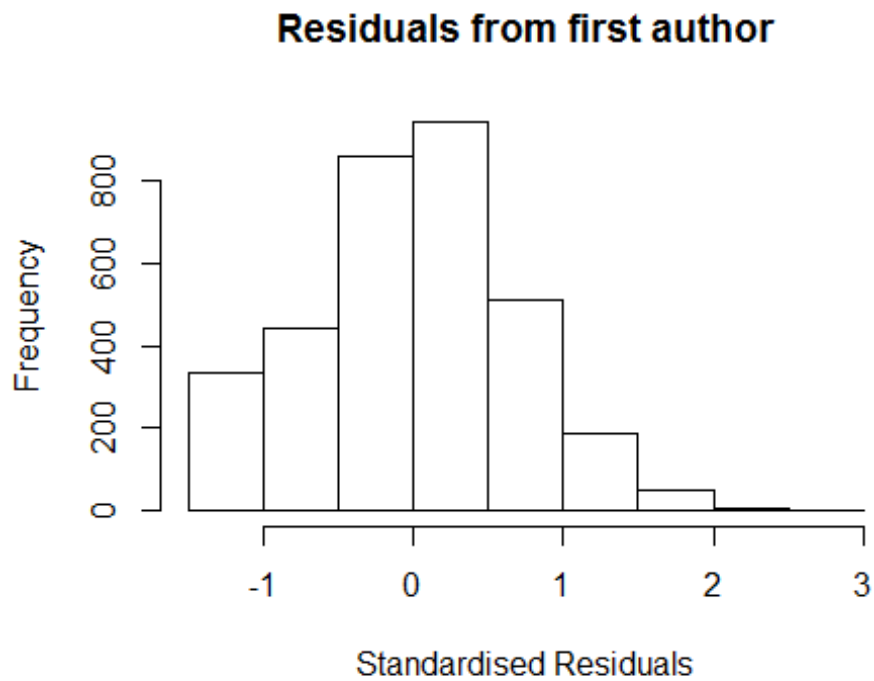

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
UniqueAuthors +
##     Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##     k.max = 1000))
##  \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.41155 -0.46119  0.00758  0.44089  2.42013
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         0.94034    0.06323   14.87  < 2e-16 ***
## FirstAuthorFemale1  0.07795    0.02947    2.64   0.0082 **
## LastAuthorFemale1  -0.00585    0.03107   -0.19   0.8507
## UniqueAuthors2      0.27397    0.02946    9.30  < 2e-16 ***
## UniqueAuthors3      0.36619    0.03191   11.48  < 2e-16 ***
## UniqueAuthors4      0.43230    0.04666    9.26  < 2e-16 ***
## UniqueAuthors5      0.44717    0.07355    6.08  1.3e-09 ***
```

```
## Year1997           -0.05090    0.08839   -0.58   0.5648
## Year1998            0.00796    0.08468    0.09   0.9251
## Year1999            0.07020    0.08019    0.88   0.3814
## Year2000            0.12754    0.07765    1.64   0.1006
## Year2001            0.11930    0.08610    1.39   0.1660
## Year2002           -0.00175    0.08436   -0.02   0.9835
## Year2003           -0.09873    0.07961   -1.24   0.2150
## Year2004            0.06611    0.07666    0.86   0.3885
## Year2005           -0.04336    0.07344   -0.59   0.5550
## Year2006           -0.08542    0.07120   -1.20   0.2303
## Year2007           -0.14250    0.07020   -2.03   0.0424 *
## Year2008           -0.08808    0.07112   -1.24   0.2156
## Year2009           -0.10728    0.07186   -1.49   0.1356
## Year2010           -0.17542    0.07429   -2.36   0.0183 *
## Year2011           -0.06874    0.07828   -0.88   0.3799
## Year2012           -0.19304    0.07749   -2.49   0.0128 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.667
## Multiple R-squared:  0.0739, Adjusted R-squared:  0.0677
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
##  326 weights are ~= 1. The remaining 3011 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.160   0.863   0.948   0.909   0.985   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          3.00e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.127  1           1.062
## LastAuthorFemale  1.108  1           1.053
## Year              1.032 16           1.001
```



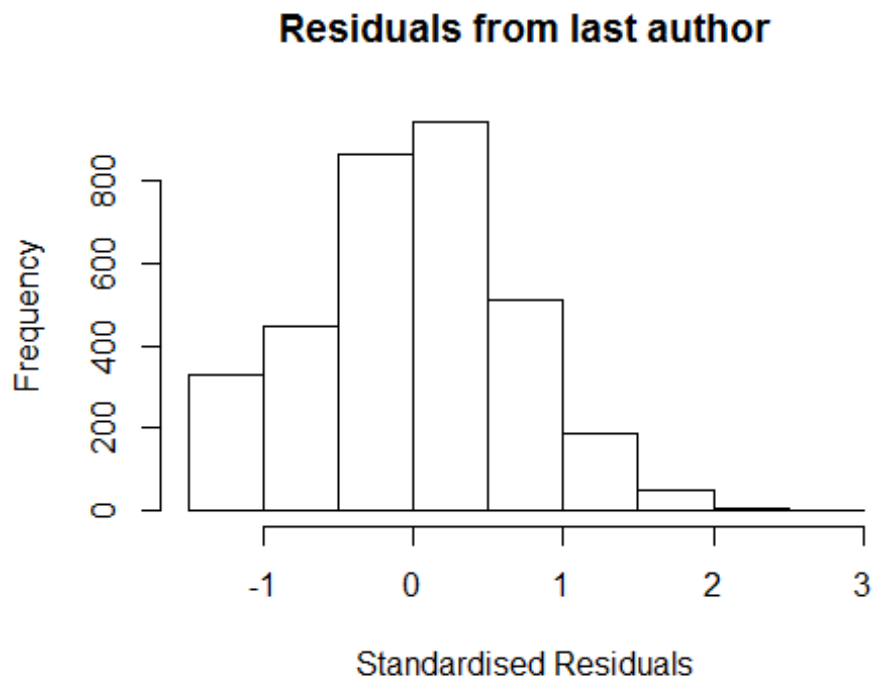

```
## [1] "List of  1 outliers with residuals above  2.5"
##        ScopusId  NLCS Year OneField Fields residuals
## 2195 1542382496 3.527 2003     1404      4     2.516
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.3713 -0.4635  0.0128  0.4509  2.5158
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)          1.13528    0.05908   19.22   <2e-16 ***
## FirstAuthorFemale1  0.09840    0.03012    3.27   0.0011 **
## LastAuthorFemale1  -0.01296    0.03169   -0.41   0.6827
## Year1997           -0.05589    0.09046   -0.62   0.5368
## Year1998           -0.01573    0.08572   -0.18   0.8544
## Year1999            0.06917    0.07899    0.88   0.3812
## Year2000            0.14889    0.07839    1.90   0.0576 .
## Year2001            0.13763    0.08529    1.61   0.1067
## Year2002            0.02333    0.08626    0.27   0.7868
## Year2003           -0.12408    0.08181   -1.52   0.1295
## Year2004            0.09632    0.07624    1.26   0.2065
## Year2005           -0.00837    0.07287   -0.11   0.9085
```

```
## Year2006           -0.04311    0.07035   -0.61   0.5401
## Year2007           -0.10759    0.06953   -1.55   0.1219
## Year2008           -0.07500    0.07079   -1.06   0.2894
## Year2009           -0.09058    0.07180   -1.26   0.2072
## Year2010           -0.12621    0.07409   -1.70   0.0886 .
## Year2011           -0.05016    0.07801   -0.64   0.5203
## Year2012           -0.16287    0.07770   -2.10   0.0362 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.677
## Multiple R-squared:  0.0194, Adjusted R-squared:  0.0141
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  282 weights are ~= 1. The remaining 3055 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.138   0.866   0.951   0.908   0.986   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          3.00e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.026  1           1.013
## Year              1.026 16           1.001
```




```
## [1] "List of  1 outliers with residuals above  2.5"
##        ScopusId  NLCS Year OneField Fields residuals
## 2195 1542382496 3.527 2003     1404      4     2.516
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##    Min     1Q Median     3Q    Max
## -1.376 -0.462  0.014  0.448  2.518
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)          1.13380    0.05893   19.24   <2e-16 ***
## FirstAuthorFemale1  0.09369    0.02888    3.24   0.0012 **
## Year1997           -0.05608    0.09047   -0.62   0.5354
## Year1998           -0.01577    0.08575   -0.18   0.8541
## Year1999            0.06917    0.07900    0.88   0.3814
## Year2000            0.14869    0.07839    1.90   0.0579 .
## Year2001            0.13722    0.08523    1.61   0.1075
## Year2002            0.02290    0.08633    0.27   0.7909
## Year2003           -0.12476    0.08179   -1.53   0.1273
## Year2004            0.09608    0.07624    1.26   0.2077
## Year2005           -0.00872    0.07288   -0.12   0.9047
## Year2006           -0.04325    0.07032   -0.62   0.5386
```

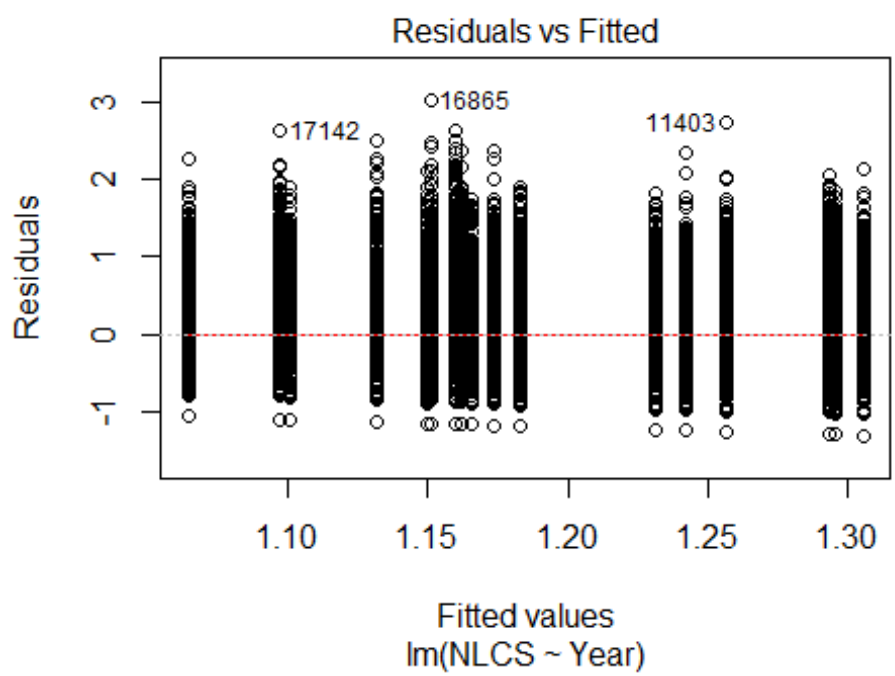
```
## Year2007           -0.10760    0.06953   -1.55   0.1218
## Year2008           -0.07554    0.07076   -1.07   0.2858
## Year2009           -0.09044    0.07178   -1.26   0.2078
## Year2010           -0.12680    0.07405   -1.71   0.0869 .
## Year2011           -0.05014    0.07800   -0.64   0.5204
## Year2012           -0.16329    0.07770   -2.10   0.0357 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.678
## Multiple R-squared:  0.0194, Adjusted R-squared:  0.0143
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  279 weights are ~= 1. The remaining 3058 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.138   0.866   0.951   0.909   0.986   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          3.00e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##                   GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.009  1           1.005
## Year             1.009 16           1.000
```



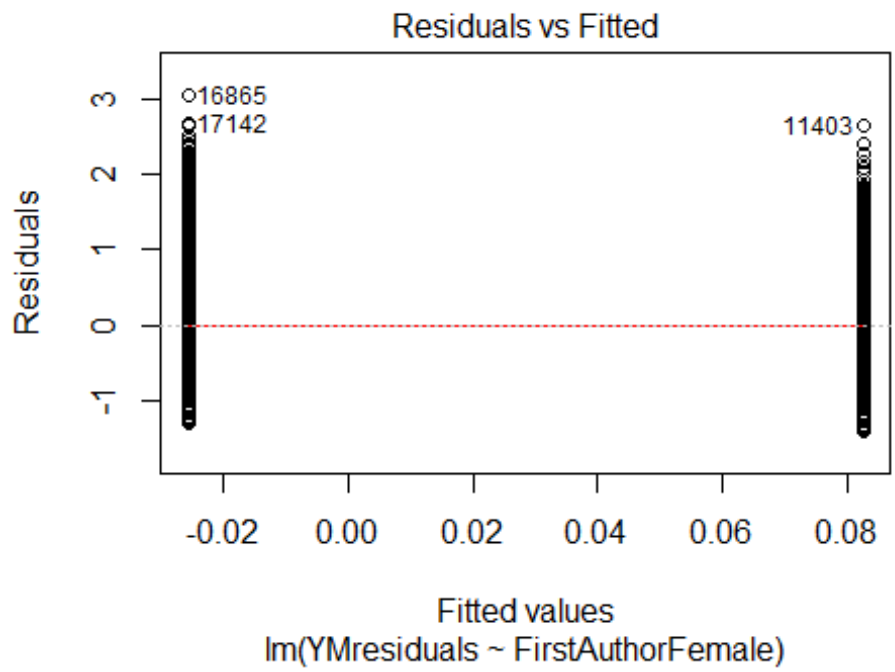

```
## [1] "List of  1 outliers with residuals above  2.5"
##        ScopusId  NLCS Year OneField Fields residuals
## 2195 1542382496 3.527 2003     1404      4     2.516
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.33073 -0.46260  0.00881  0.44991  2.50591
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.15060    0.05916   19.45   <2e-16 ***
## LastAuthorFemale1   0.02554    0.03012    0.85    0.397
## Year1997           -0.05944    0.09069   -0.66    0.512
## Year1998           -0.01873    0.08607   -0.22    0.828
## Year1999            0.06658    0.07945    0.84    0.402
## Year2000            0.15460    0.07887    1.96    0.050 .
## Year2001            0.13809    0.08510    1.62    0.105
## Year2002            0.02522    0.08653    0.29    0.771
## Year2003           -0.12950    0.08167   -1.59    0.113
## Year2004            0.09257    0.07685    1.20    0.228
## Year2005           -0.00444    0.07312   -0.06    0.952
## Year2006           -0.04082    0.07063   -0.58    0.563
```

```
## Year2007          -0.10776    0.06981   -1.54    0.123
## Year2008          -0.07666    0.07107   -1.08    0.281
## Year2009          -0.09031    0.07209   -1.25    0.210
## Year2010          -0.12169    0.07435   -1.64    0.102
## Year2011          -0.04611    0.07810   -0.59    0.555
## Year2012          -0.16002    0.07829   -2.04    0.041 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.679
## Multiple R-squared:  0.0161, Adjusted R-squared:  0.0111
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  295 weights are ~= 1. The remaining 3042 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.144   0.865   0.949   0.908   0.985   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          3.00e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Sample size for the above analysis:  3337"
## [1] ""
## [1] ""
## [1] "####################################"
## [1] "Analysis of AJSC 1405"
## [1] "####################################"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  840  764  816  808  850 1047  956  705  717  807  815  944 1019 1099 1091
## 2011 2012
## 1206  977
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  716  627  651  672  683  727  797  597  596  660  662  781  822  882  857
## 2011 2012
```

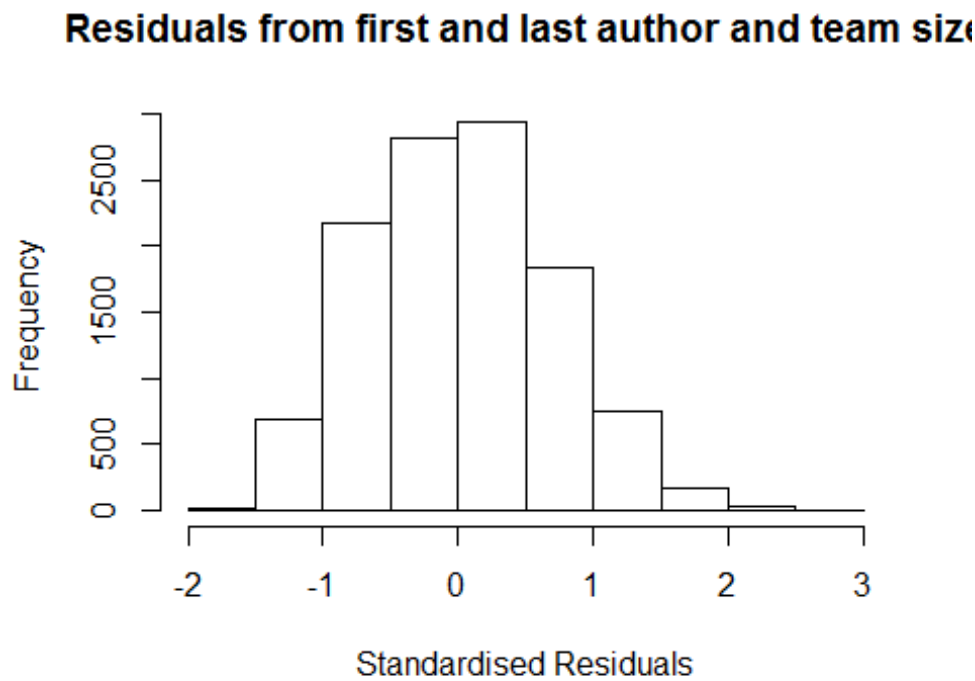
```
##  984  783
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  679  592  609  618  640  673  735  555  541  597  590  702  744  796  767
## 2011 2012
##  894  697
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
##  Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 69, df = 16, p-value = 1e-08
```




```
##
##  Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 3.8, df = 1, p-value = 0.05
```



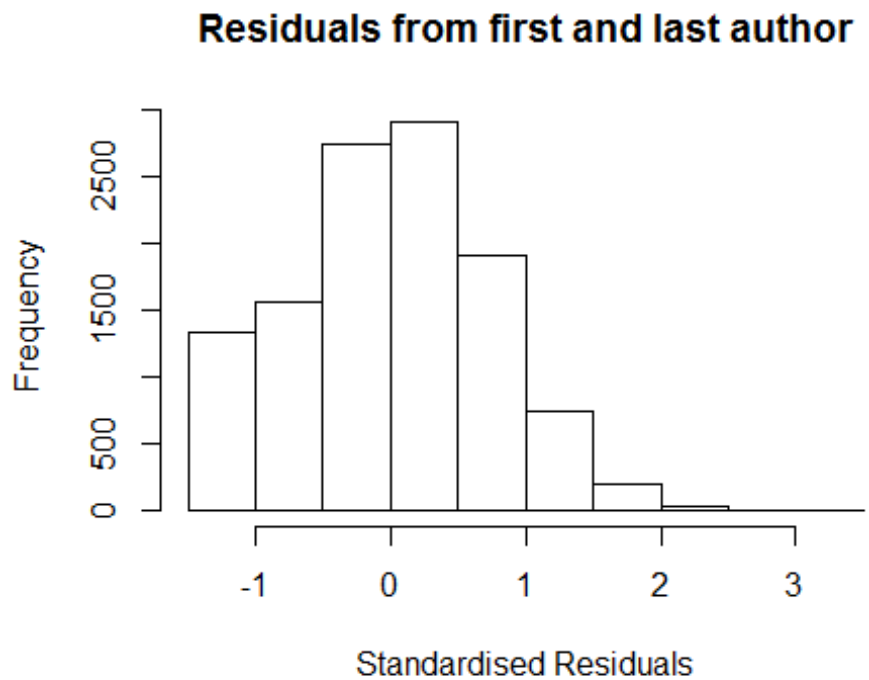

```
## [1] "Female first author team size 2018 geometric mean: 1.99512517476206"
## [1] "Male first author team size 2018 geometric mean: 1.84923247355099"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 56000, p-value = 0.09
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 1.89170407965048"
## [1] "Male last author team size 2018 geometric mean: 1.88799616991945"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 46000, p-value = 0.9
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.300  1           1.140
## LastAuthorFemale  1.293  1           1.137
## UniqueAuthors     1.058  4           1.007
## Year              1.059 16           1.002
```




```
## [1] "List of  6 outliers with residuals above  2.5"
##          ScopusId  NLCS Year OneField Fields residuals
## 629    0030505108 3.545 1996     1405      3     2.593
## 670    0030530698 3.505 1996     1400      3     2.553
## 11403 34047207566 3.996 2007     1400      4     2.579
## 16315 79955593175 3.584 2011     1400      4     2.698
## 16865 84906331429 4.168 2011     1400      5     2.986
## 17142 84866640011 3.733 2012     1400      5     2.631
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
UniqueAuthors +
##     Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##     k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.6140 -0.5019  0.0029  0.4876  2.9865
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         0.95216    0.03009   31.64  < 2e-16 ***
## FirstAuthorFemale1  0.07673    0.01792    4.28  1.9e-05 ***
## LastAuthorFemale1   0.04583    0.01823    2.51   0.0119 *
## UniqueAuthors2      0.29565    0.01597   18.51  < 2e-16 ***
## UniqueAuthors3      0.41513    0.01922   21.59  < 2e-16 ***
```

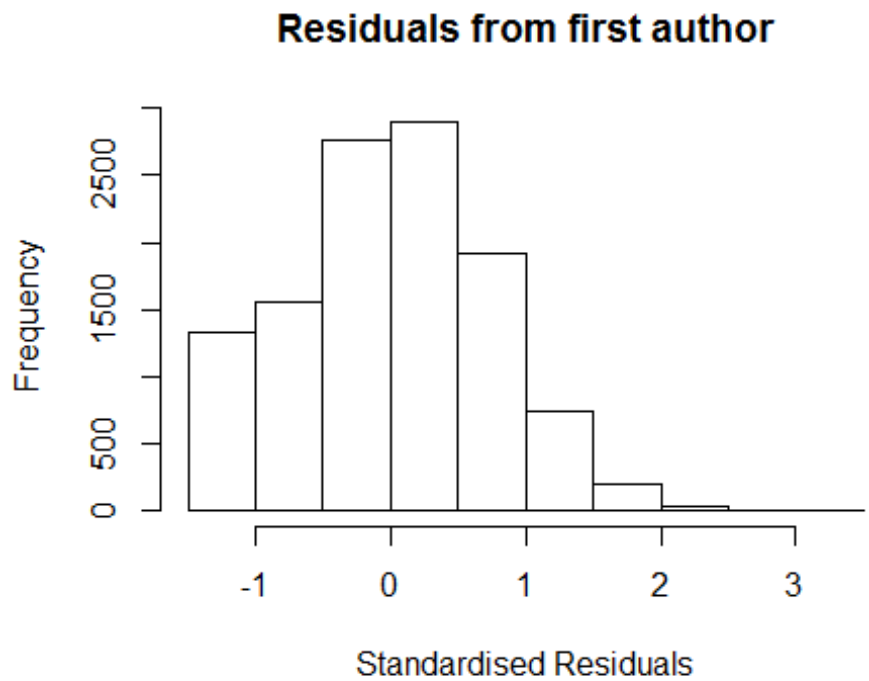
```
## UniqueAuthors4      0.43535    0.03279   13.28  < 2e-16 ***
## UniqueAuthors5      0.31994    0.04323    7.40  1.4e-13 ***
## Year1997            0.01794    0.04535    0.40   0.6924
## Year1998           -0.02998    0.04318   -0.69   0.4875
## Year1999           -0.04758    0.04065   -1.17   0.2418
## Year2000           -0.10485    0.04058   -2.58   0.0098 **
## Year2001            0.12416    0.04271    2.91   0.0037 **
## Year2002            0.00166    0.03991    0.04   0.9668
## Year2003           -0.01823    0.04124   -0.44   0.6585
## Year2004            0.11212    0.04217    2.66   0.0078 **
## Year2005            0.03309    0.04028    0.82   0.4114
## Year2006            0.10223    0.04023    2.54   0.0111 *
## Year2007            0.04626    0.04065    1.14   0.2552
## Year2008            0.02364    0.03890    0.61   0.5434
## Year2009           -0.05757    0.03824   -1.51   0.1322
## Year2010           -0.02835    0.03931   -0.72   0.4709
## Year2011           -0.06629    0.03848   -1.72   0.0850 .
## Year2012           -0.14586    0.04030   -3.62   0.0003 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.731
## Multiple R-squared:  0.0717, Adjusted R-squared:  0.0699
## Convergence in 10 IRWLS iterations
##
## Robustness weights:
##  966 weights are ~= 1. The remaining 10463 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0577  0.8680  0.9500  0.9150  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          8.75e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.283  1           1.133
```

```
## LastAuthorFemale  1.278  1           1.131
## Year              1.010 16           1.000
```



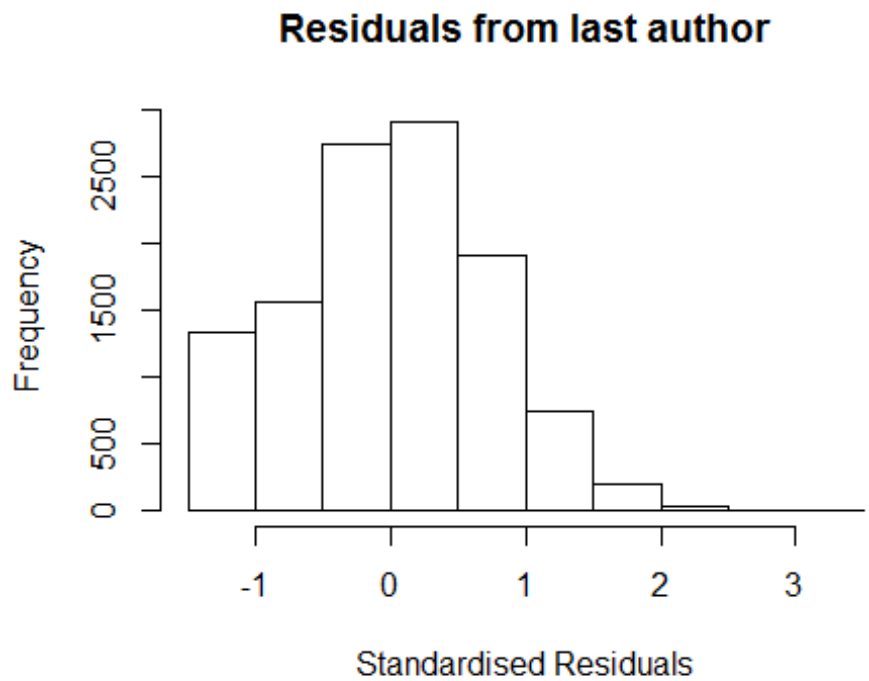

```
## [1] "List of  5 outliers with residuals above  2.5"
##          ScopusId  NLCS Year OneField Fields residuals
## 553    0030305327 3.789 1996     1400      4     2.649
## 2242   0000384657 3.626 1998     1403      2     2.524
## 11403 34047207566 3.996 2007     1400      4     2.660
## 16865 84906331429 4.168 2011     1400      5     3.082
## 17142 84866640011 3.733 2012     1400      5     2.719
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.4158 -0.5074  0.0141  0.5062  3.0815
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.085793   0.030435   35.68  < 2e-16 ***
## FirstAuthorFemale1  0.095563   0.018313    5.22  1.8e-07 ***
## LastAuthorFemale1   0.054376   0.018606    2.92  0.00348 **
## Year1997            0.017225   0.046639    0.37  0.71189
## Year1998           -0.038058   0.044793   -0.85  0.39555
```

```
## Year1999           -0.024412   0.041846   -0.58  0.55965
## Year2000           -0.083519   0.042300   -1.97  0.04835 *
## Year2001            0.148921   0.043960    3.39  0.00071 ***
## Year2002            0.039570   0.041338    0.96  0.33847
## Year2003            0.022539   0.042730    0.53  0.59788
## Year2004            0.172841   0.043477    3.98  7.1e-05 ***
## Year2005            0.090668   0.041523    2.18  0.02902 *
## Year2006            0.180071   0.041625    4.33  1.5e-05 ***
## Year2007            0.100415   0.041417    2.42  0.01535 *
## Year2008            0.090447   0.039787    2.27  0.02303 *
## Year2009            0.008240   0.039441    0.21  0.83452
## Year2010            0.036711   0.040566    0.90  0.36550
## Year2011            0.000688   0.039674    0.02  0.98617
## Year2012           -0.071472   0.041329   -1.73  0.08378 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.752
## Multiple R-squared:  0.0161, Adjusted R-squared:  0.0145
## Convergence in 10 IRWLS iterations
##
## Robustness weights:
##  974 weights are ~= 1. The remaining 10455 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0553  0.8680  0.9500  0.9150  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          8.75e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.006  1           1.003
## Year              1.006 16           1.000
```




```
## [1] "List of  5 outliers with residuals above  2.5"
##          ScopusId  NLCS Year OneField Fields residuals
## 553    0030305327 3.789 1996     1400      4     2.649
## 2242   0000384657 3.626 1998     1403      2     2.524
## 11403 34047207566 3.996 2007     1400      4     2.660
## 16865 84906331429 4.168 2011     1400      5     3.082
## 17142 84866640011 3.733 2012     1400      5     2.719
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.3948 -0.5060  0.0135  0.5050  3.0753
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.09101    0.03036   35.94  < 2e-16 ***
## FirstAuthorFemale1  0.12299    0.01619    7.60  3.3e-14 ***
## Year1997            0.01687    0.04668    0.36   0.7177
## Year1998           -0.03595    0.04476   -0.80   0.4219
## Year1999           -0.02347    0.04189   -0.56   0.5753
## Year2000           -0.08296    0.04232   -1.96   0.0500 *
## Year2001            0.14911    0.04397    3.39   0.0007 ***
## Year2002            0.04097    0.04132    0.99   0.3214
```

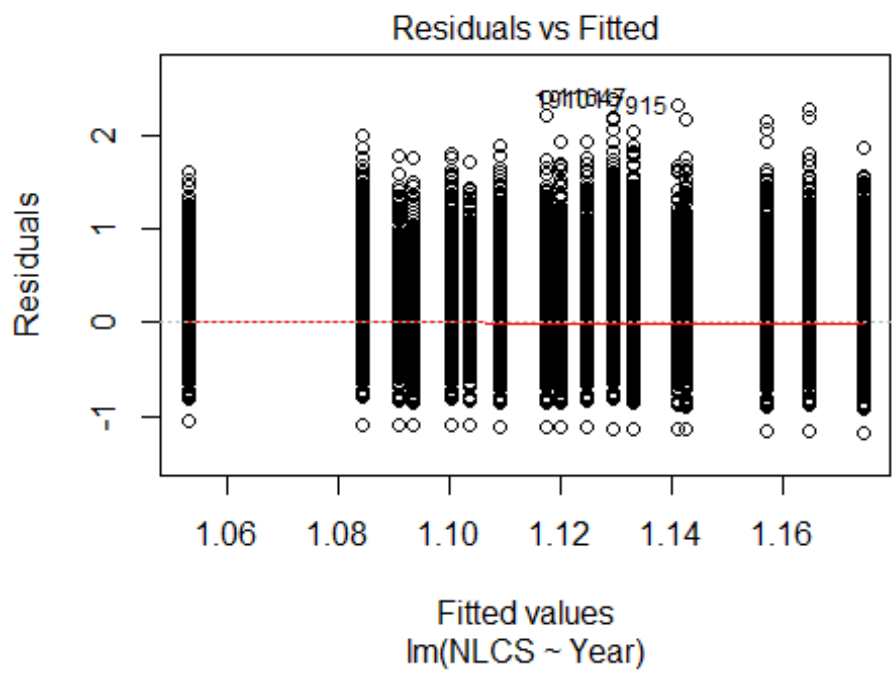
```
## Year2003            0.02235    0.04278    0.52   0.6013
## Year2004            0.17385    0.04349    4.00  6.4e-05 ***
## Year2005            0.09211    0.04148    2.22   0.0264 *
## Year2006            0.18082    0.04162    4.34  1.4e-05 ***
## Year2007            0.10319    0.04145    2.49   0.0128 *
## Year2008            0.09120    0.03979    2.29   0.0219 *
## Year2009            0.00898    0.03946    0.23   0.8200
## Year2010            0.03744    0.04059    0.92   0.3563
## Year2011            0.00165    0.03970    0.04   0.9668
## Year2012           -0.07025    0.04132   -1.70   0.0892 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.752
## Multiple R-squared:  0.0153, Adjusted R-squared:  0.0138
## Convergence in 10 IRWLS iterations
##
## Robustness weights:
##  985 weights are ~= 1. The remaining 10444 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0567  0.8680  0.9500  0.9150  0.9860  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          8.75e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                    psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##                   GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.003  1           1.002
## Year             1.003 16           1.000
```



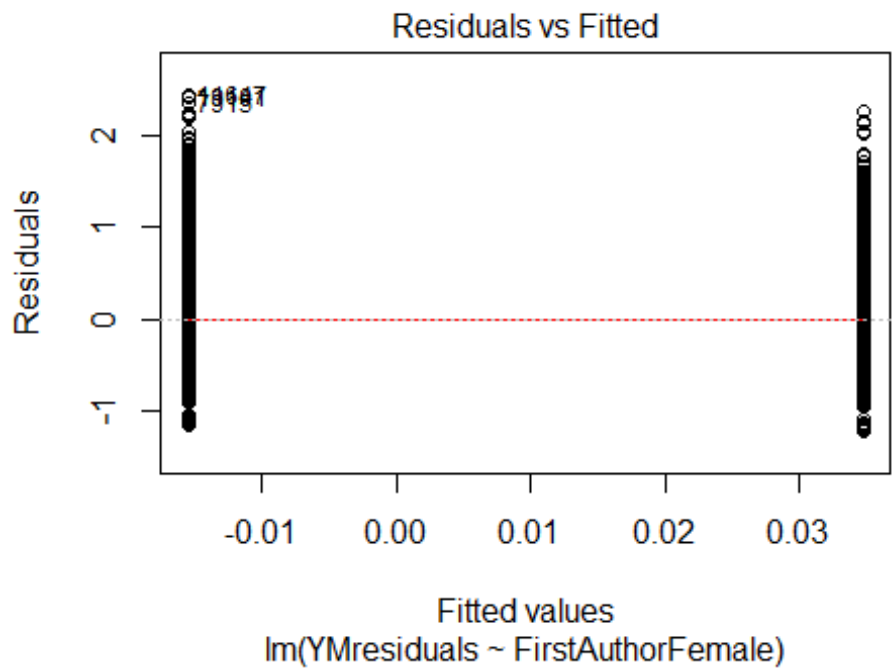

```
## [1] "List of  5 outliers with residuals above  2.5"
##          ScopusId  NLCS Year OneField Fields residuals
## 553    0030305327 3.789 1996     1400      4     2.649
## 2242   0000384657 3.626 1998     1403      2     2.524
## 11403 34047207566 3.996 2007     1400      4     2.660
## 16865 84906331429 4.168 2011     1400      5     3.082
## 17142 84866640011 3.733 2012     1400      5     2.719
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.3804 -0.5070  0.0146  0.5066  3.0685
##
## Coefficients:
##                   Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.0943     0.0305   35.93  < 2e-16 ***
## LastAuthorFemale1   0.1034     0.0164    6.29  3.2e-10 ***
## Year1997            0.0188     0.0467    0.40  0.68800
## Year1998           -0.0387     0.0449   -0.86  0.38781
## Year1999           -0.0229     0.0420   -0.55  0.58462
## Year2000           -0.0809     0.0424   -1.91  0.05640 .
## Year2001            0.1548     0.0440    3.52  0.00044 ***
## Year2002            0.0431     0.0415    1.04  0.29865
```

```
## Year2003            0.0253     0.0428    0.59  0.55441
## Year2004            0.1744     0.0435    4.01  6.2e-05 ***
## Year2005            0.0945     0.0416    2.27  0.02304 *
## Year2006            0.1827     0.0417    4.38  1.2e-05 ***
## Year2007            0.1021     0.0414    2.47  0.01371 *
## Year2008            0.0937     0.0399    2.35  0.01886 *
## Year2009            0.0111     0.0395    0.28  0.77952
## Year2010            0.0414     0.0406    1.02  0.30759
## Year2011            0.0052     0.0397    0.13  0.89584
## Year2012           -0.0657     0.0414   -1.59  0.11271
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.754
## Multiple R-squared:  0.0138, Adjusted R-squared:  0.0124
## Convergence in 10 IRWLS iterations
##
## Robustness weights:
##  984 weights are ~= 1. The remaining 10445 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0598  0.8680  0.9500  0.9150  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          8.75e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Sample size for the above analysis:  11429"
## [1] ""
## [1] ""
## [1] "##################################"
## [1] "Analysis of AJSC 1406"
## [1] "##################################"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  765  683  723  664  877  888  940  643  830  952  984 1042 1251 1514 1442
## 2011 2012
## 1418 1257
```

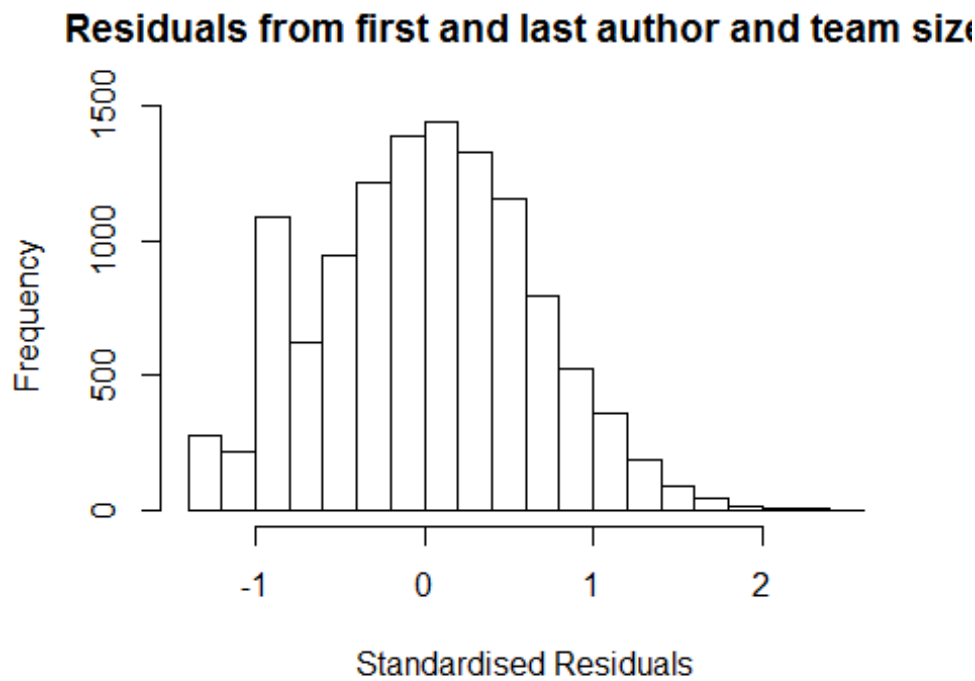
```
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  614  550  580  524  737  726  763  519  654  763  782  815  965 1140 1093
## 2011 2012
## 1034  949
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  567  489  520  481  665  664  690  469  583  669  700  709  851 1007  954
## 2011 2012
##  893  828
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
##  Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 130, df = 16, p-value <2e-16
```




```
##
##  Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 0.32, df = 1, p-value = 0.6
```

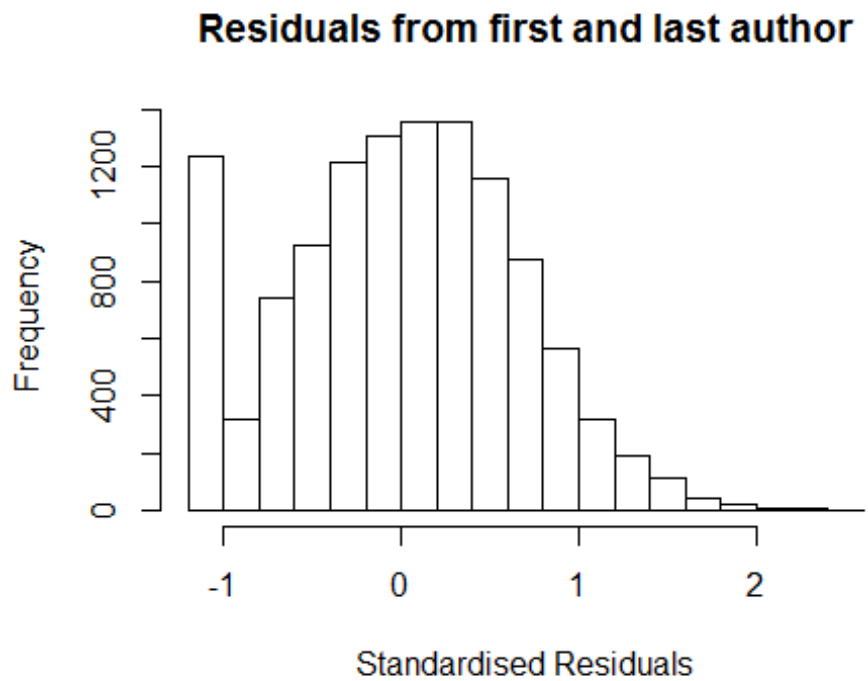



```
## [1] "Female first author team size 2018 geometric mean: 2.04890255574143"
## [1] "Male first author team size 2018 geometric mean: 2.03348452677508"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 62000, p-value = 0.8
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 1.92381398801688"
## [1] "Male last author team size 2018 geometric mean: 2.11775942729096"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 53000, p-value = 0.007
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##                     GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.231  1           1.110
## LastAuthorFemale  1.237  1           1.112
## UniqueAuthors     1.043  4           1.005
## Year              1.045 16           1.001
```



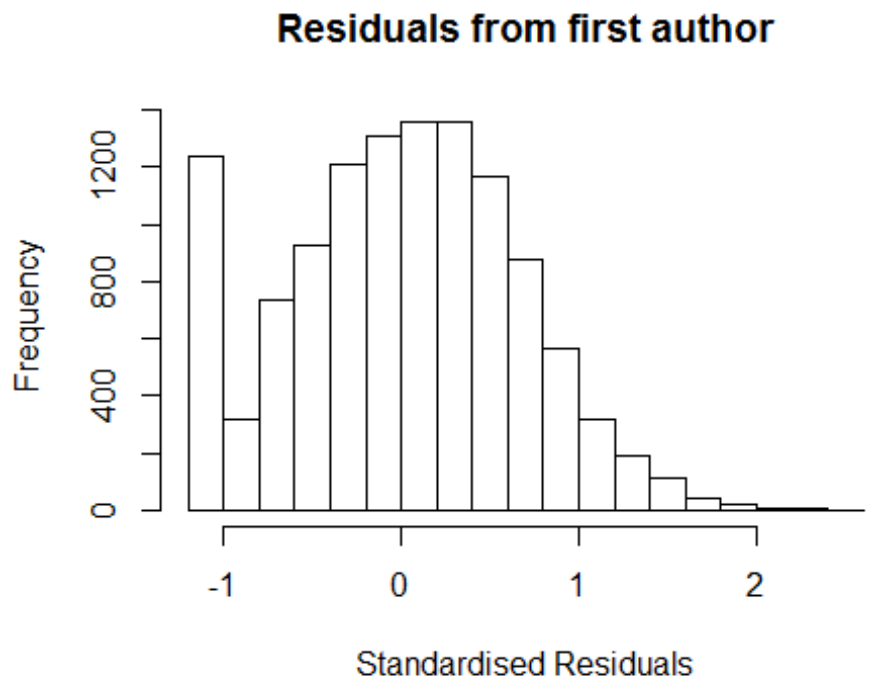

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
UniqueAuthors +
##     Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##     k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.3514 -0.4491  0.0137  0.4427  2.4418
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         0.92918    0.03367   27.60  < 2e-16 ***
## FirstAuthorFemale1  0.04586    0.01469    3.12  0.00180 **
## LastAuthorFemale1  -0.00758    0.01471   -0.52  0.60654
## UniqueAuthors2      0.29928    0.01437   20.83  < 2e-16 ***
## UniqueAuthors3      0.33558    0.01701   19.73  < 2e-16 ***
## UniqueAuthors4      0.27372    0.02987    9.16  < 2e-16 ***
## UniqueAuthors5      0.19271    0.05235    3.68  0.00023 ***
## Year1997            0.04074    0.04817    0.85  0.39773
## Year1998            0.03412    0.04746    0.72  0.47222
## Year1999            0.00181    0.04633    0.04  0.96879
```

```
## Year2000            0.04087    0.04222    0.97  0.33307
## Year2001            0.00530    0.04123    0.13  0.89779
## Year2002           -0.02705    0.04144   -0.65  0.51397
## Year2003           -0.08514    0.04559   -1.87  0.06189 .
## Year2004           -0.02318    0.04071   -0.57  0.56906
## Year2005           -0.04907    0.03952   -1.24  0.21441
## Year2006           -0.04617    0.03911   -1.18  0.23787
## Year2007           -0.04801    0.03946   -1.22  0.22372
## Year2008           -0.03764    0.03929   -0.96  0.33804
## Year2009           -0.06671    0.03858   -1.73  0.08384 .
## Year2010           -0.06694    0.03901   -1.72  0.08617 .
## Year2011           -0.04585    0.03973   -1.15  0.24855
## Year2012           -0.05024    0.04052   -1.24  0.21500
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.658
## Multiple R-squared:  0.0548, Adjusted R-squared:  0.0531
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  1006 weights are ~= 1. The remaining 10733 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.139   0.858   0.950   0.913   0.985   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          8.52e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.218  1           1.104
## LastAuthorFemale  1.216  1           1.103
## Year              1.016 16           1.000
```



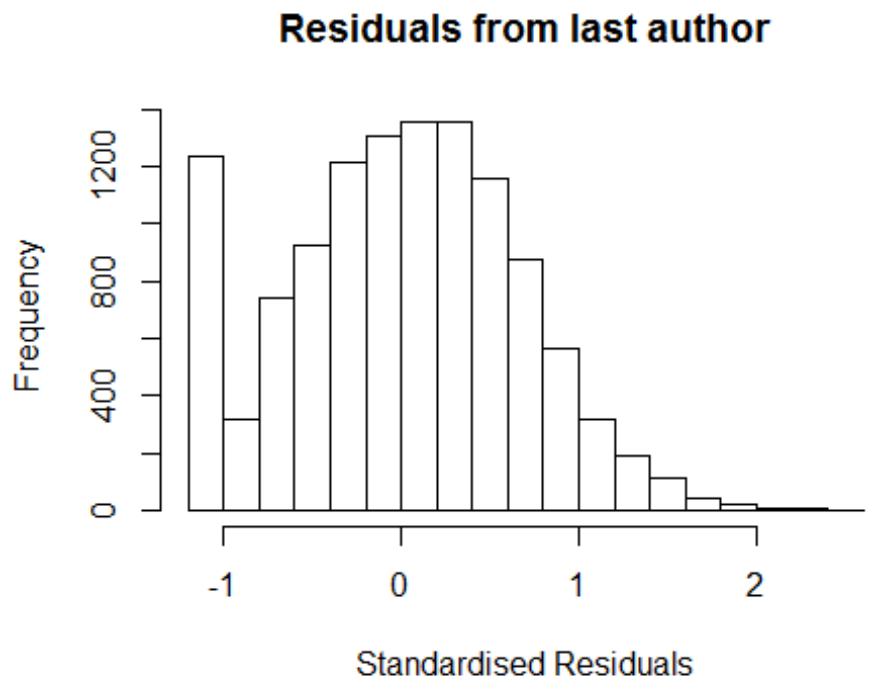

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.1961 -0.4601  0.0198  0.4577  2.4631
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)          1.10973    0.03272   33.91  < 2e-16 ***
## FirstAuthorFemale1   0.05863    0.01506    3.89  9.9e-05 ***
## LastAuthorFemale1   -0.00373    0.01506   -0.25    0.804
## Year1997             0.02448    0.04744    0.52    0.606
## Year1998             0.02765    0.04812    0.57    0.566
## Year1999            -0.00553    0.04690   -0.12    0.906
## Year2000             0.02774    0.04240    0.65    0.513
## Year2001            -0.01063    0.04109   -0.26    0.796
## Year2002            -0.02954    0.04204   -0.70    0.482
## Year2003            -0.08872    0.04665   -1.90    0.057 .
## Year2004            -0.00606    0.04111   -0.15    0.883
## Year2005            -0.04860    0.04023   -1.21    0.227
```

```
## Year2006           -0.04270    0.03951   -1.08    0.280
## Year2007           -0.03714    0.03975   -0.93    0.350
## Year2008           -0.04285    0.03990   -1.07    0.283
## Year2009           -0.06027    0.03904   -1.54    0.123
## Year2010           -0.06426    0.03960   -1.62    0.105
## Year2011           -0.03679    0.04018   -0.92    0.360
## Year2012           -0.03082    0.04078   -0.76    0.450
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.676
## Multiple R-squared:  0.00347,    Adjusted R-squared:  0.00194
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  923 weights are ~= 1. The remaining 10816 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.157   0.873   0.951   0.914   0.986   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          8.52e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.009  1           1.005
## Year              1.009 16           1.000
```




```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.1938 -0.4599  0.0203  0.4576  2.4637
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)          1.10915    0.03262   34.01  < 2e-16 ***
## FirstAuthorFemale1   0.05688    0.01373    4.14  3.5e-05 ***
## Year1997             0.02452    0.04744    0.52    0.605
## Year1998             0.02767    0.04812    0.58    0.565
## Year1999            -0.00540    0.04689   -0.12    0.908
## Year2000             0.02778    0.04240    0.66    0.512
## Year2001            -0.01069    0.04108   -0.26    0.795
## Year2002            -0.02963    0.04203   -0.70    0.481
## Year2003            -0.08889    0.04664   -1.91    0.057 .
## Year2004            -0.00603    0.04111   -0.15    0.883
## Year2005            -0.04852    0.04022   -1.21    0.228
## Year2006            -0.04269    0.03950   -1.08    0.280
```

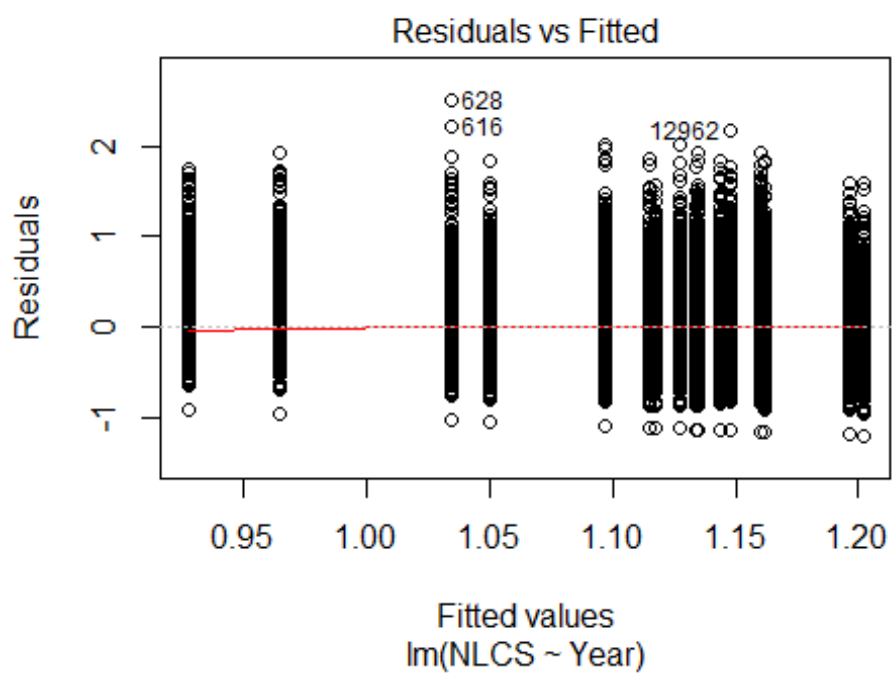
```
## Year2007           -0.03718    0.03975   -0.94    0.350
## Year2008           -0.04289    0.03990   -1.07    0.282
## Year2009           -0.06034    0.03903   -1.55    0.122
## Year2010           -0.06433    0.03960   -1.62    0.104
## Year2011           -0.03679    0.04018   -0.92    0.360
## Year2012           -0.03075    0.04077   -0.75    0.451
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.676
## Multiple R-squared:  0.00347,    Adjusted R-squared:  0.00202
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  923 weights are ~= 1. The remaining 10816 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.156   0.873   0.951   0.914   0.986   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          8.52e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##                   GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.008  1           1.004
## Year             1.008 16           1.000
```



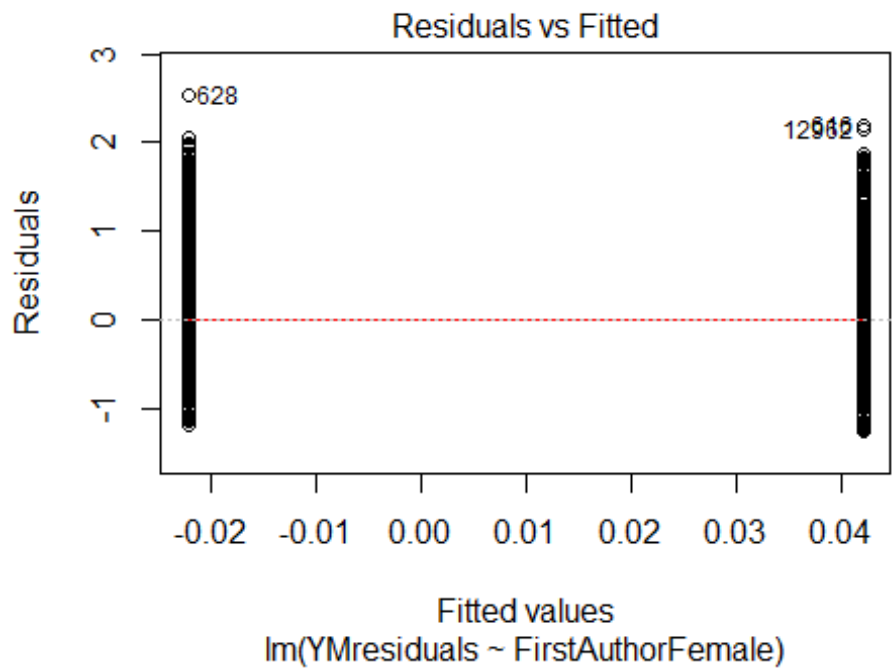

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.1707 -0.4593  0.0224  0.4570  2.4532
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.11551    0.03274   34.08   <2e-16 ***
## LastAuthorFemale1   0.02402    0.01369    1.75    0.079 .
## Year1997            0.02544    0.04753    0.54    0.593
## Year1998            0.02979    0.04817    0.62    0.536
## Year1999           -0.00242    0.04694   -0.05    0.959
## Year2000            0.03114    0.04242    0.73    0.463
## Year2001           -0.00673    0.04110   -0.16    0.870
## Year2002           -0.02606    0.04209   -0.62    0.536
## Year2003           -0.08522    0.04675   -1.82    0.068 .
## Year2004           -0.00259    0.04114   -0.06    0.950
## Year2005           -0.04305    0.04018   -1.07    0.284
## Year2006           -0.03729    0.03949   -0.94    0.345
```

```
## Year2007          -0.03296    0.03977   -0.83    0.407
## Year2008          -0.03873    0.03990   -0.97    0.332
## Year2009          -0.05719    0.03912   -1.46    0.144
## Year2010          -0.05858    0.03963   -1.48    0.139
## Year2011          -0.03221    0.04020   -0.80    0.423
## Year2012          -0.02439    0.04080   -0.60    0.550
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.677
## Multiple R-squared:  0.00219,    Adjusted R-squared:  0.000742
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  930 weights are ~= 1. The remaining 10809 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.161   0.872   0.950   0.914   0.985   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          8.52e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Sample size for the above analysis:  11739"
## [1] ""
## [1] ""
## [1] "####################################"
## [1] "Analysis of AJSC 1407"
## [1] "####################################"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  663  604  637  597  732  829  737  582  721  790  870  889 1011 1154 1263
## 2011 2012
## 1267 1201
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  593  527  574  532  655  688  656  511  642  694  760  762  870  987 1093
## 2011 2012
```

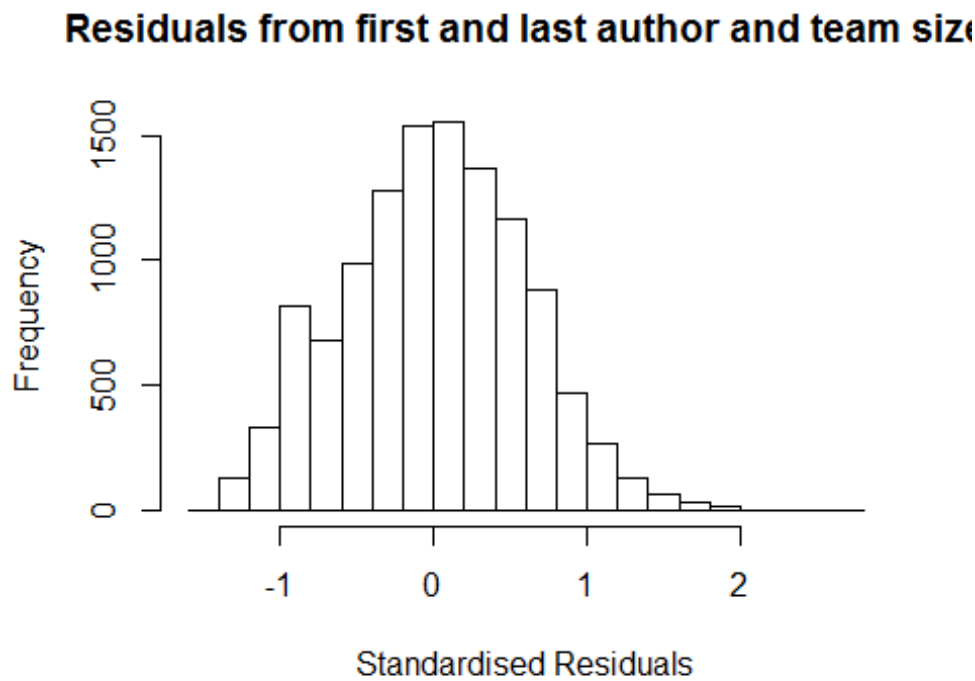
```
## 1087 1018
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  568  495  543  499  616  636  617  471  592  650  691  699  797  903 1007
## 2011 2012
##  996  930
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
##  Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 88, df = 16, p-value = 6e-12
```




```
##
##  Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 9.2, df = 1, p-value = 0.002
```

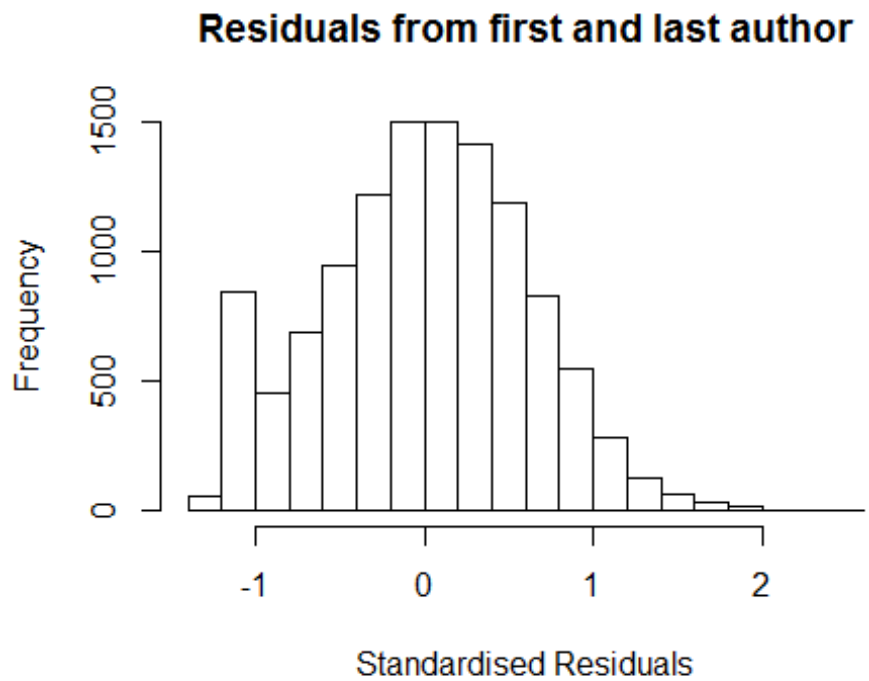



```
## [1] "Female first author team size 2018 geometric mean: 2.18587489787493"
## [1] "Male first author team size 2018 geometric mean: 2.07032406140681"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 77000, p-value = 0.2
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 2.13388917664031"
## [1] "Male last author team size 2018 geometric mean: 2.10878454533845"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 70000, p-value = 0.9
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.297  1           1.139
## LastAuthorFemale  1.291  1           1.136
## UniqueAuthors     1.056  4           1.007
## Year              1.060 16           1.002
```



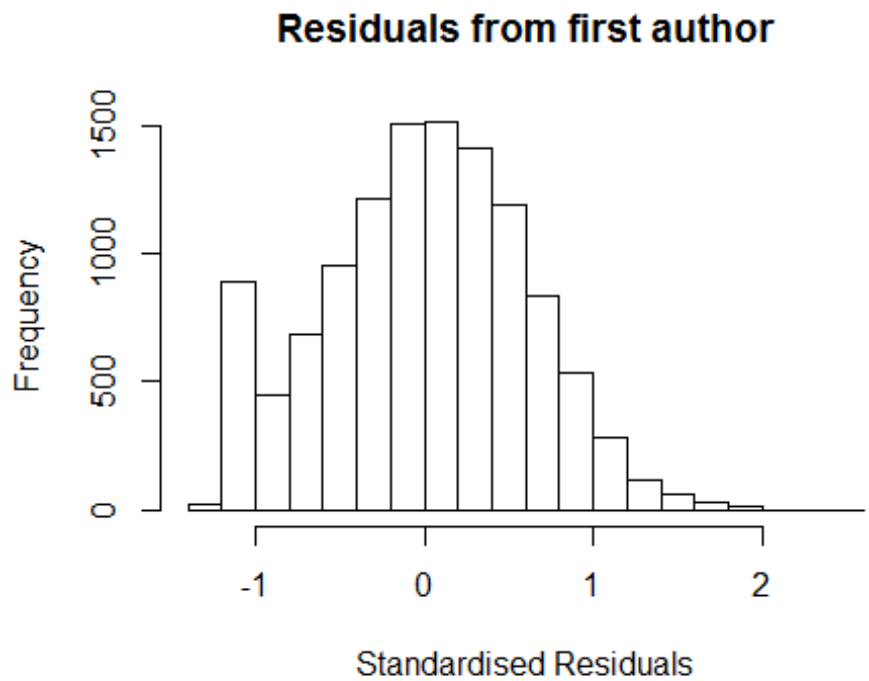

```
## [1] "List of  1 outliers with residuals above  2.5"
##       ScopusId  NLCS Year OneField Fields residuals
## 628 0030505108 3.545 1996     1405      3     2.667
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
UniqueAuthors +
##     Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##     k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.4180 -0.4027  0.0123  0.4144  2.6669
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)         8.78e-01   2.80e-02   31.37  < 2e-16 ***
## FirstAuthorFemale1  5.93e-02   1.34e-02    4.41    1e-05 ***
## LastAuthorFemale1  -5.43e-05   1.35e-02    0.00  0.99678
## UniqueAuthors2      2.29e-01   1.35e-02   17.03  < 2e-16 ***
## UniqueAuthors3      3.27e-01   1.60e-02   20.41  < 2e-16 ***
## UniqueAuthors4      3.87e-01   2.56e-02   15.08  < 2e-16 ***
## UniqueAuthors5      3.69e-01   3.46e-02   10.65  < 2e-16 ***
## Year1997           -1.09e-01   4.14e-02   -2.63  0.00860 **
## Year1998           -7.86e-02   3.78e-02   -2.08  0.03742 *
## Year1999            2.50e-03   3.85e-02    0.06  0.94823
```

```
## Year2000            5.24e-02   3.66e-02    1.43  0.15189
## Year2001            8.30e-02   3.76e-02    2.21  0.02717 *
## Year2002            9.86e-02   3.57e-02    2.76  0.00575 **
## Year2003            1.21e-01   3.72e-02    3.24  0.00120 **
## Year2004            7.47e-02   3.76e-02    1.99  0.04692 *
## Year2005            5.46e-02   3.54e-02    1.54  0.12304
## Year2006            4.99e-02   3.45e-02    1.44  0.14860
## Year2007            5.56e-02   3.47e-02    1.60  0.10896
## Year2008            1.12e-01   3.31e-02    3.37  0.00074 ***
## Year2009            7.30e-02   3.45e-02    2.11  0.03451 *
## Year2010            7.47e-02   3.43e-02    2.17  0.02966 *
## Year2011            6.48e-02   3.40e-02    1.91  0.05638 .
## Year2012            6.38e-03   3.53e-02    0.18  0.85683
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.604
## Multiple R-squared:  0.0684, Adjusted R-squared:  0.0666
## Convergence in 10 IRWLS iterations
##
## Robustness weights:
##  1024 weights are ~= 1. The remaining 10686 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0124  0.8680  0.9500  0.9100  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          8.54e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.299  1           1.140
## LastAuthorFemale  1.294  1           1.138
## Year              1.017 16           1.001
```



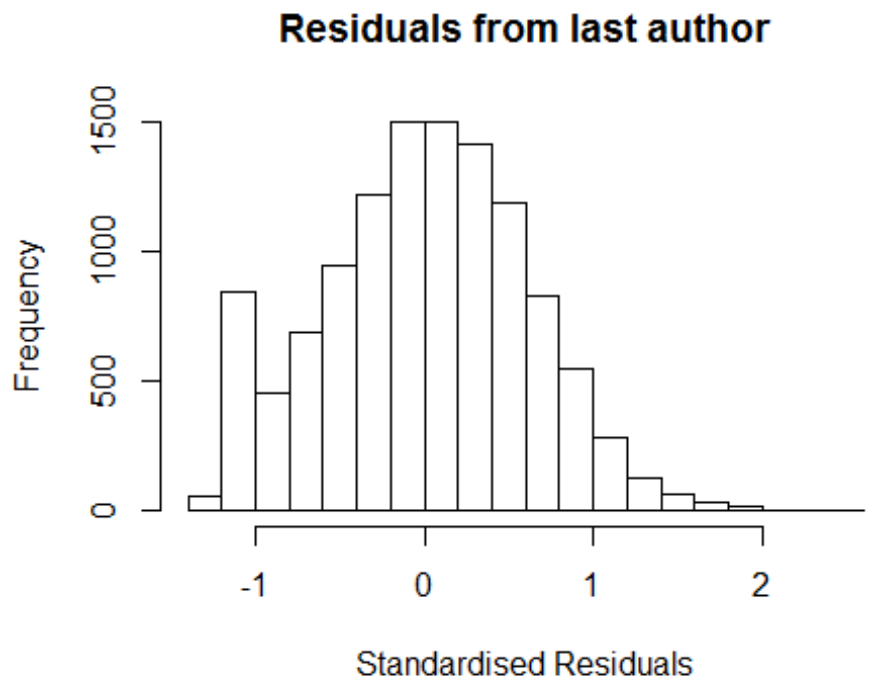

```
## [1] "List of  1 outliers with residuals above  2.5"
##       ScopusId  NLCS Year OneField Fields residuals
## 628 0030505108 3.545 1996     1405      3     2.554
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.2467 -0.4123  0.0205  0.4225  2.5543
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)          0.99074    0.02781   35.62  < 2e-16 ***
## FirstAuthorFemale1   0.07170    0.01385    5.18  2.3e-07 ***
## LastAuthorFemale1    0.00855    0.01393    0.61  0.53942
## Year1997            -0.11008    0.04183   -2.63  0.00851 **
## Year1998            -0.07716    0.03864   -2.00  0.04585 *
## Year1999             0.02205    0.03932    0.56  0.57490
## Year2000             0.08968    0.03734    2.40  0.01634 *
## Year2001             0.10124    0.03823    2.65  0.00810 **
## Year2002             0.13036    0.03643    3.58  0.00035 ***
## Year2003             0.16932    0.03792    4.46  8.1e-06 ***
## Year2004             0.10696    0.03816    2.80  0.00507 **
## Year2005             0.08532    0.03608    2.36  0.01806 *
```

```
## Year2006            0.09289    0.03552    2.61  0.00894 **
## Year2007            0.10051    0.03549    2.83  0.00463 **
## Year2008            0.17566    0.03394    5.18  2.3e-07 ***
## Year2009            0.11542    0.03541    3.26  0.00112 **
## Year2010            0.13037    0.03499    3.73  0.00020 ***
## Year2011            0.11801    0.03468    3.40  0.00067 ***
## Year2012            0.07082    0.03609    1.96  0.04973 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.618
## Multiple R-squared:  0.016,  Adjusted R-squared:  0.0145
## Convergence in 10 IRWLS iterations
##
## Robustness weights:
##  938 weights are ~= 1. The remaining 10772 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0489  0.8660  0.9500  0.9100  0.9860  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          8.54e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.013  1           1.006
## Year              1.013 16           1.000
```




```
## [1] "List of  1 outliers with residuals above  2.5"
##       ScopusId  NLCS Year OneField Fields residuals
## 628 0030505108 3.545 1996     1405      3     2.554
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.2440 -0.4124  0.0208  0.4225  2.5532
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)           0.9918     0.0277   35.77  < 2e-16 ***
## FirstAuthorFemale1    0.0761     0.0122    6.22  5.3e-10 ***
## Year1997             -0.1102     0.0418   -2.63  0.00845 **
## Year1998             -0.0767     0.0386   -1.98  0.04719 *
## Year1999              0.0224     0.0393    0.57  0.56884
## Year2000              0.0899     0.0373    2.41  0.01612 *
## Year2001              0.1015     0.0382    2.66  0.00792 **
## Year2002              0.1304     0.0364    3.58  0.00034 ***
## Year2003              0.1695     0.0379    4.47  7.9e-06 ***
## Year2004              0.1072     0.0382    2.81  0.00496 **
## Year2005              0.0855     0.0361    2.37  0.01782 *
## Year2006              0.0935     0.0355    2.63  0.00849 **
```

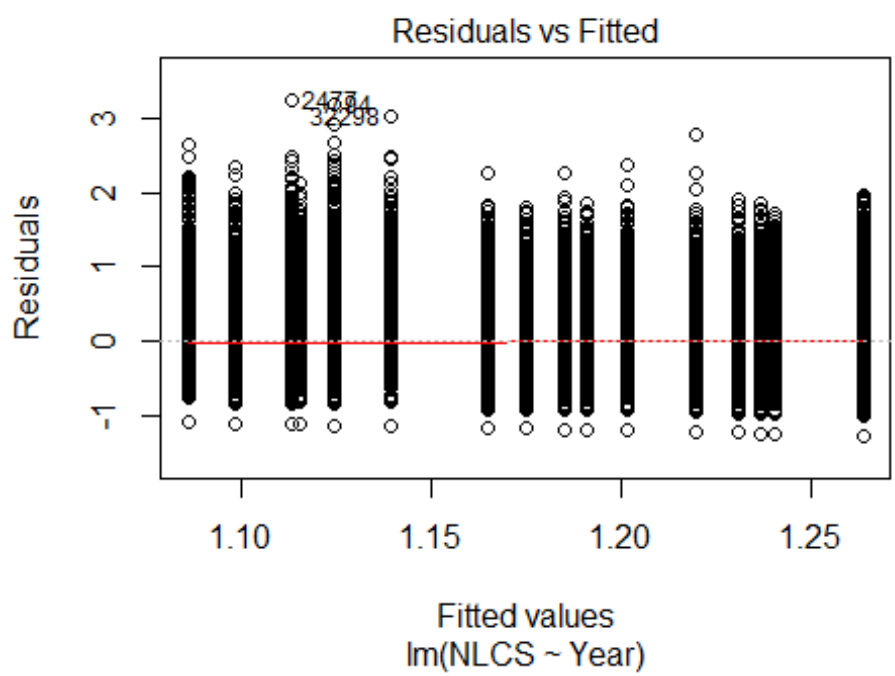
```
## Year2007             0.1011     0.0355    2.85  0.00438 **
## Year2008             0.1761     0.0339    5.19  2.1e-07 ***
## Year2009             0.1159     0.0354    3.27  0.00106 **
## Year2010             0.1307     0.0350    3.74  0.00019 ***
## Year2011             0.1185     0.0347    3.42  0.00064 ***
## Year2012             0.0715     0.0361    1.98  0.04759 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.618
## Multiple R-squared:  0.016,  Adjusted R-squared:  0.0145
## Convergence in 10 IRWLS iterations
##
## Robustness weights:
##  925 weights are ~= 1. The remaining 10785 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0492  0.8660  0.9500  0.9100  0.9860  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          8.54e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##                  GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.01  1           1.005
## Year             1.01 16           1.000
```



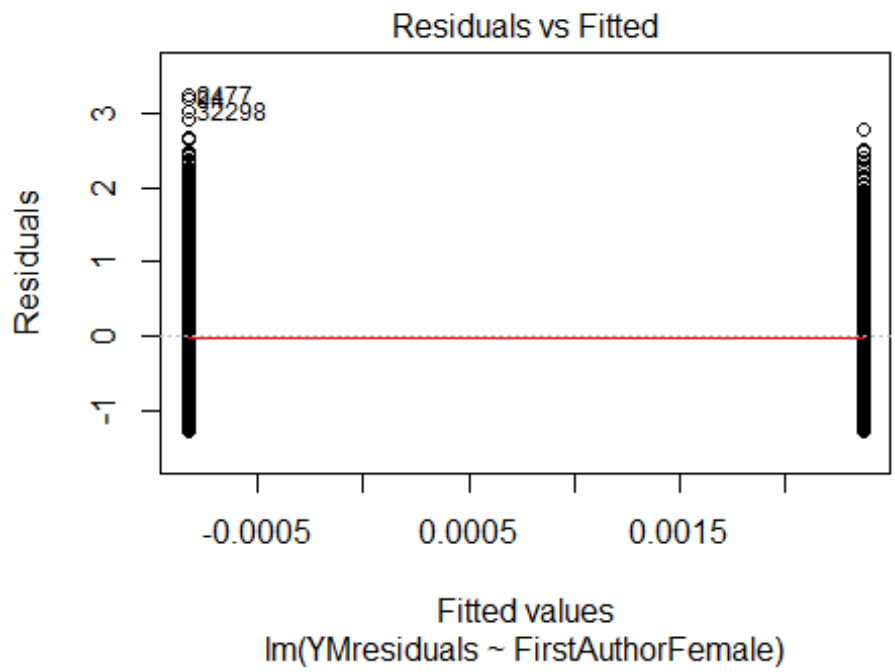

```
## [1] "List of  1 outliers with residuals above  2.5"
##       ScopusId  NLCS Year OneField Fields residuals
## 628 0030505108 3.545 1996     1405      3     2.554
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.2243 -0.4123  0.0186  0.4225  2.5437
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)          1.0013     0.0278   36.00  < 2e-16 ***
## LastAuthorFemale1    0.0459     0.0123    3.74  0.00019 ***
## Year1997            -0.1093     0.0420   -2.60  0.00932 **
## Year1998            -0.0790     0.0387   -2.04  0.04117 *
## Year1999             0.0211     0.0395    0.53  0.59370
## Year2000             0.0908     0.0375    2.43  0.01530 *
## Year2001             0.1028     0.0384    2.68  0.00743 **
## Year2002             0.1323     0.0365    3.62  0.00029 ***
## Year2003             0.1716     0.0380    4.51  6.4e-06 ***
## Year2004             0.1074     0.0383    2.81  0.00503 **
## Year2005             0.0897     0.0362    2.48  0.01326 *
## Year2006             0.0935     0.0356    2.63  0.00860 **
```

```
## Year2007            0.1020     0.0355    2.87  0.00408 **
## Year2008            0.1770     0.0340    5.20  2.0e-07 ***
## Year2009            0.1167     0.0355    3.28  0.00103 **
## Year2010            0.1341     0.0350    3.83  0.00013 ***
## Year2011            0.1194     0.0348    3.43  0.00060 ***
## Year2012            0.0749     0.0361    2.07  0.03825 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.618
## Multiple R-squared:  0.0138, Adjusted R-squared:  0.0124
## Convergence in 10 IRWLS iterations
##
## Robustness weights:
##  949 weights are ~= 1. The remaining 10761 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0526  0.8650  0.9490  0.9100  0.9860  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          8.54e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Sample size for the above analysis:  11710"
## [1] ""
## [1] ""
## [1] "####################################"
## [1] "Analysis of AJSC 1408"
## [1] "####################################"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 1612 1555 1521 1507 1555 1738 1528 1306 1354 1534 1663 1825 2036 2400 2375
## 2011 2012
## 2445 2273
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 1255 1166 1114 1162 1201 1209 1233 1064 1064 1209 1346 1434 1583 1856 1814
## 2011 2012
```

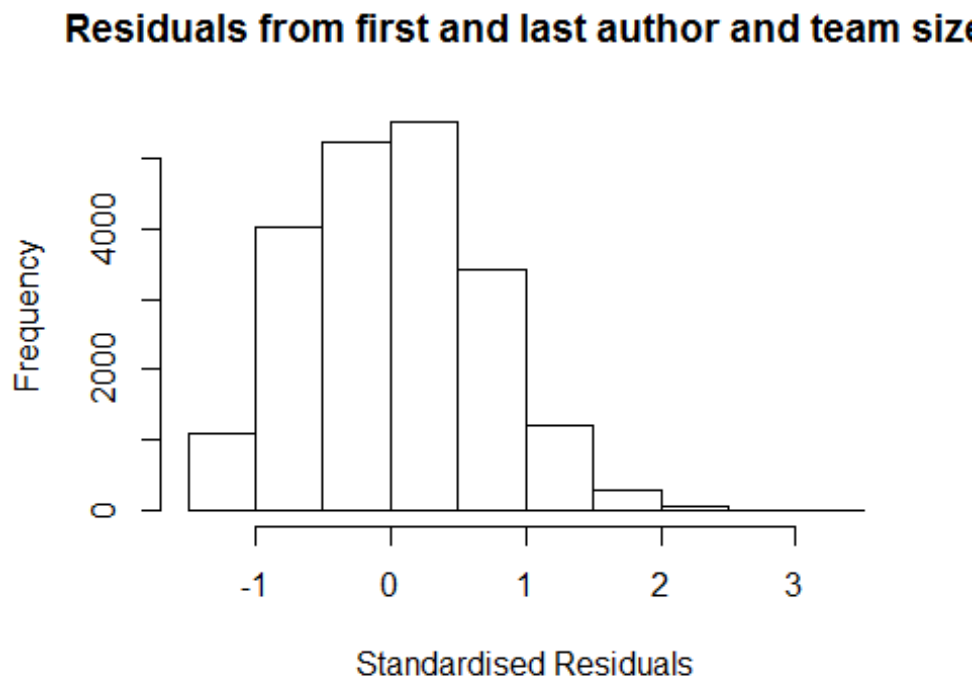
```
## 1924 1755
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 1166 1078 1024 1046 1100 1090 1120  960  928 1072 1188 1251 1399 1620 1598
## 2011 2012
## 1687 1556
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
##  Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 170, df = 16, p-value <2e-16
```




```
##
##  Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 0.92, df = 1, p-value = 0.3
```



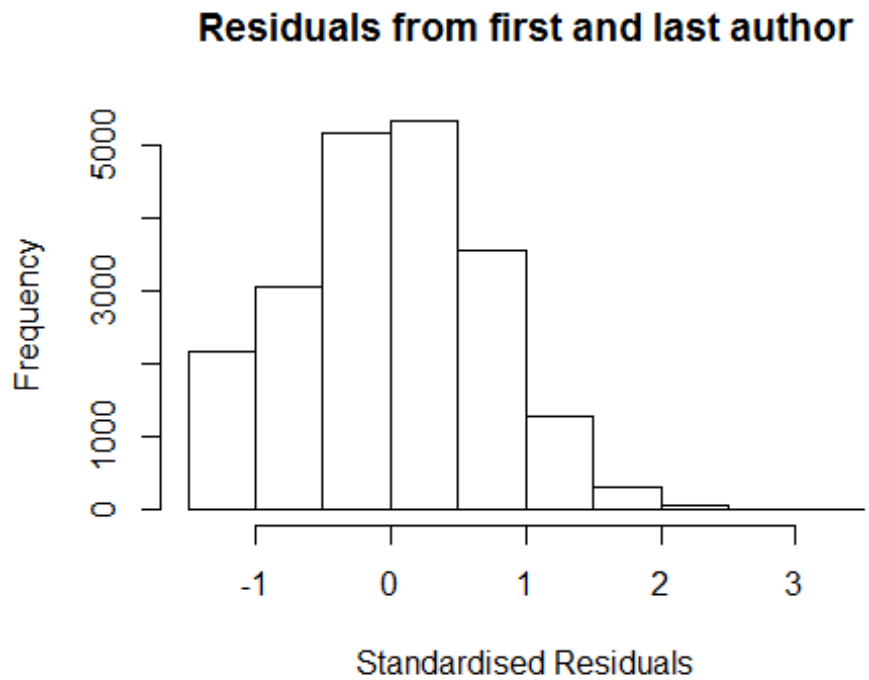

```
## [1] "Female first author team size 2018 geometric mean: 2.10807618038354"
## [1] "Male first author team size 2018 geometric mean: 1.88130372051053"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 310000, p-value = 5e-04
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 1.99855600933172"
## [1] "Male last author team size 2018 geometric mean: 1.93695941270657"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 270000, p-value = 0.5
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.321  1           1.149
## LastAuthorFemale  1.324  1           1.151
## UniqueAuthors     1.040  4           1.005
## Year              1.051 16           1.002
```




```
## [1] "List of  11 outliers with residuals above  2.5"
##          ScopusId  NLCS Year OneField Fields residuals
## 71     0041751098 4.048 1996     1403      2     3.121
## 94     2942649403 4.310 1996     1403      2     3.383
## 1149   0030305327 3.789 1996     1400      4     2.579
## 1246   0030505108 3.545 1996     1405      3     2.618
## 1301   0030530698 3.505 1996     1400      3     2.578
## 2025   0031271567 3.528 1997     1405      3     2.598
## 2477   0342775775 4.354 1997     1403      2     3.050
## 20961 34047207566 3.996 2007     1400      4     2.713
## 31244 79955593175 3.584 2011     1400      4     2.686
## 32298 84906331429 4.168 2011     1400      5     2.968
## 32950 84866640011 3.733 2012     1400      5     2.600
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
UniqueAuthors +
##     Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##     k.max = 1000))
##  \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.47069 -0.48861  0.00568  0.47429  3.38266
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
```

```
## (Intercept)         0.92734    0.02348   39.49  < 2e-16 ***
## FirstAuthorFemale1  0.01384    0.01265    1.09  0.27411
## LastAuthorFemale1  -0.01883    0.01278   -1.47  0.14058
## UniqueAuthors2      0.30116    0.01150   26.19  < 2e-16 ***
## UniqueAuthors3      0.38780    0.01382   28.07  < 2e-16 ***
## UniqueAuthors4      0.38969    0.02277   17.12  < 2e-16 ***
## UniqueAuthors5      0.27665    0.02937    9.42  < 2e-16 ***
## Year1997            0.00805    0.03361    0.24  0.81079
## Year1998           -0.02371    0.03355   -0.71  0.47977
## Year1999            0.06516    0.03064    2.13  0.03344 *
## Year2000            0.04411    0.03026    1.46  0.14492
## Year2001            0.14172    0.03152    4.50    7e-06 ***
## Year2002            0.10567    0.03025    3.49  0.00048 ***
## Year2003            0.10527    0.03044    3.46  0.00054 ***
## Year2004            0.10420    0.03182    3.28  0.00106 **
## Year2005            0.04332    0.03112    1.39  0.16391
## Year2006            0.03707    0.03018    1.23  0.21934
## Year2007            0.05912    0.03018    1.96  0.05017 .
## Year2008            0.04518    0.02954    1.53  0.12607
## Year2009           -0.03853    0.02911   -1.32  0.18573
## Year2010           -0.05253    0.02926   -1.80  0.07259 .
## Year2011           -0.02899    0.02898   -1.00  0.31717
## Year2012           -0.09532    0.02965   -3.22  0.00131 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.709
## Multiple R-squared:  0.0626, Adjusted R-squared:  0.0616
## Convergence in 10 IRWLS iterations
##
## Robustness weights:
##  observation 58 is an outlier with |weight| = 0 ( < 4.8e-06);
##  1734 weights are ~= 1. The remaining 19148 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0138  0.8680  0.9500  0.9150  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          4.79e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
```

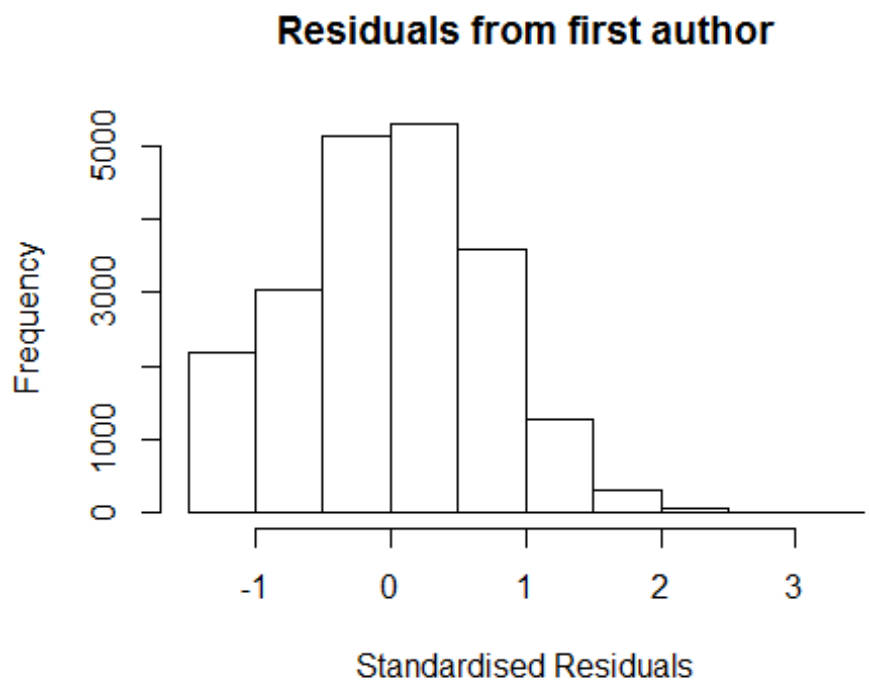
```
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.302  1           1.141
## LastAuthorFemale  1.302  1           1.141
## Year              1.013 16           1.000
```




```
## [1] "List of  9 outliers with residuals above  2.5"
##          ScopusId  NLCS Year OneField Fields residuals
## 71     0041751098 4.048 1996     1403      2     2.982
## 94     2942649403 4.310 1996     1403      2     3.244
## 1149   0030305327 3.789 1996     1400      4     2.744
## 2477   0342775775 4.354 1997     1403      2     3.293
## 20961 34047207566 3.996 2007     1400      4     2.805
## 31170 79952793118 3.630 2011     1400      3     2.527
## 31652 79551514315 3.634 2011     1408      4     2.530
## 32298 84906331429 4.168 2011     1400      5     3.064
## 32950 84866640011 3.733 2012     1400      5     2.686
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
```

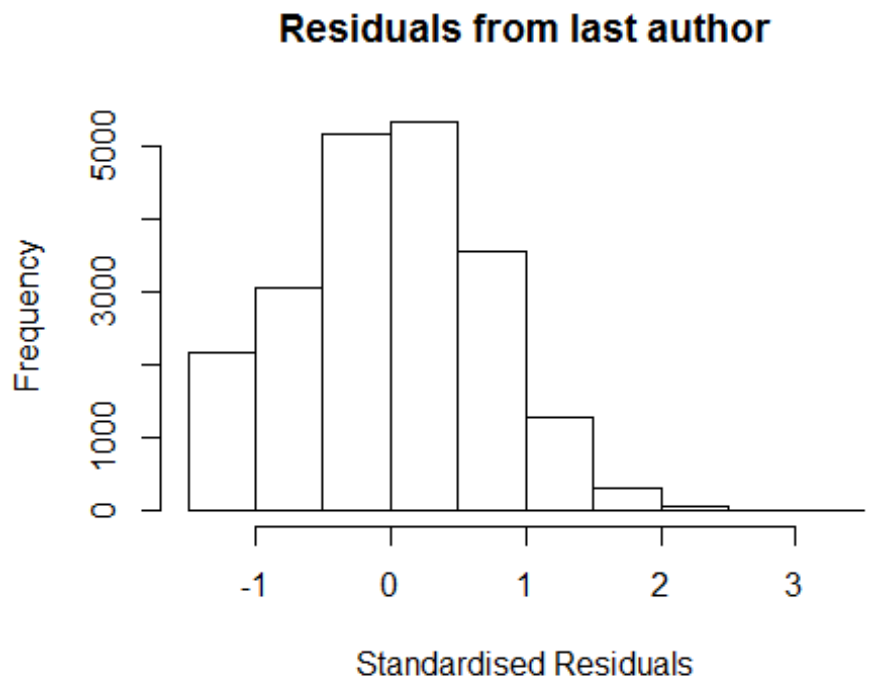
```
## -1.26310 -0.49988  0.00532  0.49702  3.29295
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.06611    0.02350   45.37  < 2e-16 ***
## FirstAuthorFemale1  0.02078    0.01294    1.61  0.10826
## LastAuthorFemale1  -0.02121    0.01305   -1.63  0.10402
## Year1997            0.01616    0.03440    0.47  0.63862
## Year1998           -0.00643    0.03423   -0.19  0.85112
## Year1999            0.10496    0.03110    3.37  0.00074 ***
## Year2000            0.08751    0.03092    2.83  0.00465 **
## Year2001            0.17621    0.03226    5.46  4.8e-08 ***
## Year2002            0.16001    0.03084    5.19  2.1e-07 ***
## Year2003            0.15074    0.03117    4.84  1.3e-06 ***
## Year2004            0.16042    0.03264    4.92  8.9e-07 ***
## Year2005            0.10259    0.03182    3.22  0.00127 **
## Year2006            0.10434    0.03085    3.38  0.00072 ***
## Year2007            0.12509    0.03078    4.06  4.8e-05 ***
## Year2008            0.11323    0.03021    3.75  0.00018 ***
## Year2009            0.02562    0.02972    0.86  0.38879
## Year2010            0.02027    0.02986    0.68  0.49717
## Year2011            0.03751    0.02974    1.26  0.20727
## Year2012           -0.01901    0.03006   -0.63  0.52719
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.732
## Multiple R-squared:  0.00784,    Adjusted R-squared:  0.00698
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  1748 weights are ~= 1. The remaining 19135 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0061  0.8720  0.9500  0.9160  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          4.79e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                    psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
```

```
## [1] "Regression 3: First author gender, Year as factors"
##                     GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.008  1           1.004
## Year              1.008 16           1.000
```




```
## [1] "List of  9 outliers with residuals above  2.5"
##          ScopusId  NLCS Year OneField Fields residuals
## 71     0041751098 4.048 1996     1403      2     2.982
## 94     2942649403 4.310 1996     1403      2     3.244
## 1149   0030305327 3.789 1996     1400      4     2.744
## 2477   0342775775 4.354 1997     1403      2     3.293
## 20961 34047207566 3.996 2007     1400      4     2.805
## 31170 79952793118 3.630 2011     1400      3     2.527
## 31652 79551514315 3.634 2011     1408      4     2.530
## 32298 84906331429 4.168 2011     1400      5     3.064
## 32950 84866640011 3.733 2012     1400      5     2.686
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.25019 -0.49914  0.00577  0.49802  3.27398
##
## Coefficients:
```

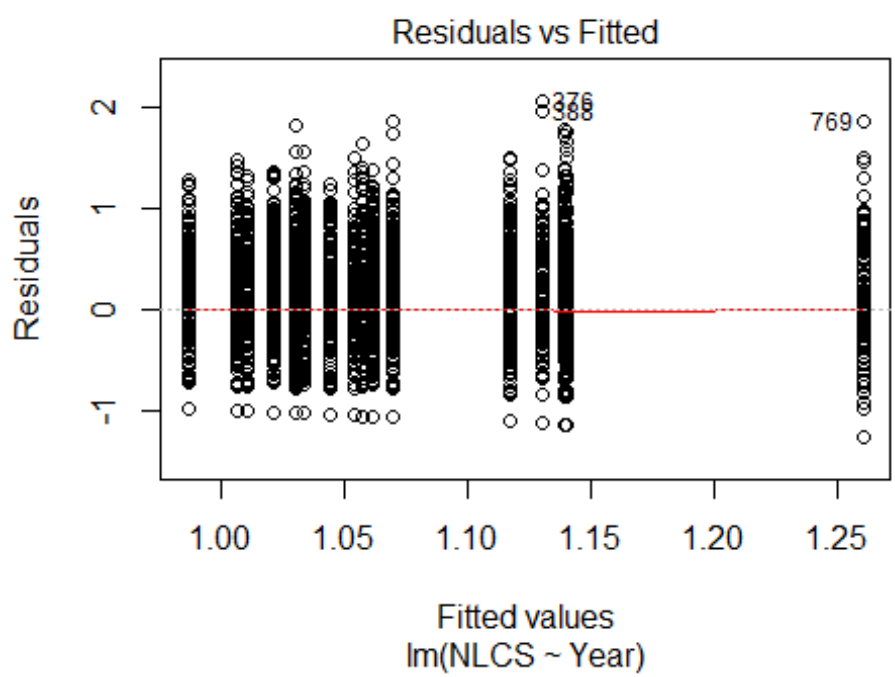
```
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)          1.06386    0.02344   45.38  < 2e-16 ***
## FirstAuthorFemale1  0.01020    0.01141    0.89  0.37140
## Year1997            0.01617    0.03439    0.47  0.63828
## Year1998           -0.00679    0.03424   -0.20  0.84287
## Year1999            0.10489    0.03110    3.37  0.00075 ***
## Year2000            0.08738    0.03092    2.83  0.00472 **
## Year2001            0.17614    0.03226    5.46  4.8e-08 ***
## Year2002            0.15926    0.03085    5.16  2.5e-07 ***
## Year2003            0.15094    0.03117    4.84  1.3e-06 ***
## Year2004            0.16009    0.03265    4.90  9.5e-07 ***
## Year2005            0.10252    0.03184    3.22  0.00129 **
## Year2006            0.10406    0.03085    3.37  0.00075 ***
## Year2007            0.12410    0.03076    4.03  5.5e-05 ***
## Year2008            0.11258    0.03022    3.73  0.00020 ***
## Year2009            0.02497    0.02973    0.84  0.40084
## Year2010            0.01980    0.02986    0.66  0.50733
## Year2011            0.03670    0.02974    1.23  0.21716
## Year2012           -0.02007    0.03005   -0.67  0.50426
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.732
## Multiple R-squared:  0.00772,    Adjusted R-squared:  0.00691
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  1735 weights are ~= 1. The remaining 19148 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.008   0.872   0.950   0.916   0.985   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          4.79e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##                   GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.008  1           1.004
## Year             1.008 16           1.000
```



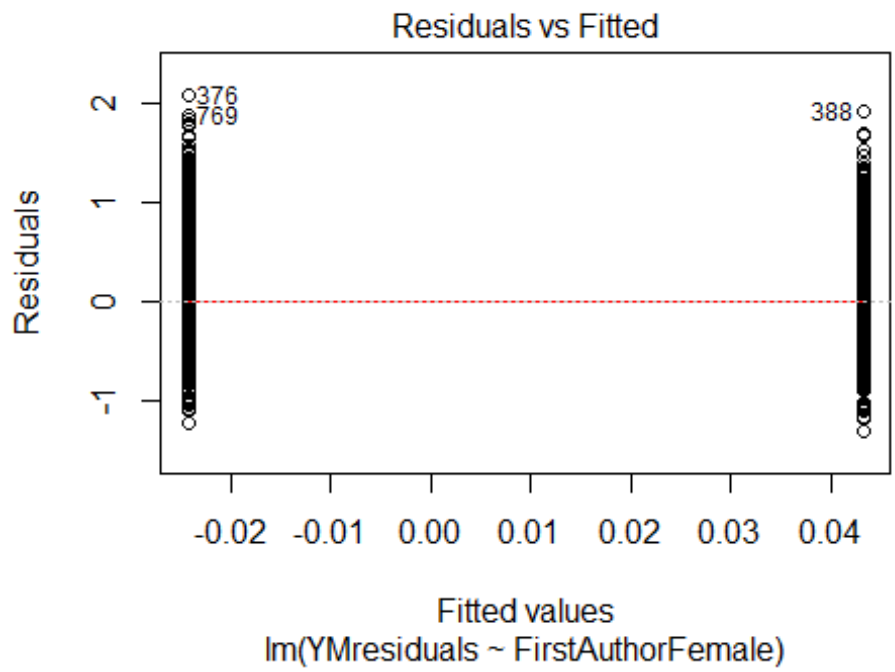

```
## [1] "List of  9 outliers with residuals above  2.5"
##          ScopusId  NLCS Year OneField Fields residuals
## 71     0041751098 4.048 1996     1403      2     2.982
## 94     2942649403 4.310 1996     1403      2     3.244
## 1149   0030305327 3.789 1996     1400      4     2.744
## 2477   0342775775 4.354 1997     1403      2     3.293
## 20961 34047207566 3.996 2007     1400      4     2.805
## 31170 79952793118 3.630 2011     1400      3     2.527
## 31652 79551514315 3.634 2011     1408      4     2.530
## 32298 84906331429 4.168 2011     1400      5     3.064
## 32950 84866640011 3.733 2012     1400      5     2.686
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.24517 -0.50028  0.00572  0.49683  3.28010
##
## Coefficients:
##                   Estimate Std. Error t value Pr(>|t|)
## (Intercept)        1.06829    0.02344   45.58  < 2e-16 ***
## LastAuthorFemale1 -0.01069    0.01147   -0.93  0.35113
## Year1997           0.01630    0.03440    0.47  0.63564
## Year1998          -0.00663    0.03422   -0.19  0.84643
```

```
## Year1999           0.10526    0.03110    3.38  0.00072 ***
## Year2000           0.08794    0.03092    2.84  0.00446 **
## Year2001           0.17688    0.03225    5.48  4.2e-08 ***
## Year2002           0.16015    0.03084    5.19  2.1e-07 ***
## Year2003           0.15186    0.03117    4.87  1.1e-06 ***
## Year2004           0.16095    0.03264    4.93  8.3e-07 ***
## Year2005           0.10332    0.03181    3.25  0.00116 **
## Year2006           0.10519    0.03083    3.41  0.00065 ***
## Year2007           0.12535    0.03076    4.08  4.6e-05 ***
## Year2008           0.11415    0.03020    3.78  0.00016 ***
## Year2009           0.02607    0.02972    0.88  0.38042
## Year2010           0.02145    0.02984    0.72  0.47212
## Year2011           0.03827    0.02973    1.29  0.19811
## Year2012          -0.01801    0.03005   -0.60  0.54896
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.732
## Multiple R-squared:  0.00772,    Adjusted R-squared:  0.00691
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  1727 weights are ~= 1. The remaining 19156 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0073  0.8720  0.9500  0.9160  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          4.79e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Sample size for the above analysis:  20883"
## [1] ""
## [1] ""
## [1] "####################################"
## [1] "Analysis of AJSC 1409"
## [1] "####################################"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
```

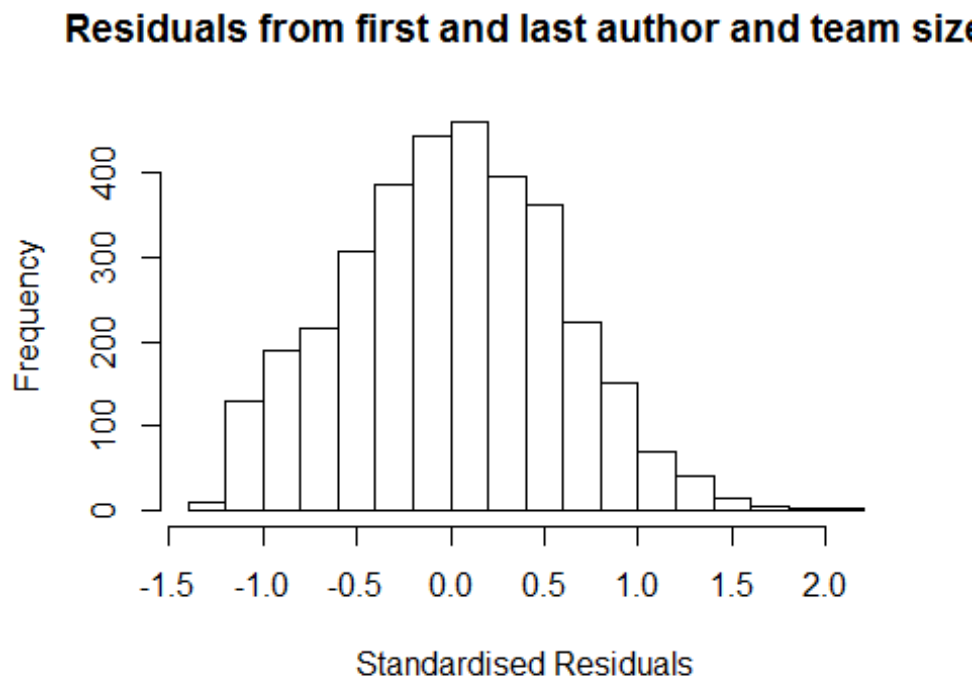
```
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  196  167  194  190  244  273  286  245  305  269  309  312  388  454  435
## 2011 2012
##  468  508
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  163  136  148  144  180  220  228  196  244  212  234  228  284  329  285
## 2011 2012
##  321  357
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  152  130  129  126  168  194  204  178  206  189  201  201  243  277  236
## 2011 2012
##  272  303
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
##  Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 57, df = 16, p-value = 1e-06
```




```
##
##  Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 3.1, df = 1, p-value = 0.08
```

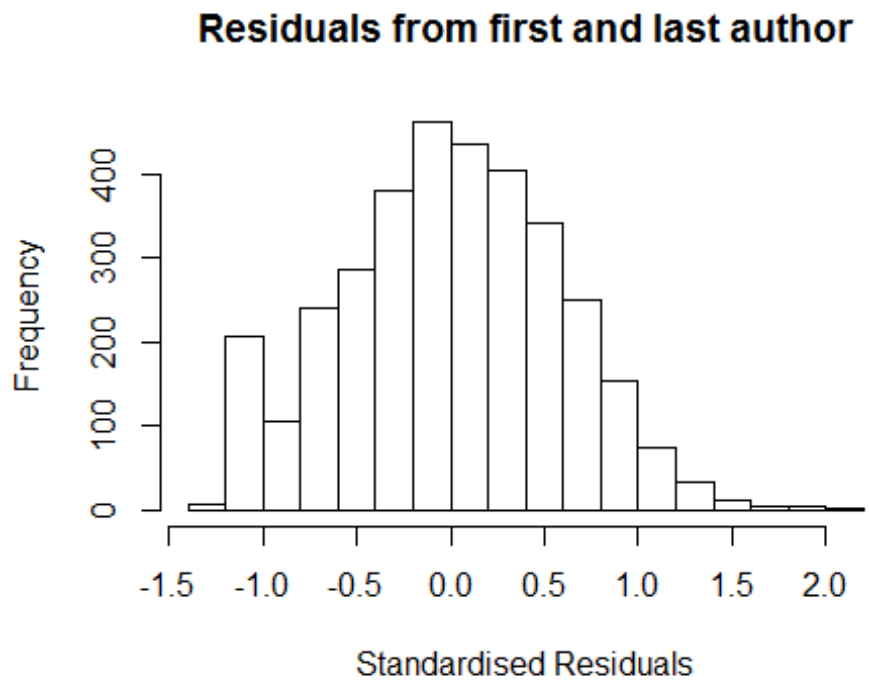



```
## [1] "Female first author team size 2018 geometric mean: 1.93661167595403"
## [1] "Male first author team size 2018 geometric mean: 2.07998567904896"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 13000, p-value = 0.3
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 1.87658226877107"
## [1] "Male last author team size 2018 geometric mean: 2.11317747622411"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 12000, p-value = 0.06
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##                     GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.277  1           1.130
## LastAuthorFemale  1.266  1           1.125
## UniqueAuthors     1.102  4           1.012
## Year              1.152 16           1.004
```



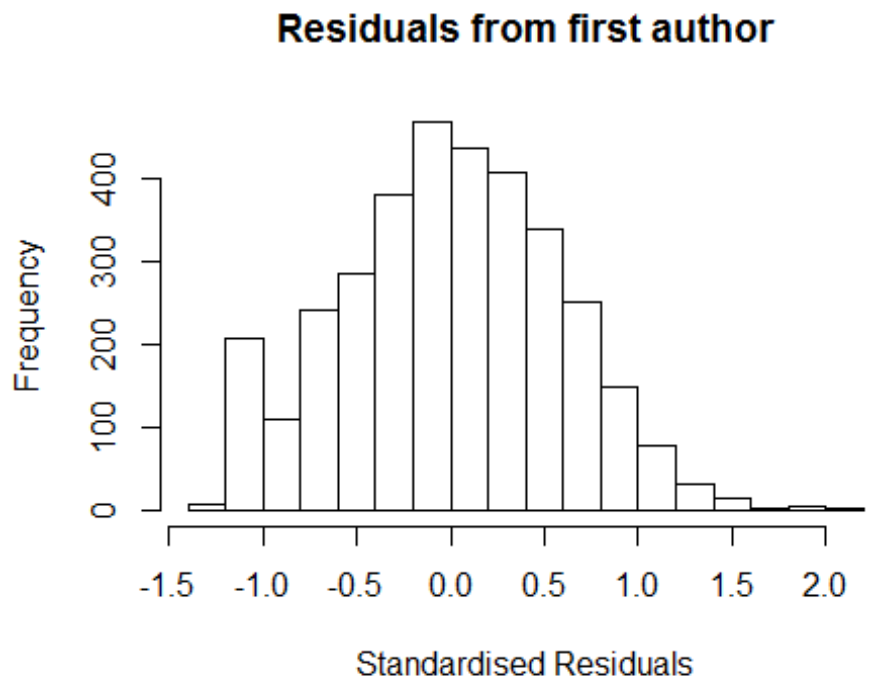

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
UniqueAuthors +
##     Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##     k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.2651 -0.4003  0.0092  0.4103  2.0400
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)         0.938000   0.059715   15.71  < 2e-16 ***
## FirstAuthorFemale1  0.065433   0.023757    2.75   0.0059 **
## LastAuthorFemale1   0.008060   0.023798    0.34   0.7349
## UniqueAuthors2      0.125716   0.024469    5.14  2.9e-07 ***
## UniqueAuthors3      0.153039   0.028444    5.38  7.9e-08 ***
## UniqueAuthors4      0.245781   0.044923    5.47  4.8e-08 ***
## UniqueAuthors5      0.282595   0.071789    3.94  8.4e-05 ***
## Year1997            0.079267   0.078597    1.01   0.3133
## Year1998            0.129934   0.082850    1.57   0.1169
## Year1999            0.200317   0.077743    2.58   0.0100 *
```

```
## Year2000            0.000128   0.076478    0.00   0.9987
## Year2001            0.070371   0.080700    0.87   0.3833
## Year2002            0.058510   0.074263    0.79   0.4308
## Year2003           -0.044317   0.074543   -0.59   0.5522
## Year2004           -0.046290   0.071532   -0.65   0.5176
## Year2005           -0.009076   0.070439   -0.13   0.8975
## Year2006           -0.100850   0.069474   -1.45   0.1467
## Year2007           -0.014030   0.070719   -0.20   0.8428
## Year2008           -0.005896   0.068615   -0.09   0.9315
## Year2009           -0.068001   0.067291   -1.01   0.3123
## Year2010           -0.029006   0.067546   -0.43   0.6676
## Year2011           -0.045201   0.066537   -0.68   0.4970
## Year2012           -0.041548   0.067489   -0.62   0.5382
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.596
## Multiple R-squared:  0.032,  Adjusted R-squared:  0.0257
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  296 weights are ~= 1. The remaining 3113 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.217   0.873   0.950   0.913   0.986   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          2.93e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.286  1           1.134
## LastAuthorFemale  1.276  1           1.130
## Year              1.049 16           1.001
```



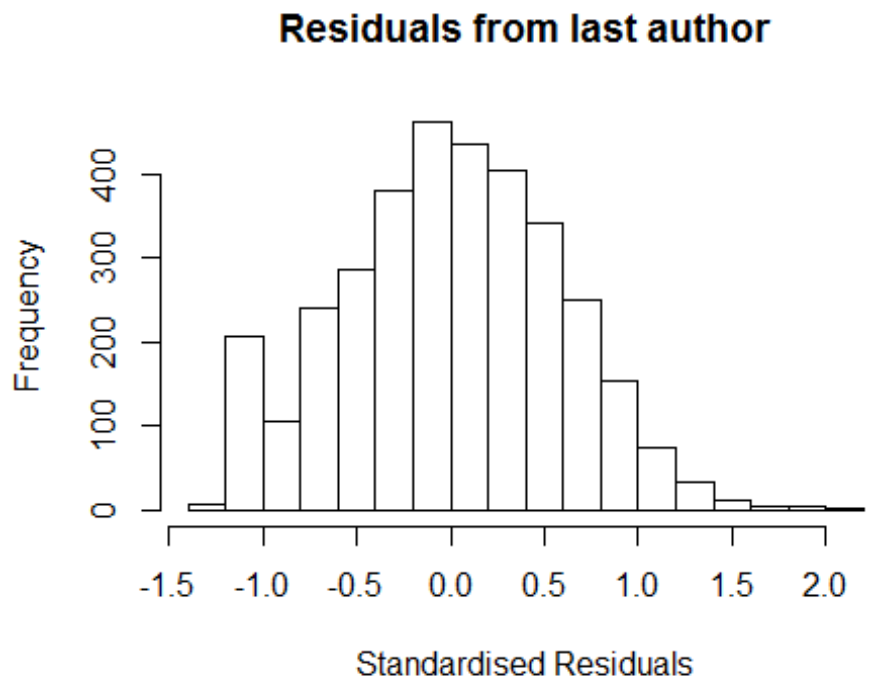

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.30440 -0.39845  0.00629  0.41349  2.09632
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)          1.00873    0.05963   16.92   <2e-16 ***
## FirstAuthorFemale1   0.07123    0.02412    2.95   0.0032 **
## LastAuthorFemale1    0.00929    0.02414    0.38   0.7004
## Year1997             0.07795    0.07884    0.99   0.3229
## Year1998             0.12872    0.08385    1.54   0.1248
## Year1999             0.21515    0.07903    2.72   0.0065 **
## Year2000            -0.00468    0.07726   -0.06   0.9517
## Year2001             0.07761    0.08216    0.94   0.3449
## Year2002             0.06905    0.07546    0.92   0.3602
## Year2003            -0.03032    0.07555   -0.40   0.6882
## Year2004            -0.03199    0.07257   -0.44   0.6594
## Year2005             0.01017    0.07147    0.14   0.8868
```

```
## Year2006           -0.08286    0.07086   -1.17   0.2423
## Year2007            0.00976    0.07189    0.14   0.8920
## Year2008            0.01579    0.06969    0.23   0.8208
## Year2009           -0.05224    0.06868   -0.76   0.4469
## Year2010           -0.00739    0.06878   -0.11   0.9145
## Year2011           -0.02942    0.06794   -0.43   0.6651
## Year2012           -0.00536    0.06825   -0.08   0.9374
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.6
## Multiple R-squared:  0.0142, Adjusted R-squared:  0.00898
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  302 weights are ~= 1. The remaining 3107 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.197   0.874   0.950   0.913   0.986   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          2.93e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.026  1           1.013
## Year              1.026 16           1.001
```




```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.3015 -0.3992  0.0048  0.4137  2.0948
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)          1.01030    0.05943   17.00  < 2e-16 ***
## FirstAuthorFemale1   0.07583    0.02158    3.51  0.00045 ***
## Year1997             0.07788    0.07887    0.99  0.32353
## Year1998             0.12791    0.08388    1.52  0.12736
## Year1999             0.21541    0.07909    2.72  0.00649 **
## Year2000            -0.00483    0.07727   -0.06  0.95012
## Year2001             0.07746    0.08221    0.94  0.34617
## Year2002             0.06901    0.07550    0.91  0.36082
## Year2003            -0.03042    0.07556   -0.40  0.68728
## Year2004            -0.03185    0.07261   -0.44  0.66094
## Year2005             0.01021    0.07149    0.14  0.88649
## Year2006            -0.08211    0.07088   -1.16  0.24676
```

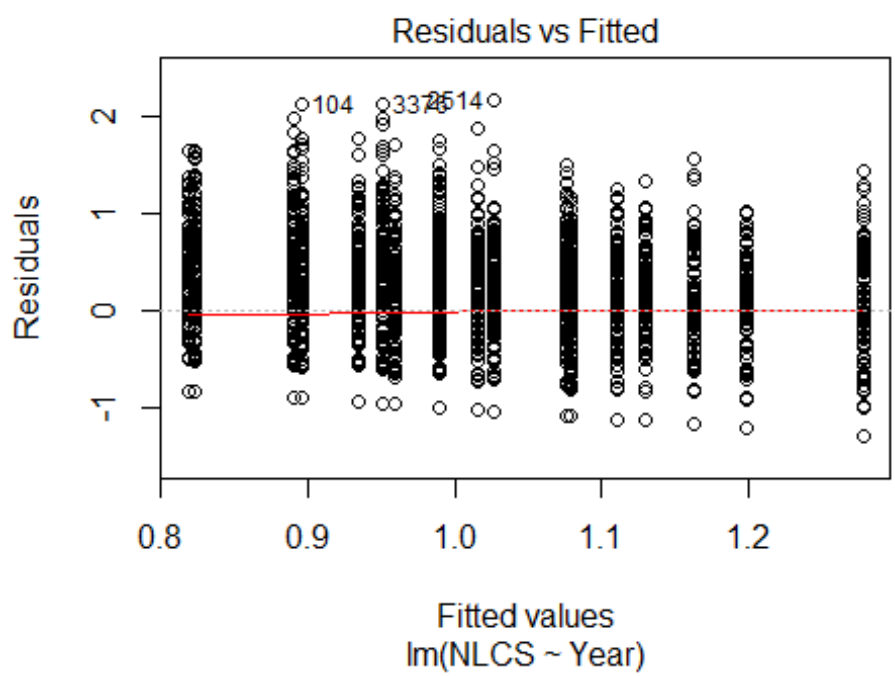
```
## Year2007            0.00931    0.07188    0.13  0.89690
## Year2008            0.01589    0.06972    0.23  0.81972
## Year2009           -0.05180    0.06871   -0.75  0.45096
## Year2010           -0.00710    0.06884   -0.10  0.91786
## Year2011           -0.02917    0.06798   -0.43  0.66783
## Year2012           -0.00599    0.06824   -0.09  0.93001
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.6
## Multiple R-squared:  0.0142, Adjusted R-squared:  0.00923
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  303 weights are ~= 1. The remaining 3106 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.198   0.874   0.950   0.913   0.986   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          2.93e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##                   GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.019  1           1.010
## Year             1.019 16           1.001
```



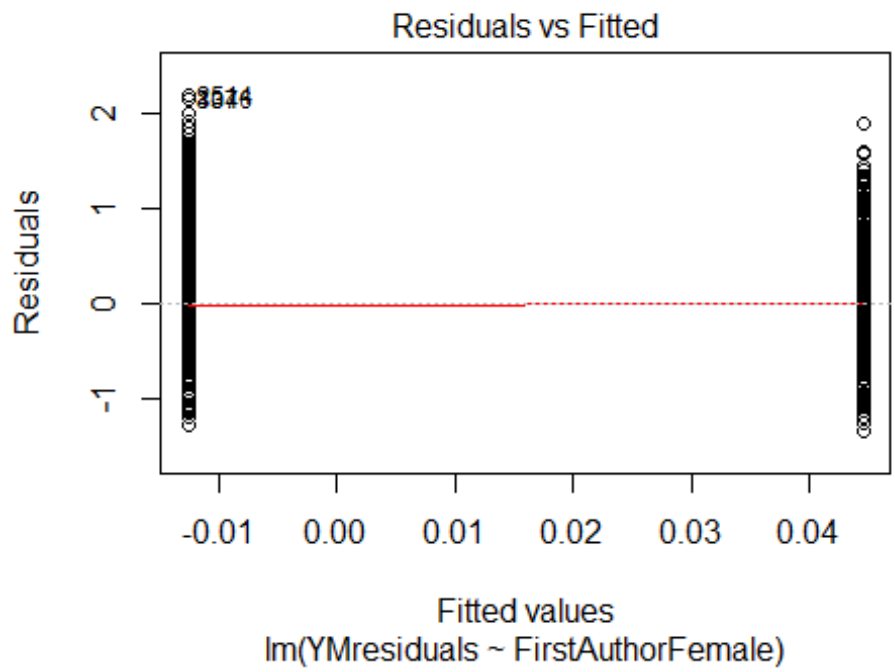

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.28200 -0.40460  0.00536  0.41021  2.09056
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)        1.02e+00   5.93e-02   17.17   <2e-16 ***
## LastAuthorFemale1  4.51e-02   2.16e-02    2.09   0.0369 *
## Year1997           7.34e-02   7.87e-02    0.93   0.3514
## Year1998           1.31e-01   8.39e-02    1.56   0.1199
## Year1999           2.18e-01   7.88e-02    2.77   0.0057 **
## Year2000          -4.23e-06   7.73e-02    0.00   1.0000
## Year2001           7.86e-02   8.21e-02    0.96   0.3385
## Year2002           6.93e-02   7.51e-02    0.92   0.3561
## Year2003          -2.61e-02   7.52e-02   -0.35   0.7286
## Year2004          -2.90e-02   7.23e-02   -0.40   0.6887
## Year2005           1.26e-02   7.13e-02    0.18   0.8595
## Year2006          -8.21e-02   7.07e-02   -1.16   0.2455
```

```
## Year2007           1.92e-02   7.17e-02    0.27   0.7892
## Year2008           1.81e-02   6.95e-02    0.26   0.7945
## Year2009          -5.22e-02   6.86e-02   -0.76   0.4472
## Year2010          -1.28e-03   6.85e-02   -0.02   0.9851
## Year2011          -2.65e-02   6.77e-02   -0.39   0.6956
## Year2012           1.64e-03   6.79e-02    0.02   0.9807
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.601
## Multiple R-squared:  0.0118, Adjusted R-squared:  0.00684
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  292 weights are ~= 1. The remaining 3117 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.201   0.875   0.950   0.913   0.986   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          2.93e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Sample size for the above analysis:  3409"
## [1] ""
## [1] ""
## [1] "####################################"
## [1] "Analysis of AJSC 1410"
## [1] "####################################"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  231  240  243  185  180  141  106  128  151  203  178  167  181  311  311
## 2011 2012
##  265  245
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  181  181  184  142  146   93   85  107  123  175  144  128  147  249  237
## 2011 2012
```

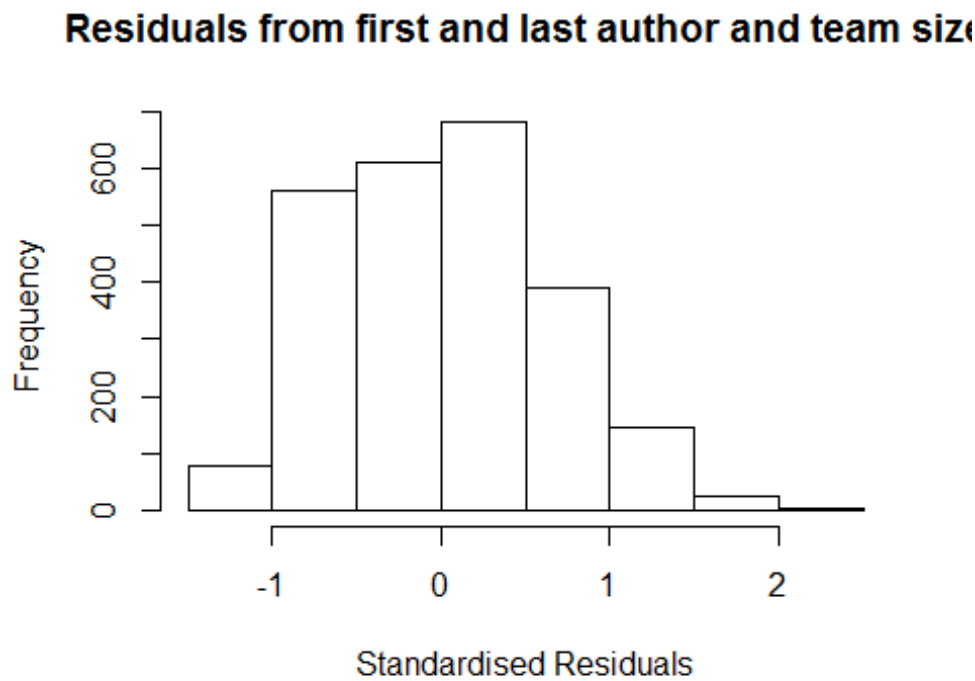
```
##  218  199
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  172  165  167  132  128   86   73   96  114  156  129  114  136  221  221
## 2011 2012
##  201  182
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
##  Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 49, df = 16, p-value = 4e-05
```




```
##
##  Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 0.13, df = 1, p-value = 0.7
```

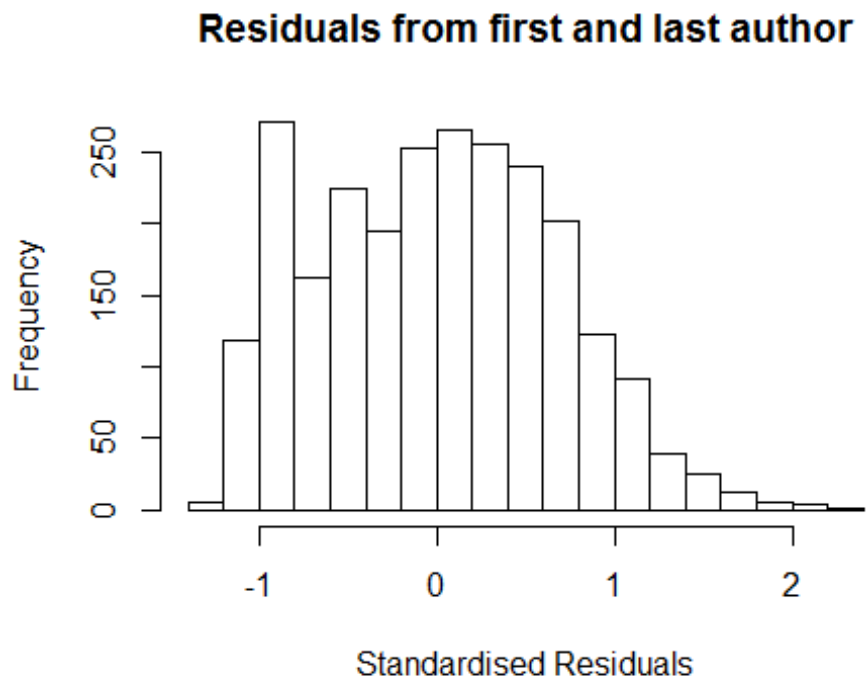



```
## [1] "Female first author team size 2018 geometric mean: 1.6571528133767"
## [1] "Male first author team size 2018 geometric mean: 1.59370581691238"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 4800, p-value = 0.8
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 1.66871211664546"
## [1] "Male last author team size 2018 geometric mean: 1.58738866513215"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 5000, p-value = 0.7
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##                     GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.326  1           1.152
## LastAuthorFemale  1.314  1           1.146
## UniqueAuthors     1.156  4           1.018
## Year              1.174 16           1.005
```



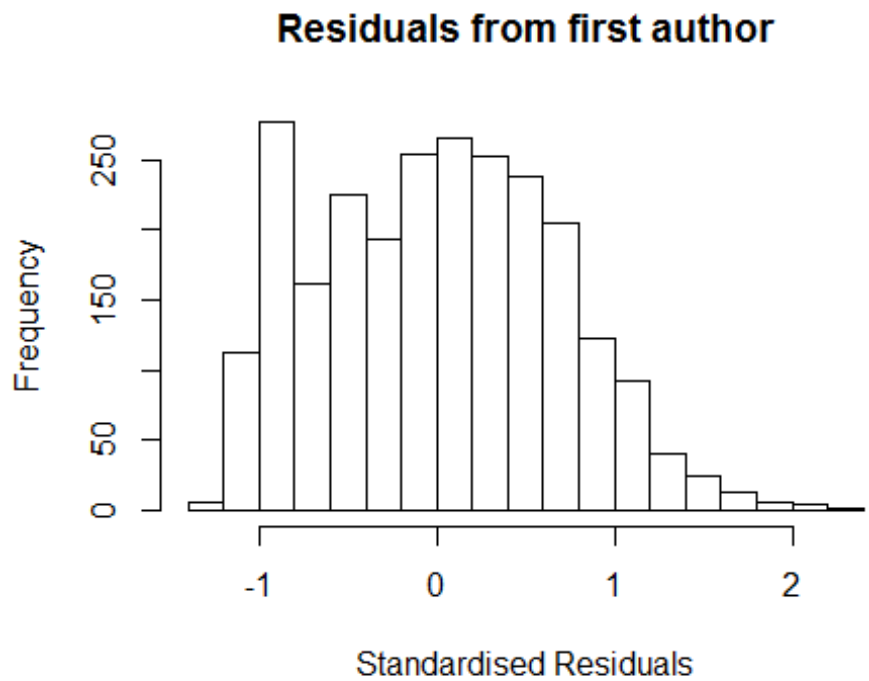

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
UniqueAuthors +
##     Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##     k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.4959 -0.5231 -0.0033  0.4452  2.3653
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)          0.6627     0.0587   11.30  < 2e-16 ***
## FirstAuthorFemale1   0.0584     0.0358    1.63   0.1034
## LastAuthorFemale1    0.0192     0.0367    0.52   0.6005
## UniqueAuthors2       0.3926     0.0317   12.37  < 2e-16 ***
## UniqueAuthors3       0.5536     0.0393   14.09  < 2e-16 ***
## UniqueAuthors4       0.5243     0.0637    8.23  2.9e-16 ***
## UniqueAuthors5      -0.0317     0.0995   -0.32   0.7501
## Year1997            -0.0583     0.0750   -0.78   0.4365
## Year1998            -0.0781     0.0774   -1.01   0.3127
## Year1999             0.0781     0.0799    0.98   0.3284
```

```
## Year2000             0.1678     0.0738    2.27   0.0231 *
## Year2001             0.2637     0.0891    2.96   0.0031 **
## Year2002             0.2797     0.0884    3.16   0.0016 **
## Year2003             0.3283     0.0803    4.09  4.5e-05 ***
## Year2004             0.2373     0.0842    2.82   0.0049 **
## Year2005             0.1328     0.0814    1.63   0.1030
## Year2006             0.0938     0.0781    1.20   0.2297
## Year2007             0.1895     0.0792    2.39   0.0168 *
## Year2008             0.1277     0.0800    1.60   0.1106
## Year2009             0.0588     0.0723    0.81   0.4159
## Year2010            -0.0484     0.0745   -0.65   0.5160
## Year2011            -0.0396     0.0796   -0.50   0.6187
## Year2012            -0.0151     0.0770   -0.20   0.8448
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.652
## Multiple R-squared:  0.147,  Adjusted R-squared:  0.139
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  188 weights are ~= 1. The remaining 2305 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.160   0.883   0.940   0.913   0.985   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          4.01e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.350  1           1.162
## LastAuthorFemale  1.354  1           1.164
## Year              1.042 16           1.001
```



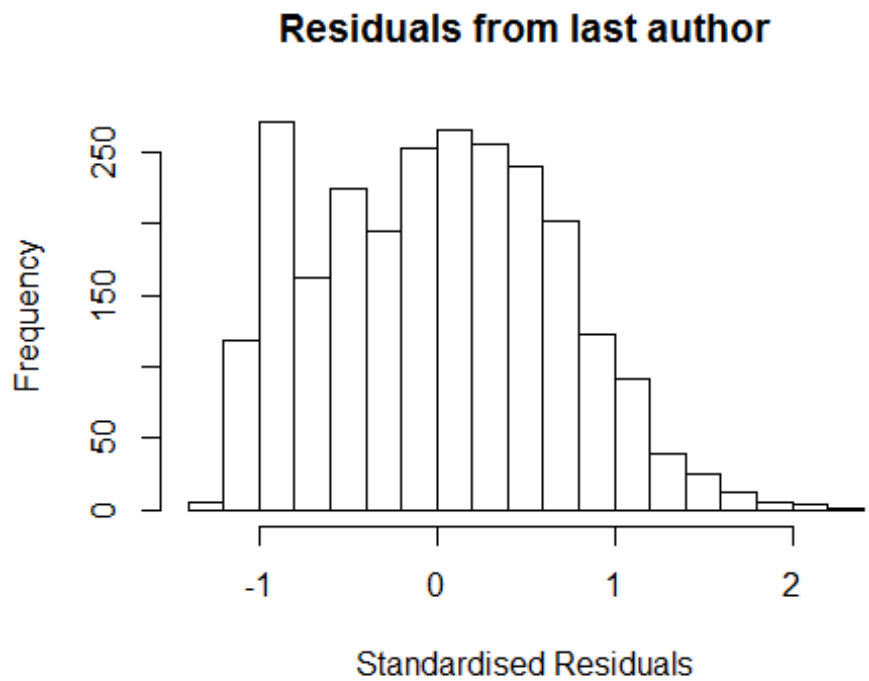

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.2863 -0.5227  0.0197  0.5021  2.2376
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)           0.8528     0.0608   14.02  < 2e-16 ***
## FirstAuthorFemale1    0.0699     0.0389    1.80   0.0723 .
## LastAuthorFemale1     0.0107     0.0397    0.27   0.7874
## Year1997             -0.0894     0.0795   -1.12   0.2609
## Year1998             -0.1055     0.0809   -1.30   0.1921
## Year1999              0.0667     0.0818    0.82   0.4149
## Year2000              0.2013     0.0780    2.58   0.0099 **
## Year2001              0.2849     0.0904    3.15   0.0016 **
## Year2002              0.3533     0.0877    4.03  5.8e-05 ***
## Year2003              0.3528     0.0831    4.24  2.3e-05 ***
## Year2004              0.2603     0.0874    2.98   0.0029 **
## Year2005              0.1866     0.0865    2.16   0.0310 *
```

```
## Year2006             0.1107     0.0836    1.32   0.1858
## Year2007             0.2204     0.0846    2.61   0.0092 **
## Year2008             0.1479     0.0878    1.69   0.0920 .
## Year2009             0.0901     0.0783    1.15   0.2497
## Year2010            -0.0259     0.0796   -0.33   0.7447
## Year2011            -0.0134     0.0848   -0.16   0.8741
## Year2012             0.0522     0.0808    0.65   0.5184
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.726
## Multiple R-squared:  0.0367, Adjusted R-squared:  0.0297
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  212 weights are ~= 1. The remaining 2281 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.321   0.882   0.947   0.920   0.984   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          4.01e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.021  1           1.011
## Year              1.021 16           1.001
```




```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.2827 -0.5188  0.0219  0.5011  2.2360
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)           0.8537     0.0607   14.06  < 2e-16 ***
## FirstAuthorFemale1    0.0760     0.0339    2.24   0.0252 *
## Year1997             -0.0892     0.0795   -1.12   0.2616
## Year1998             -0.1055     0.0809   -1.30   0.1925
## Year1999              0.0664     0.0818    0.81   0.4166
## Year2000              0.2014     0.0779    2.58   0.0098 **
## Year2001              0.2848     0.0904    3.15   0.0016 **
## Year2002              0.3529     0.0877    4.03  5.9e-05 ***
## Year2003              0.3530     0.0832    4.24  2.3e-05 ***
## Year2004              0.2602     0.0873    2.98   0.0029 **
## Year2005              0.1869     0.0864    2.16   0.0307 *
## Year2006              0.1108     0.0836    1.32   0.1853
```

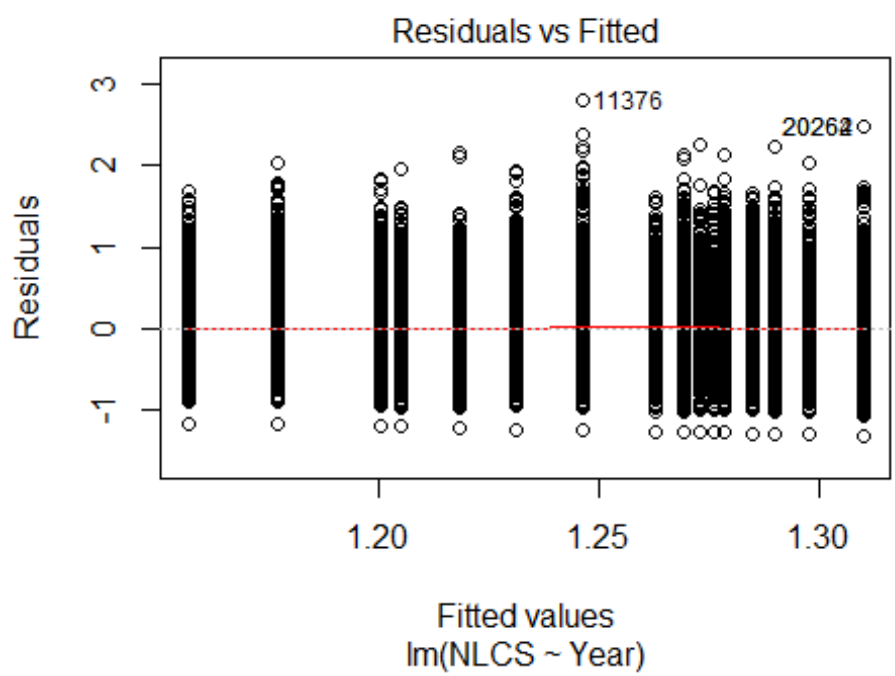
```
## Year2007             0.2212     0.0845    2.62   0.0089 **
## Year2008             0.1484     0.0877    1.69   0.0908 .
## Year2009             0.0898     0.0783    1.15   0.2515
## Year2010            -0.0257     0.0796   -0.32   0.7465
## Year2011            -0.0127     0.0847   -0.15   0.8808
## Year2012             0.0527     0.0808    0.65   0.5142
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.726
## Multiple R-squared:  0.0367, Adjusted R-squared:  0.0301
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  214 weights are ~= 1. The remaining 2279 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.322   0.881   0.947   0.920   0.984   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          4.01e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##                   GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.026  1           1.013
## Year             1.026 16           1.001
```



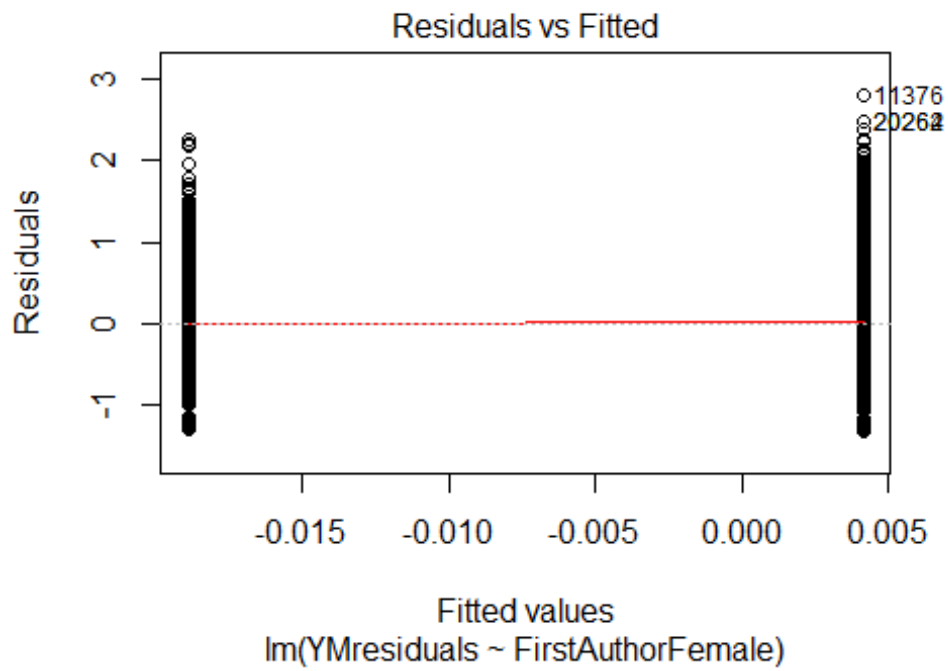

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.2671 -0.5150  0.0209  0.4966  2.2309
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)          0.8586     0.0605   14.18  < 2e-16 ***
## LastAuthorFemale1    0.0512     0.0343    1.49   0.1359
## Year1997            -0.0906     0.0795   -1.14   0.2544
## Year1998            -0.1062     0.0810   -1.31   0.1899
## Year1999             0.0681     0.0819    0.83   0.4061
## Year2000             0.2015     0.0780    2.58   0.0098 **
## Year2001             0.2875     0.0909    3.16   0.0016 **
## Year2002             0.3541     0.0879    4.03  5.8e-05 ***
## Year2003             0.3573     0.0833    4.29  1.8e-05 ***
## Year2004             0.2640     0.0877    3.01   0.0026 **
## Year2005             0.1877     0.0864    2.17   0.0300 *
## Year2006             0.1138     0.0836    1.36   0.1735
```

```
## Year2007            0.2154     0.0845    2.55   0.0109 *
## Year2008            0.1471     0.0881    1.67   0.0950 .
## Year2009            0.0925     0.0783    1.18   0.2373
## Year2010           -0.0236     0.0798   -0.30   0.7670
## Year2011           -0.0125     0.0848   -0.15   0.8824
## Year2012            0.0545     0.0808    0.67   0.5002
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.726
## Multiple R-squared:  0.0356, Adjusted R-squared:  0.029
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  214 weights are ~= 1. The remaining 2279 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.324   0.880   0.947   0.920   0.984   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          4.01e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Sample size for the above analysis:  2493"
## [1] ""
## [1] ""
## [1] "####################################"
## [1] "Analysis of AJSC 1500"
## [1] "####################################"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 2553 2263 2168 2031 2065 1839 1808 1495 1503 1630 1752 1766 1821 2129 2431
## 2011 2012
## 2554 2350
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 1286 1112 1093  983  934  579  860  689  741  801  861  950  989 1207 1439
## 2011 2012
```

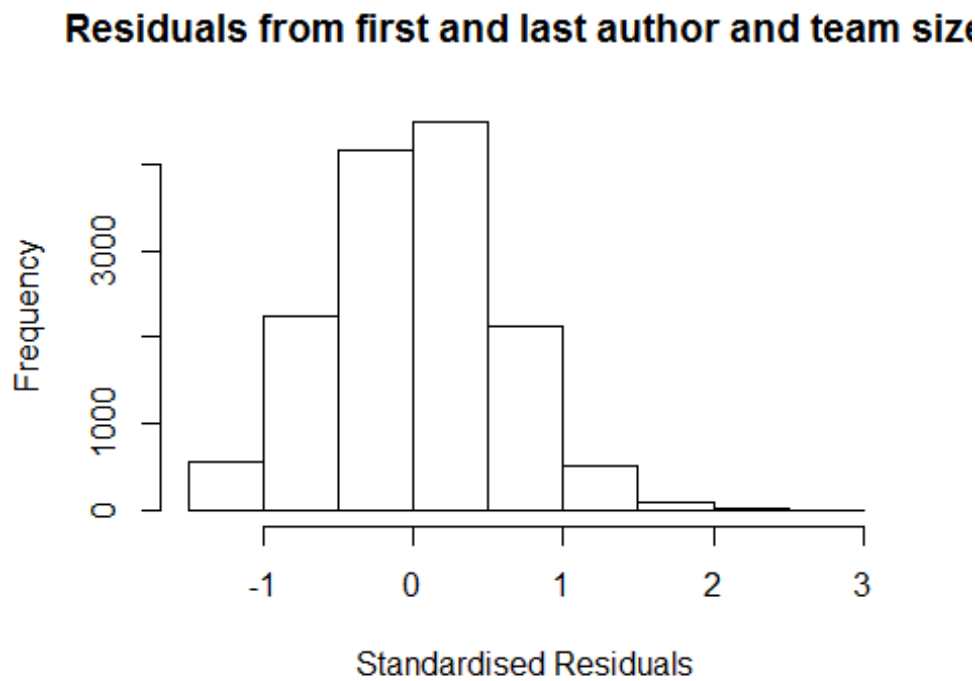
```
## 1487 1368
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 1066  901  918  825  765  473  700  570  593  651  673  769  815 1012 1155
## 2011 2012
## 1196 1101
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
##  Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 190, df = 16, p-value <2e-16
```




```
##
##  Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 20, df = 1, p-value = 9e-06
```



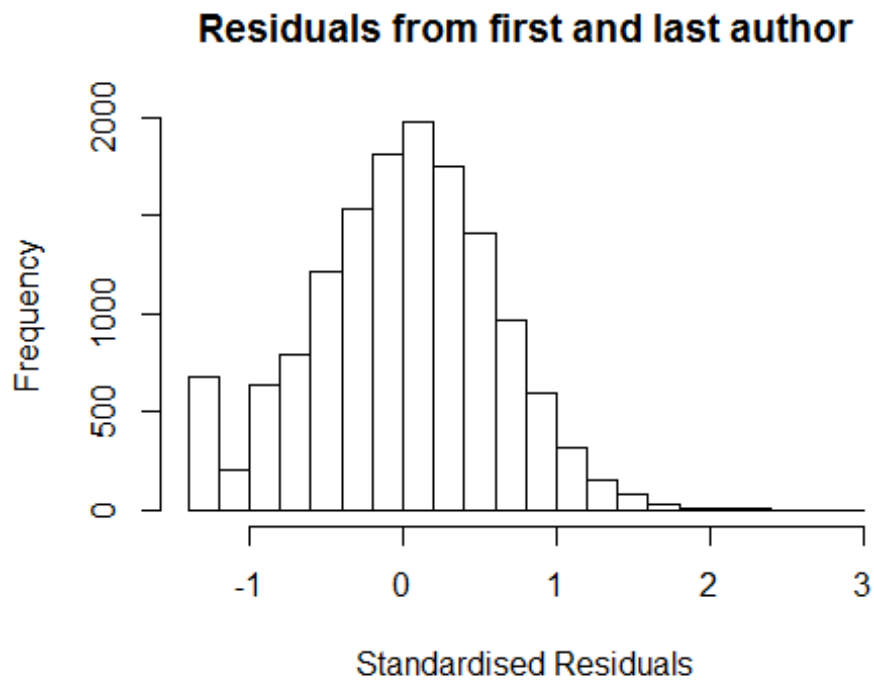

```
## [1] "Female first author team size 2018 geometric mean: 3.97118376671052"
## [1] "Male first author team size 2018 geometric mean: 3.45948677263946"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 160000, p-value = 0.002
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 3.8072417005443"
## [1] "Male last author team size 2018 geometric mean: 3.52512567103847"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 120000, p-value = 0.2
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.047  1           1.023
## LastAuthorFemale  1.035  1           1.018
## UniqueAuthors     1.049  4           1.006
## Year              1.060 16           1.002
```




```
## [1] "List of  5 outliers with residuals above  2.5"
##          ScopusId  NLCS Year OneField Fields residuals
## 11376  0034193710 4.050 2000     1500      3     2.776
## 13435  0035037776 3.537 2001     1500      1     2.572
## 14940  0036591982 3.537 2002     1500      5     2.628
## 20262 18144385855 3.797 2005     1500      3     2.858
## 20264 18144418629 3.800 2005     1500      3     2.861
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
UniqueAuthors +
##     Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##     k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.4967 -0.3902  0.0129  0.3910  2.8608
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         0.94959    0.02443   38.88  < 2e-16 ***
## FirstAuthorFemale1 -0.05363    0.01253   -4.28  1.9e-05 ***
## LastAuthorFemale1  -0.07513    0.01571   -4.78  1.8e-06 ***
## UniqueAuthors2      0.37534    0.01759   21.34  < 2e-16 ***
## UniqueAuthors3      0.40830    0.01821   22.43  < 2e-16 ***
## UniqueAuthors4      0.47465    0.02007   23.65  < 2e-16 ***
```

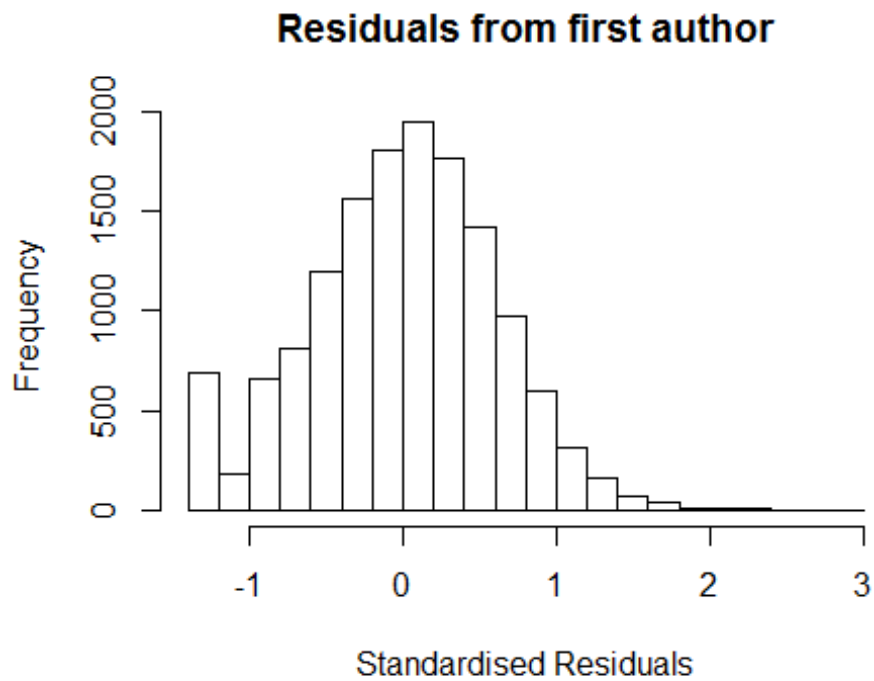
```
## UniqueAuthors5      0.53212    0.01934   27.51  < 2e-16 ***
## Year1997           -0.00975    0.03029   -0.32  0.74760
## Year1998            0.00570    0.03061    0.19  0.85230
## Year1999           -0.02514    0.03120   -0.81  0.42052
## Year2000           -0.05050    0.03290   -1.54  0.12479
## Year2001            0.01498    0.03629    0.41  0.67964
## Year2002           -0.04034    0.03094   -1.30  0.19221
## Year2003           -0.01485    0.03161   -0.47  0.63857
## Year2004           -0.02204    0.03073   -0.72  0.47339
## Year2005           -0.01034    0.03013   -0.34  0.73136
## Year2006           -0.04781    0.03102   -1.54  0.12326
## Year2007           -0.10536    0.02852   -3.69  0.00022 ***
## Year2008           -0.11103    0.02896   -3.83  0.00013 ***
## Year2009           -0.04439    0.02679   -1.66  0.09763 .
## Year2010           -0.10912    0.02631   -4.15  3.4e-05 ***
## Year2011           -0.15507    0.02612   -5.94  3.0e-09 ***
## Year2012           -0.16638    0.02666   -6.24  4.5e-10 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.577
## Multiple R-squared:  0.0798, Adjusted R-squared:  0.0784
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  3 observations c(4276,7146,7147)
##   are outliers with |weight| = 0 ( < 7.1e-06);
##  1202 weights are ~= 1. The remaining 12978 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0031  0.8670  0.9490  0.9050  0.9860  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          7.05e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                    psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##                     GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.023  1           1.012
```

```
## LastAuthorFemale  1.023  1           1.011
## Year              1.024 16           1.001
```



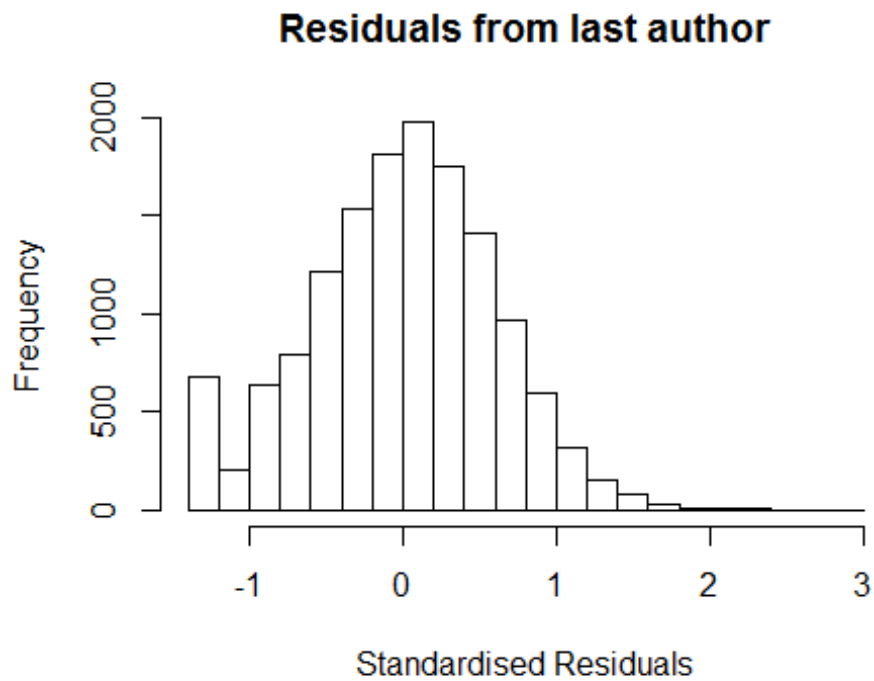

```
## [1] "List of  1 outliers with residuals above  2.5"
##         ScopusId NLCS Year OneField Fields residuals
## 11376 0034193710 4.05 2000     1500      3     2.819
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.3089 -0.3984  0.0211  0.4041  2.8192
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.26829    0.02172   58.39  < 2e-16 ***
## FirstAuthorFemale1 -0.00975    0.01289   -0.76  0.44951
## LastAuthorFemale1  -0.06835    0.01639   -4.17  3.1e-05 ***
## Year1997           -0.01015    0.03158   -0.32  0.74790
## Year1998            0.02279    0.03183    0.72  0.47395
## Year1999           -0.01790    0.03281   -0.55  0.58538
## Year2000           -0.03748    0.03506   -1.07  0.28513
## Year2001            0.01744    0.03760    0.46  0.64271
## Year2002            0.00524    0.03222    0.16  0.87092
```

```
## Year2003            0.03217    0.03254    0.99  0.32276
## Year2004            0.02917    0.03204    0.91  0.36261
## Year2005            0.04060    0.03146    1.29  0.19687
## Year2006            0.01079    0.03214    0.34  0.73710
## Year2007           -0.06107    0.03041   -2.01  0.04460 *
## Year2008           -0.06234    0.03074   -2.03  0.04257 *
## Year2009            0.00397    0.02814    0.14  0.88795
## Year2010           -0.05639    0.02733   -2.06  0.03913 *
## Year2011           -0.08244    0.02701   -3.05  0.00228 **
## Year2012           -0.10312    0.02750   -3.75  0.00018 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.593
## Multiple R-squared:  0.0069, Adjusted R-squared:  0.00564
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  observation 4276 is an outlier with |weight| = 0 ( < 7.1e-06);
##  1195 weights are ~= 1. The remaining 12987 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0384  0.8650  0.9500  0.9030  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          7.05e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.013  1           1.007
## Year              1.013 16           1.000
```




```
## [1] "List of  1 outliers with residuals above  2.5"
##         ScopusId NLCS Year OneField Fields residuals
## 11376 0034193710 4.05 2000     1500      3     2.819
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.3035 -0.3972  0.0217  0.4050  2.8241
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.26399    0.02168   58.29  < 2e-16 ***
## FirstAuthorFemale1 -0.01900    0.01295   -1.47   0.1423
## Year1997           -0.01088    0.03162   -0.34   0.7308
## Year1998            0.02317    0.03181    0.73   0.4664
## Year1999           -0.01911    0.03281   -0.58   0.5602
## Year2000           -0.03806    0.03501   -1.09   0.2770
## Year2001            0.01636    0.03758    0.44   0.6633
## Year2002            0.00481    0.03222    0.15   0.8813
## Year2003            0.02965    0.03254    0.91   0.3621
## Year2004            0.02762    0.03208    0.86   0.3893
## Year2005            0.03952    0.03146    1.26   0.2090
## Year2006            0.00835    0.03212    0.26   0.7948
```

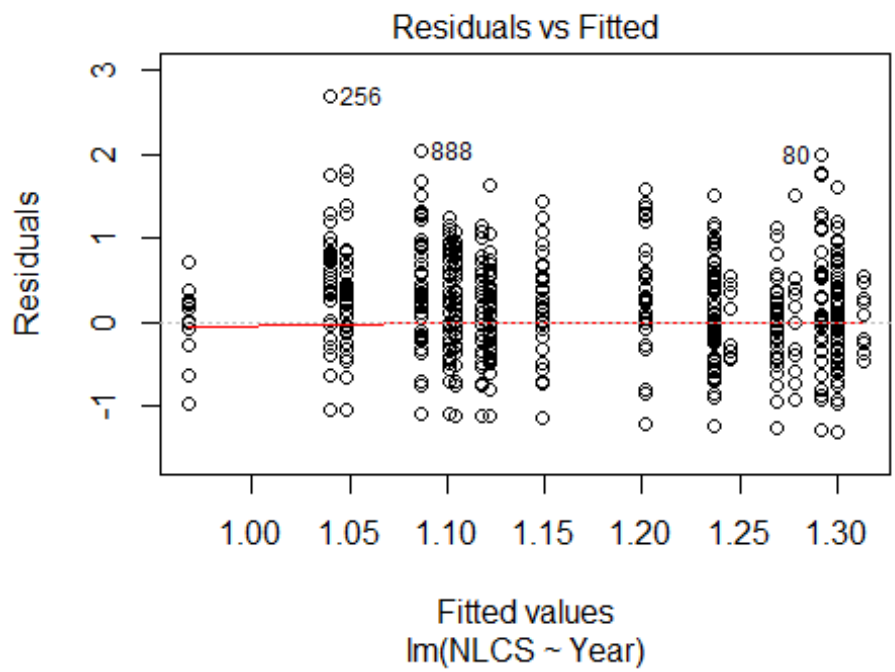
```
## Year2007           -0.06345    0.03042   -2.09   0.0370 *
## Year2008           -0.06471    0.03079   -2.10   0.0356 *
## Year2009            0.00268    0.02817    0.10   0.9242
## Year2010           -0.06096    0.02731   -2.23   0.0256 *
## Year2011           -0.08562    0.02702   -3.17   0.0015 **
## Year2012           -0.10738    0.02752   -3.90  9.6e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.594
## Multiple R-squared:  0.00565,    Adjusted R-squared:  0.00446
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  observation 4276 is an outlier with |weight| = 0 ( < 7.1e-06);
##  1201 weights are ~= 1. The remaining 12981 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0377  0.8650  0.9500  0.9030  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          7.05e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##                   GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.013  1           1.006
## Year             1.013 16           1.000
```



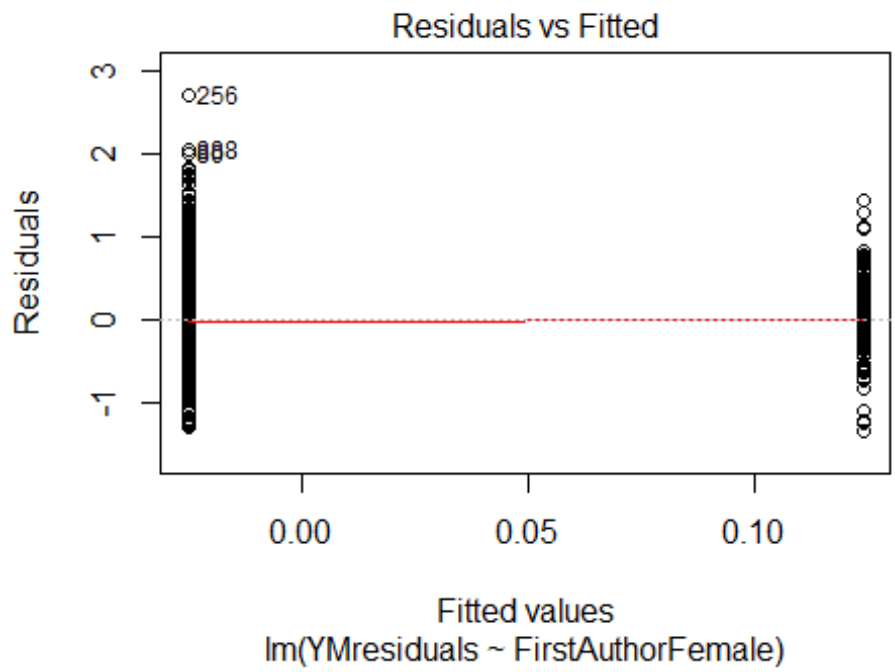

```
## [1] "List of  1 outliers with residuals above  2.5"
##         ScopusId NLCS Year OneField Fields residuals
## 11376 0034193710 4.05 2000     1500      3     2.819
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.3074 -0.3980  0.0203  0.4040  2.8203
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.26700    0.02163   58.57  < 2e-16 ***
## LastAuthorFemale1 -0.07034    0.01637   -4.30  1.7e-05 ***
## Year1997          -0.01018    0.03157   -0.32  0.74716
## Year1998           0.02266    0.03181    0.71  0.47637
## Year1999          -0.01797    0.03280   -0.55  0.58388
## Year2000          -0.03726    0.03506   -1.06  0.28783
## Year2001           0.01748    0.03758    0.47  0.64189
## Year2002           0.00516    0.03222    0.16  0.87286
## Year2003           0.03204    0.03253    0.99  0.32460
## Year2004           0.02899    0.03203    0.91  0.36541
## Year2005           0.04044    0.03144    1.29  0.19837
## Year2006           0.01060    0.03214    0.33  0.74145
```

```
## Year2007          -0.06128    0.03039   -2.02  0.04380 *
## Year2008          -0.06311    0.03070   -2.06  0.03984 *
## Year2009           0.00336    0.02811    0.12  0.90494
## Year2010          -0.05711    0.02730   -2.09  0.03644 *
## Year2011          -0.08329    0.02697   -3.09  0.00202 **
## Year2012          -0.10391    0.02747   -3.78  0.00016 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.593
## Multiple R-squared:  0.00686,    Adjusted R-squared:  0.00566
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  observation 4276 is an outlier with |weight| = 0 ( < 7.1e-06);
##  1193 weights are ~= 1. The remaining 12989 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0383  0.8650  0.9500  0.9040  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          7.05e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                    psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Sample size for the above analysis:  14183"
## [1] ""
## [1] ""
## [1] "####################################"
## [1] "Analysis of AJSC 1501"
## [1] "####################################"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##   93  117   99   91   96  117   96   81  125  124  125  100   67   18   22
## 2011 2012
##   25   17
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##   44   56   51   43   45   45   49   45   72   75   77   56   38    8   11
```

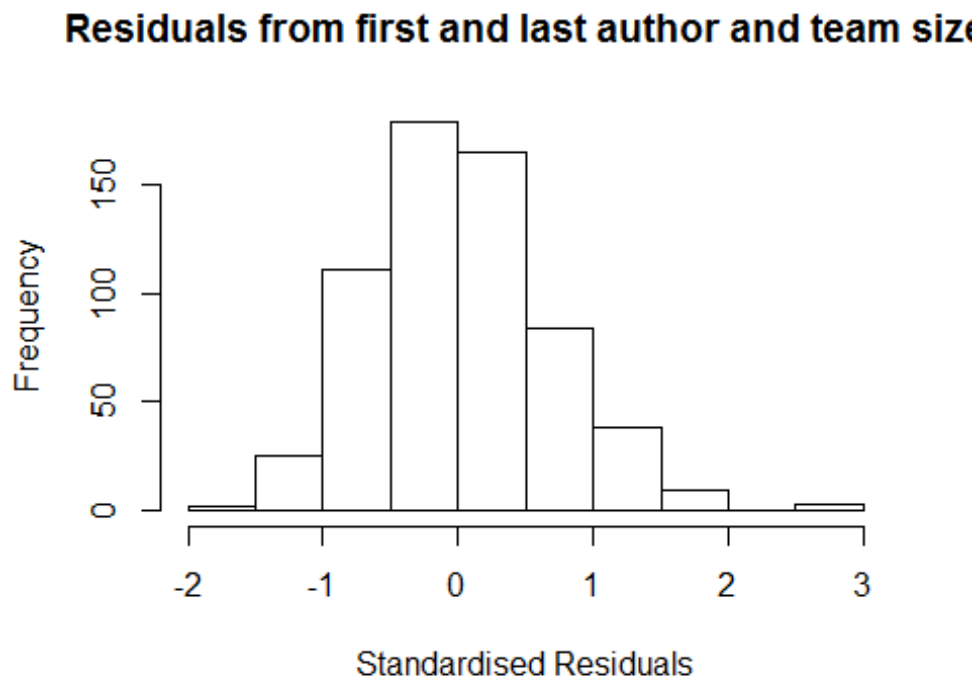
```
## 2011 2012
##   13   13
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##   35   46   43   40   35   40   41   39   61   62   63   50   28    6    7
## 2011 2012
##    9   11
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
##  Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 55, df = 16, p-value = 4e-06
```




```
##
##  Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 17, df = 1, p-value = 4e-05
```

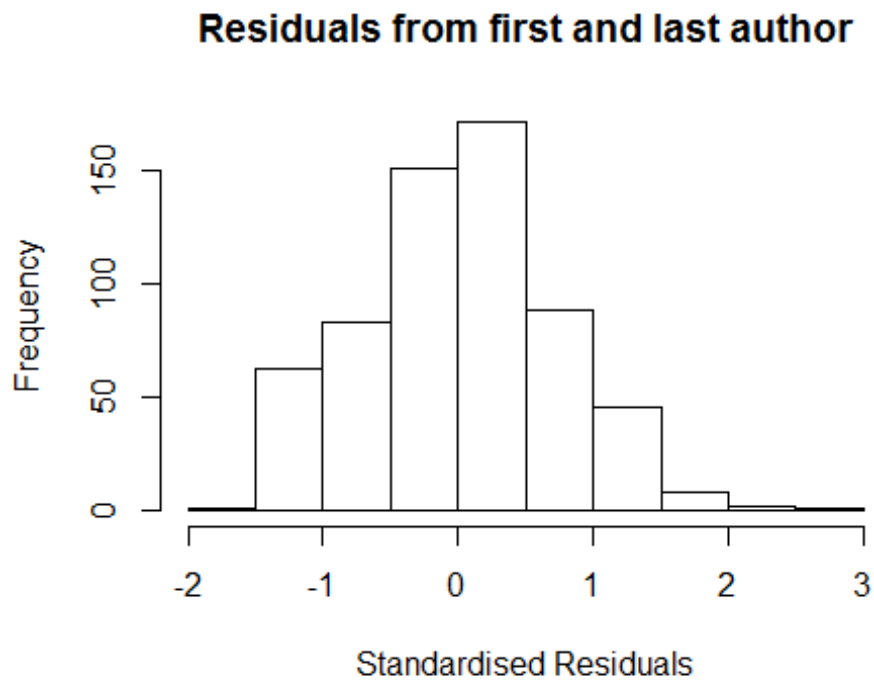



```
## [1] "Female first author team size 2018 geometric mean: 2.59910094796657"
## [1] "Male first author team size 2018 geometric mean: 3.52487600789573"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 470, p-value = 0.1
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 2.18515221382051"
## [1] "Male last author team size 2018 geometric mean: 3.654838353021"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 320, p-value = 0.006
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##                     GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.311  1           1.145
## LastAuthorFemale  1.279  1           1.131
## UniqueAuthors     2.365  4           1.114
## Year              3.031 16           1.035
```



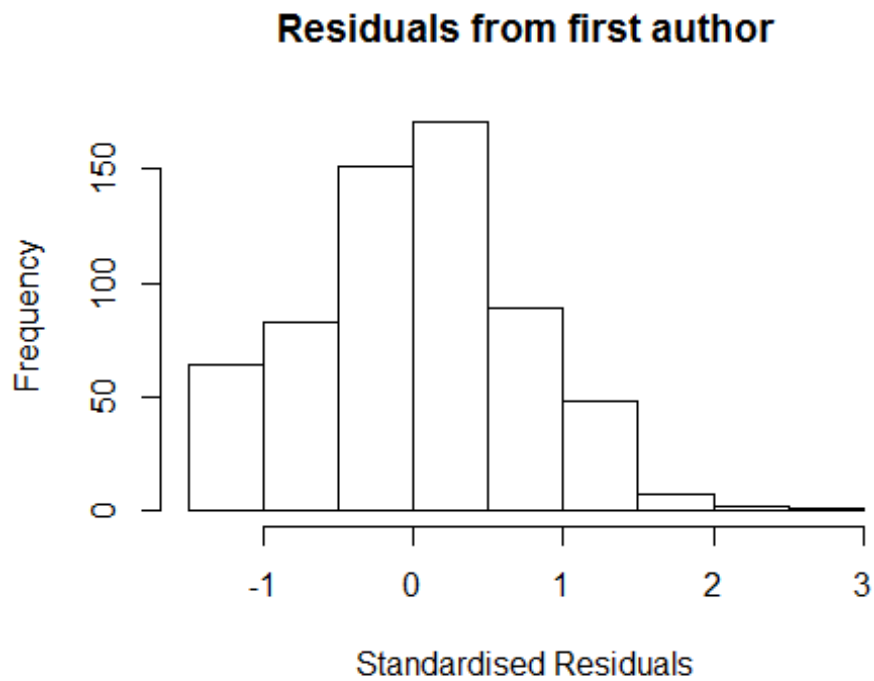

```
## [1] "List of  3 outliers with residuals above  2.5"
##       ScopusId  NLCS Year OneField Fields residuals
## 80  0030484769 3.279 1996     1501      2     2.616
## 256 0032092047 3.725 1998     1501      2     2.646
## 888 4143111238 3.121 2004     1501      1     2.538
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
UniqueAuthors +
##     Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##     k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.5757 -0.4682 -0.0245  0.4543  2.6458
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)           0.6631     0.1383    4.79  2.1e-06 ***
## FirstAuthorFemale1    0.1423     0.0700    2.03   0.0426 *
## LastAuthorFemale1     0.1427     0.0983    1.45   0.1469
## UniqueAuthors2        0.5972     0.0877    6.81  2.5e-11 ***
## UniqueAuthors3        0.6996     0.0920    7.61  1.1e-13 ***
## UniqueAuthors4        0.8495     0.0883    9.62  < 2e-16 ***
## UniqueAuthors5        0.8939     0.1260    7.09  3.8e-12 ***
## Year1997             -0.2761     0.1602   -1.72   0.0854 .
```

```
## Year1998            -0.1811     0.1673   -1.08   0.2794
## Year1999            -0.1346     0.1585   -0.85   0.3959
## Year2000            -0.0779     0.1664   -0.47   0.6398
## Year2001            -0.1108     0.1595   -0.69   0.4875
## Year2002            -0.0130     0.1625   -0.08   0.9363
## Year2003             0.0694     0.2042    0.34   0.7342
## Year2004            -0.0796     0.1638   -0.49   0.6272
## Year2005            -0.0807     0.1458   -0.55   0.5801
## Year2006            -0.0600     0.1552   -0.39   0.6992
## Year2007            -0.2797     0.1556   -1.80   0.0728 .
## Year2008            -0.0416     0.1714   -0.24   0.8083
## Year2009            -0.3362     0.2126   -1.58   0.1143
## Year2010            -0.1248     0.2496   -0.50   0.6172
## Year2011            -0.5326     0.1829   -2.91   0.0037 **
## Year2012            -0.1043     0.1745   -0.60   0.5505
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.607
## Multiple R-squared:  0.222,  Adjusted R-squared:  0.193
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
##  42 weights are ~= 1. The remaining 574 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0177  0.8730  0.9430  0.8970  0.9840  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.62e-04          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##                     GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.215  1           1.102
## LastAuthorFemale  1.261  1           1.123
## Year              1.431 16           1.011
```



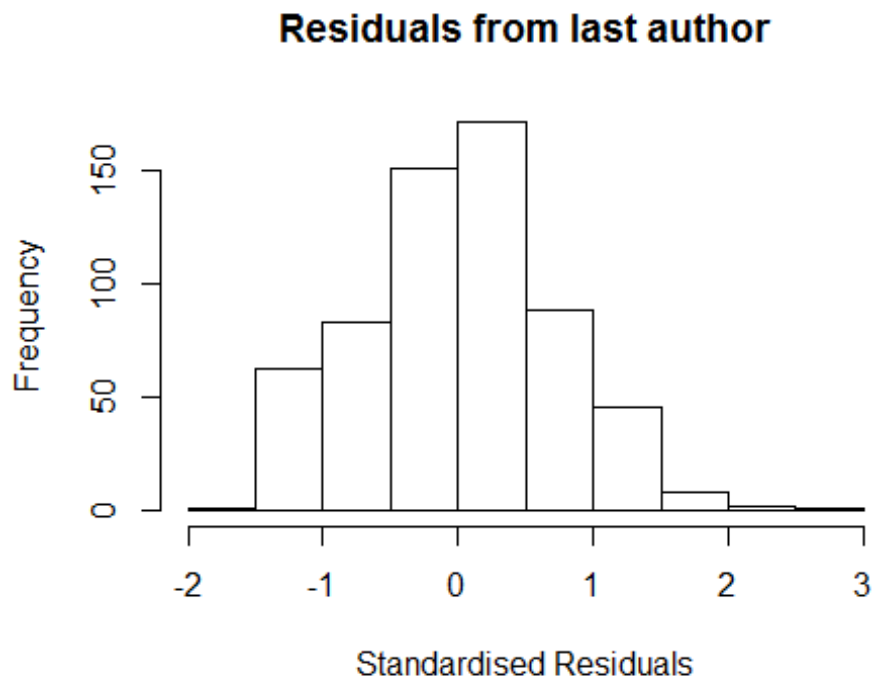

```
## [1] "List of  1 outliers with residuals above  2.5"
##       ScopusId  NLCS Year OneField Fields residuals
## 256 0032092047 3.725 1998     1501      2     2.827
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.5166 -0.4780  0.0361  0.4726  2.8274
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)           1.1393     0.1697    6.71  4.5e-11 ***
## FirstAuthorFemale1    0.2532     0.0701    3.61  0.00033 ***
## LastAuthorFemale1     0.0940     0.1004    0.94  0.34937
## Year1997             -0.2849     0.1985   -1.43  0.15183
## Year1998             -0.2417     0.2243   -1.08  0.28175
## Year1999             -0.0465     0.2126   -0.22  0.82679
## Year2000             -0.0123     0.2053   -0.06  0.95208
## Year2001             -0.0882     0.2008   -0.44  0.66053
## Year2002             -0.0787     0.1991   -0.40  0.69289
## Year2003              0.0301     0.2398    0.13  0.90024
## Year2004             -0.0991     0.2106   -0.47  0.63826
## Year2005              0.0332     0.1828    0.18  0.85606
```

```
## Year2006             0.0227     0.1893    0.12  0.90451
## Year2007            -0.1020     0.1827   -0.56  0.57686
## Year2008             0.0883     0.1928    0.46  0.64727
## Year2009            -0.1164     0.2237   -0.52  0.60318
## Year2010             0.0996     0.3467    0.29  0.77398
## Year2011            -0.3683     0.2178   -1.69  0.09131 .
## Year2012             0.1678     0.2094    0.80  0.42326
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.717
## Multiple R-squared:  0.0399, Adjusted R-squared:  0.011
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
##  49 weights are ~= 1. The remaining 567 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0854  0.8620  0.9510  0.9090  0.9870  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.62e-04          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.185  1           1.089
## Year              1.185 16           1.005
```




```
## [1] "List of  1 outliers with residuals above  2.5"
##       ScopusId  NLCS Year OneField Fields residuals
## 256 0032092047 3.725 1998     1501      2     2.827
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##    Min     1Q Median     3Q    Max
## -1.452 -0.483  0.045  0.474  2.816
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)           1.1450     0.1686    6.79  2.7e-11 ***
## FirstAuthorFemale1    0.2702     0.0694    3.89  0.00011 ***
## Year1997             -0.2729     0.1985   -1.37  0.16987
## Year1998             -0.2365     0.2254   -1.05  0.29451
## Year1999             -0.0462     0.2121   -0.22  0.82773
## Year2000             -0.0126     0.2043   -0.06  0.95068
## Year2001             -0.0948     0.1999   -0.47  0.63545
## Year2002             -0.0782     0.1979   -0.40  0.69278
## Year2003              0.0372     0.2395    0.16  0.87674
## Year2004             -0.1045     0.2094   -0.50  0.61812
## Year2005              0.0317     0.1826    0.17  0.86242
## Year2006              0.0184     0.1883    0.10  0.92209
```

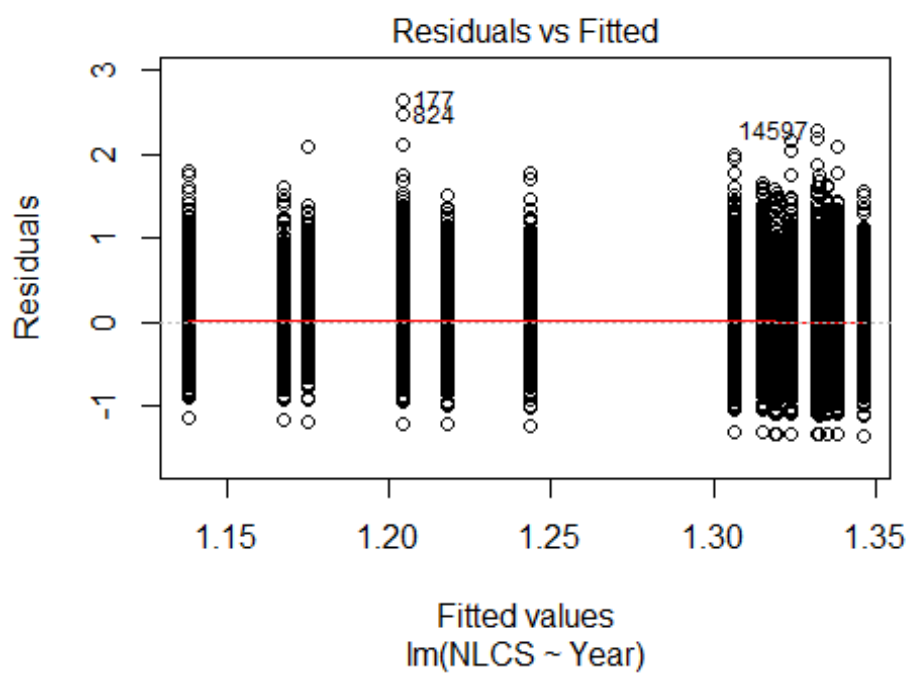
```
## Year2007            -0.0959     0.1827   -0.52  0.59991
## Year2008             0.0868     0.1927    0.45  0.65269
## Year2009            -0.1121     0.2183   -0.51  0.60784
## Year2010             0.1397     0.3427    0.41  0.68371
## Year2011            -0.3537     0.2171   -1.63  0.10378
## Year2012             0.1674     0.2064    0.81  0.41772
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.721
## Multiple R-squared:  0.0382, Adjusted R-squared:  0.0108
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
##  50 weights are ~= 1. The remaining 566 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0934  0.8610  0.9500  0.9100  0.9870  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.62e-04          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##                  GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.21  1           1.100
## Year             1.21 16           1.006
```



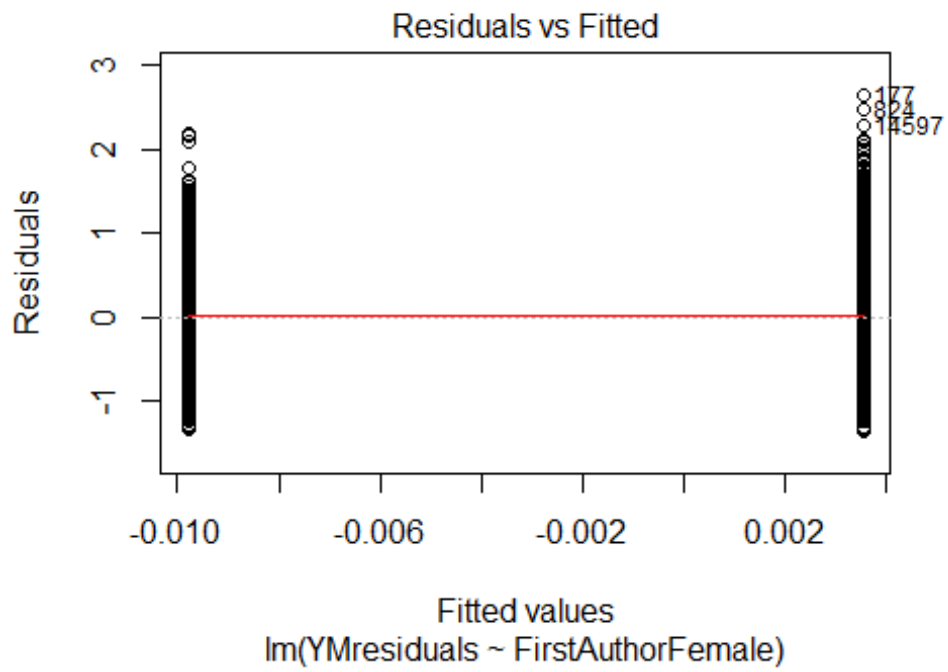

```
## [1] "List of  1 outliers with residuals above  2.5"
##       ScopusId  NLCS Year OneField Fields residuals
## 256 0032092047 3.725 1998     1501      2     2.827
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.3909 -0.5053  0.0314  0.5110  2.7993
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.18748    0.17098    6.95  9.9e-12 ***
## LastAuthorFemale1  0.16622    0.09883    1.68    0.093 .
## Year1997           -0.28619    0.20355   -1.41    0.160
## Year1998           -0.26173    0.23030   -1.14    0.256
## Year1999           -0.09009    0.21074   -0.43    0.669
## Year2000           -0.04769    0.20880   -0.23    0.819
## Year2001           -0.09217    0.20683   -0.45    0.656
## Year2002           -0.08827    0.20548   -0.43    0.668
## Year2003           -0.00946    0.24024   -0.04    0.969
## Year2004           -0.11549    0.21351   -0.54    0.589
## Year2005            0.01017    0.18552    0.05    0.956
## Year2006            0.03718    0.19151    0.19    0.846
```

```
## Year2007          -0.09848    0.18621   -0.53    0.597
## Year2008           0.08205    0.19271    0.43    0.670
## Year2009          -0.09470    0.23571   -0.40    0.688
## Year2010           0.06166    0.36705    0.17    0.867
## Year2011          -0.33290    0.21804   -1.53    0.127
## Year2012           0.13262    0.20527    0.65    0.518
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.731
## Multiple R-squared:  0.0247, Adjusted R-squared:  -0.00305
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
##  49 weights are ~= 1. The remaining 567 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.110   0.859   0.950   0.911   0.987   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.62e-04          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Sample size for the above analysis:  616"
## [1] ""
## [1] ""
## [1] "####################################"
## [1] "Analysis of AJSC 1502"
## [1] "####################################"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  817  767  715  690  825  699  923  791 1008 1243 1372 1721 2118 2331 2587
## 2011 2012
## 2701 2668
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  500  467  416  398  410  323  499  421  507  708  774  989 1190 1344 1491
## 2011 2012
```

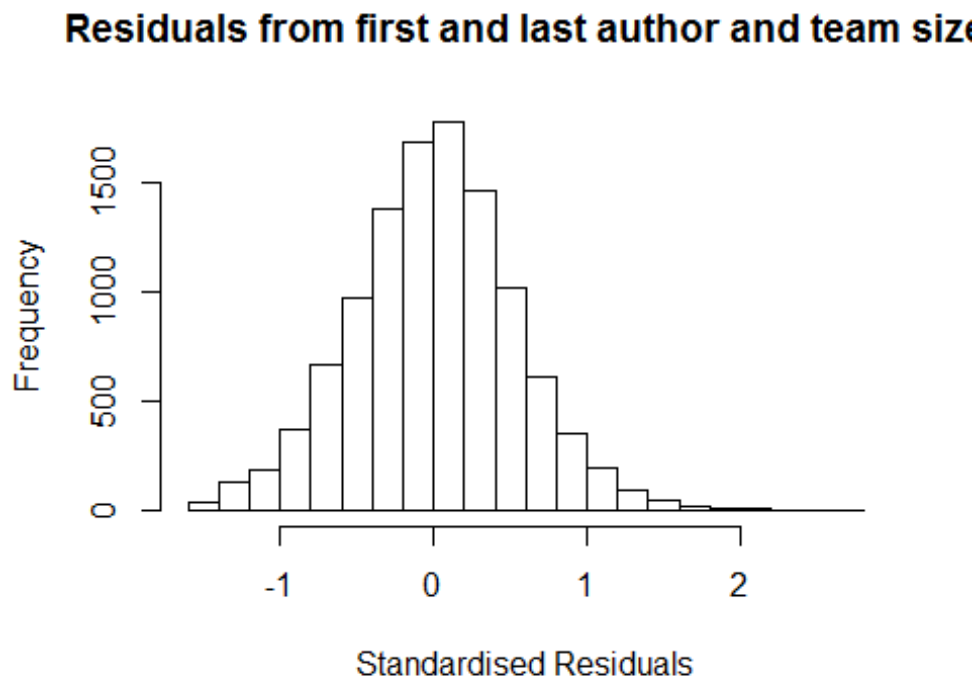
```
## 1577 1475
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  438  424  363  338  346  277  429  351  417  569  613  787  957 1046 1196
## 2011 2012
## 1251 1181
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
##  Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 190, df = 16, p-value <2e-16
```




```
##
##  Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 29, df = 1, p-value = 6e-08
```

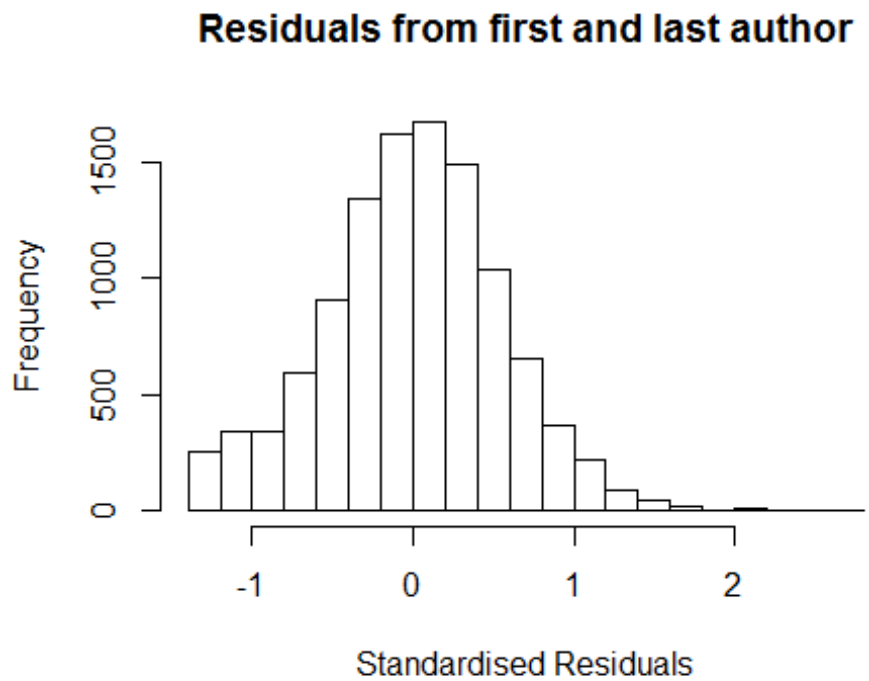



```
## [1] "Female first author team size 2018 geometric mean: 4.91452791819774"
## [1] "Male first author team size 2018 geometric mean: 4.58576484188178"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 1e+05, p-value = 0.2
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 4.37174053252667"
## [1] "Male last author team size 2018 geometric mean: 4.75141949445826"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 70000, p-value = 0.05
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.029  1           1.014
## LastAuthorFemale  1.025  1           1.012
## UniqueAuthors     1.124  4           1.015
## Year              1.134 16           1.004
```



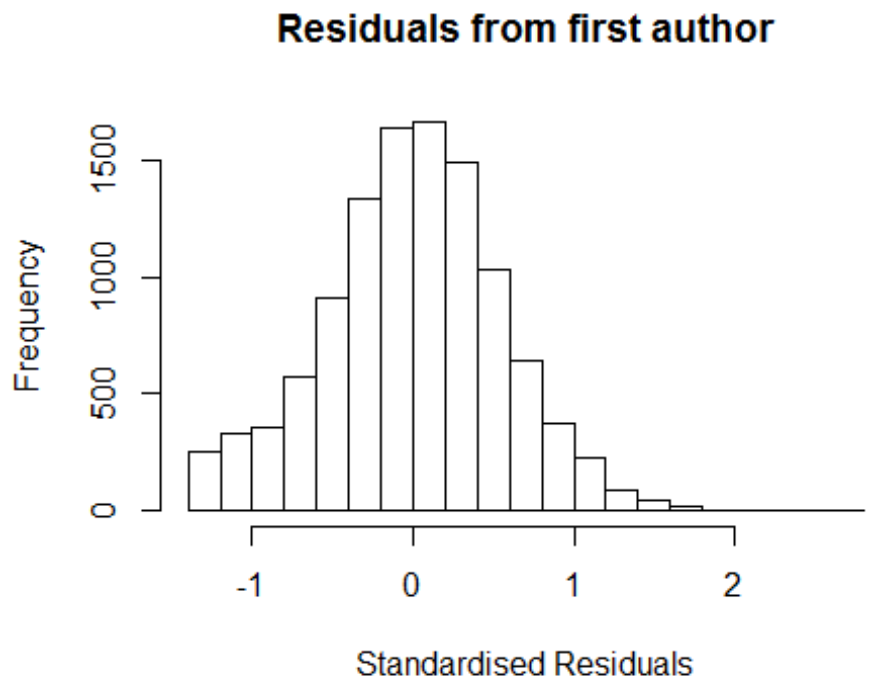

```
## [1] "List of  1 outliers with residuals above  2.5"
##       ScopusId  NLCS Year OneField Fields residuals
## 177 0030232761 3.847 1996     1502      6     2.655
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
UniqueAuthors +
##     Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##     k.max = 1000))
##  \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.46295 -0.33693  0.00913  0.33350  2.65499
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         0.77593    0.03578   21.68   <2e-16 ***
## FirstAuthorFemale1 -0.02682    0.01121   -2.39    0.017 *
## LastAuthorFemale1  -0.01626    0.01379   -1.18    0.239
## UniqueAuthors2      0.41609    0.02490   16.71   <2e-16 ***
## UniqueAuthors3      0.49030    0.02417   20.28   <2e-16 ***
## UniqueAuthors4      0.55089    0.02458   22.41   <2e-16 ***
## UniqueAuthors5      0.63474    0.02301   27.58   <2e-16 ***
## Year1997            0.02333    0.04992    0.47    0.640
## Year1998           -0.00959    0.04479   -0.21    0.830
## Year1999           -0.05665    0.04460   -1.27    0.204
```

```
## Year2000           -0.02532    0.04731   -0.54    0.592
## Year2001            0.04561    0.04992    0.91    0.361
## Year2002           -0.00901    0.04294   -0.21    0.834
## Year2003            0.01658    0.04222    0.39    0.695
## Year2004            0.02650    0.04098    0.65    0.518
## Year2005            0.02597    0.03795    0.68    0.494
## Year2006            0.04164    0.03734    1.12    0.265
## Year2007            0.05229    0.03727    1.40    0.161
## Year2008            0.02816    0.03586    0.79    0.432
## Year2009            0.02928    0.03521    0.83    0.406
## Year2010            0.00546    0.03507    0.16    0.876
## Year2011           -0.01334    0.03499   -0.38    0.703
## Year2012           -0.00388    0.03484   -0.11    0.911
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.502
## Multiple R-squared:  0.11,   Adjusted R-squared:  0.109
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
##  2 observations c(69,397) are outliers with |weight| = 0 ( < 9.1e-06);
##  952 weights are ~= 1. The remaining 10029 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.016   0.865   0.951   0.899   0.986   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          9.10e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.015  1           1.007
## LastAuthorFemale  1.012  1           1.006
## Year              1.018 16           1.001
```



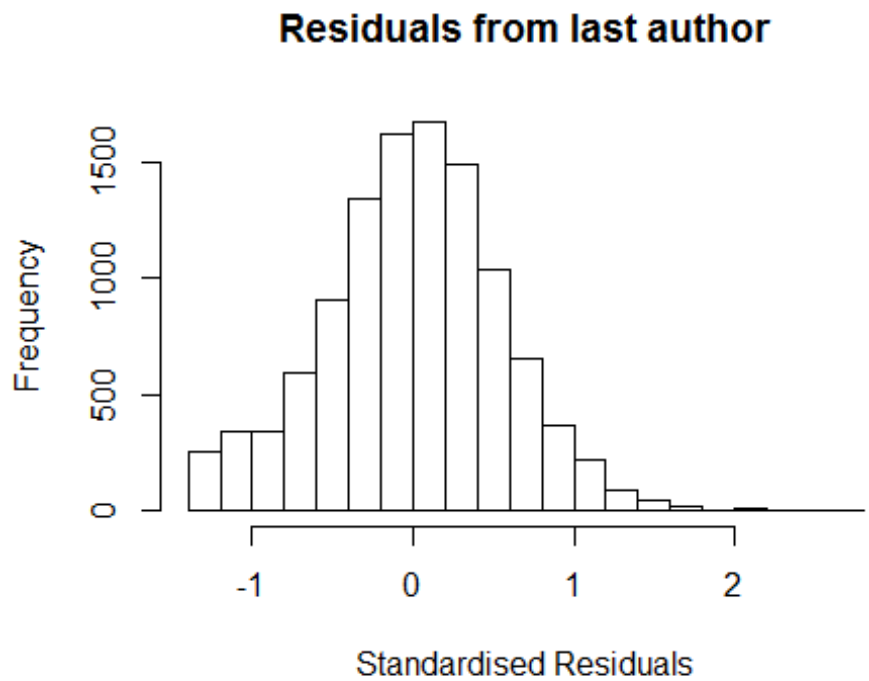

```
## [1] "List of  2 outliers with residuals above  2.5"
##       ScopusId  NLCS Year OneField Fields residuals
## 177 0030232761 3.847 1996     1502      6     2.683
## 824 0029670262 3.675 1996     1305      6     2.511
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.3293 -0.3445  0.0108  0.3493  2.6831
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)          1.16389    0.03562   32.67  < 2e-16 ***
## FirstAuthorFemale1 -0.00504    0.01164   -0.43  0.66500
## LastAuthorFemale1  -0.03467    0.01469   -2.36  0.01826 *
## Year1997           -0.00948    0.05352   -0.18  0.85938
## Year1998            0.02930    0.04781    0.61  0.54002
## Year1999           -0.00110    0.04899   -0.02  0.98204
## Year2000            0.02807    0.05070    0.55  0.57981
## Year2001            0.08601    0.05315    1.62  0.10563
## Year2002            0.08031    0.04656    1.73  0.08454 .
## Year2003            0.13793    0.04609    2.99  0.00277 **
## Year2004            0.14916    0.04475    3.33  0.00086 ***
```

```
## Year2005            0.15723    0.04133    3.80  0.00014 ***
## Year2006            0.16209    0.04108    3.95  8.0e-05 ***
## Year2007            0.16325    0.04065    4.02  6.0e-05 ***
## Year2008            0.16148    0.03961    4.08  4.6e-05 ***
## Year2009            0.16537    0.03875    4.27  2.0e-05 ***
## Year2010            0.15732    0.03857    4.08  4.6e-05 ***
## Year2011            0.12838    0.03881    3.31  0.00094 ***
## Year2012            0.14479    0.03857    3.75  0.00017 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.52
## Multiple R-squared:  0.0109, Adjusted R-squared:  0.00931
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
##  2 observations c(69,397) are outliers with |weight| = 0 ( < 9.1e-06);
##  929 weights are ~= 1. The remaining 10052 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0358  0.8610  0.9510  0.8970  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          9.10e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.011  1           1.006
## Year              1.011 16           1.000
```




```
## [1] "List of  2 outliers with residuals above  2.5"
##       ScopusId  NLCS Year OneField Fields residuals
## 177 0030232761 3.847 1996     1502      6     2.683
## 824 0029670262 3.675 1996     1305      6     2.511
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.3249 -0.3432  0.0112  0.3483  2.6865
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.16046    0.03565   32.55  < 2e-16 ***
## FirstAuthorFemale1 -0.00902    0.01167   -0.77  0.43988
## Year1997           -0.00941    0.05363   -0.18  0.86070
## Year1998            0.03014    0.04784    0.63  0.52876
## Year1999           -0.00141    0.04904   -0.03  0.97707
## Year2000            0.02779    0.05069    0.55  0.58345
## Year2001            0.08428    0.05312    1.59  0.11261
## Year2002            0.07931    0.04657    1.70  0.08859 .
## Year2003            0.13760    0.04614    2.98  0.00287 **
## Year2004            0.14775    0.04477    3.30  0.00097 ***
## Year2005            0.15628    0.04136    3.78  0.00016 ***
```

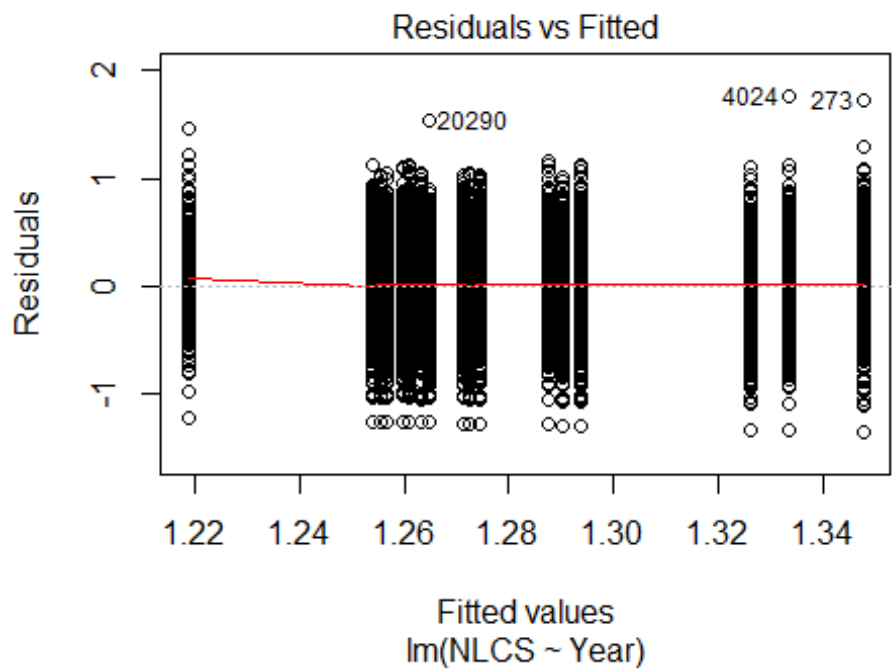
```
## Year2006            0.16038    0.04111    3.90  9.6e-05 ***
## Year2007            0.16235    0.04070    3.99  6.7e-05 ***
## Year2008            0.16056    0.03965    4.05  5.2e-05 ***
## Year2009            0.16447    0.03880    4.24  2.3e-05 ***
## Year2010            0.15618    0.03862    4.04  5.3e-05 ***
## Year2011            0.12728    0.03886    3.28  0.00106 **
## Year2012            0.14248    0.03860    3.69  0.00022 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.52
## Multiple R-squared:  0.0104, Adjusted R-squared:  0.00886
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
##  2 observations c(69,397) are outliers with |weight| = 0 ( < 9.1e-06);
##  920 weights are ~= 1. The remaining 10061 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0334  0.8610  0.9510  0.8970  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          9.10e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                  psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##                   GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.008  1           1.004
## Year             1.008 16           1.000
```



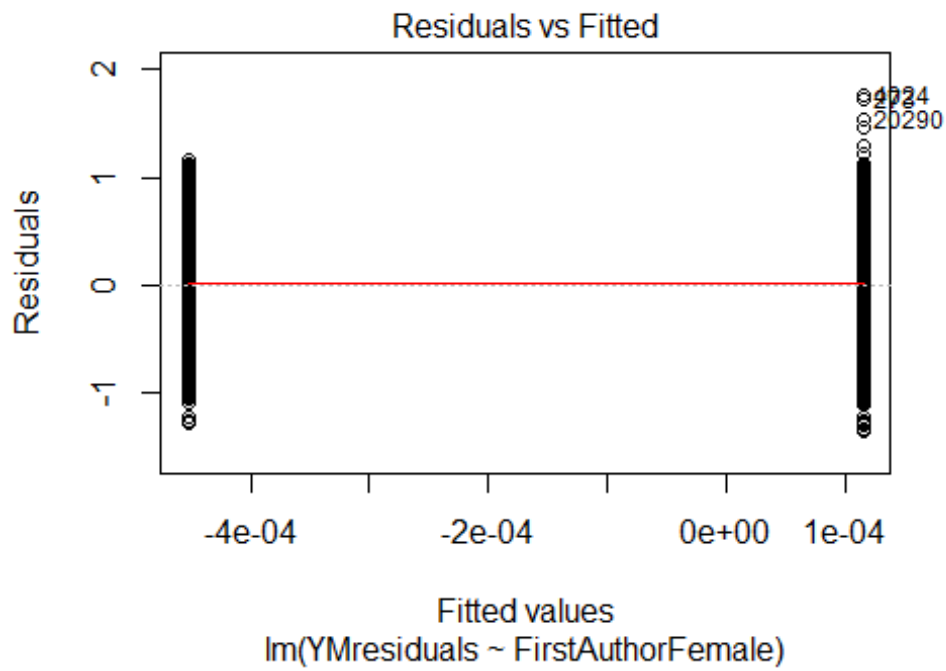

```
## [1] "List of  2 outliers with residuals above  2.5"
##       ScopusId  NLCS Year OneField Fields residuals
## 177 0030232761 3.847 1996     1502      6     2.683
## 824 0029670262 3.675 1996     1305      6     2.511
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.3279 -0.3440  0.0113  0.3488  2.6840
##
## Coefficients:
##                   Estimate Std. Error t value Pr(>|t|)
## (Intercept)        1.16299    0.03554   32.72  < 2e-16 ***
## LastAuthorFemale1 -0.03553    0.01471   -2.42  0.01571 *
## Year1997          -0.00937    0.05352   -0.18  0.86101
## Year1998           0.02920    0.04780    0.61  0.54122
## Year1999          -0.00119    0.04897   -0.02  0.98056
## Year2000           0.02781    0.05068    0.55  0.58323
## Year2001           0.08579    0.05314    1.61  0.10644
## Year2002           0.07991    0.04653    1.72  0.08597 .
## Year2003           0.13767    0.04607    2.99  0.00281 **
## Year2004           0.14866    0.04471    3.32  0.00089 ***
## Year2005           0.15689    0.04130    3.80  0.00015 ***
```

```
## Year2006           0.16176    0.04106    3.94  8.2e-05 ***
## Year2007           0.16300    0.04063    4.01  6.1e-05 ***
## Year2008           0.16110    0.03958    4.07  4.7e-05 ***
## Year2009           0.16494    0.03871    4.26  2.1e-05 ***
## Year2010           0.15696    0.03854    4.07  4.7e-05 ***
## Year2011           0.12788    0.03876    3.30  0.00097 ***
## Year2012           0.14426    0.03852    3.75  0.00018 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.52
## Multiple R-squared:  0.0109, Adjusted R-squared:  0.00938
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
##  2 observations c(69,397) are outliers with |weight| = 0 ( < 9.1e-06);
##  940 weights are ~= 1. The remaining 10041 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0368  0.8600  0.9510  0.8970  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          9.10e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                  psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Sample size for the above analysis:  10983"
## [1] ""
## [1] ""
## [1] "##################################"
## [1] "Analysis of AJSC 1503"
## [1] "##################################"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 1778 1514 1323 1416 1339 1311 1890 1909 1849 1881 1885 1779 1961 2356 2268
## 2011 2012
## 2619 2514
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
```

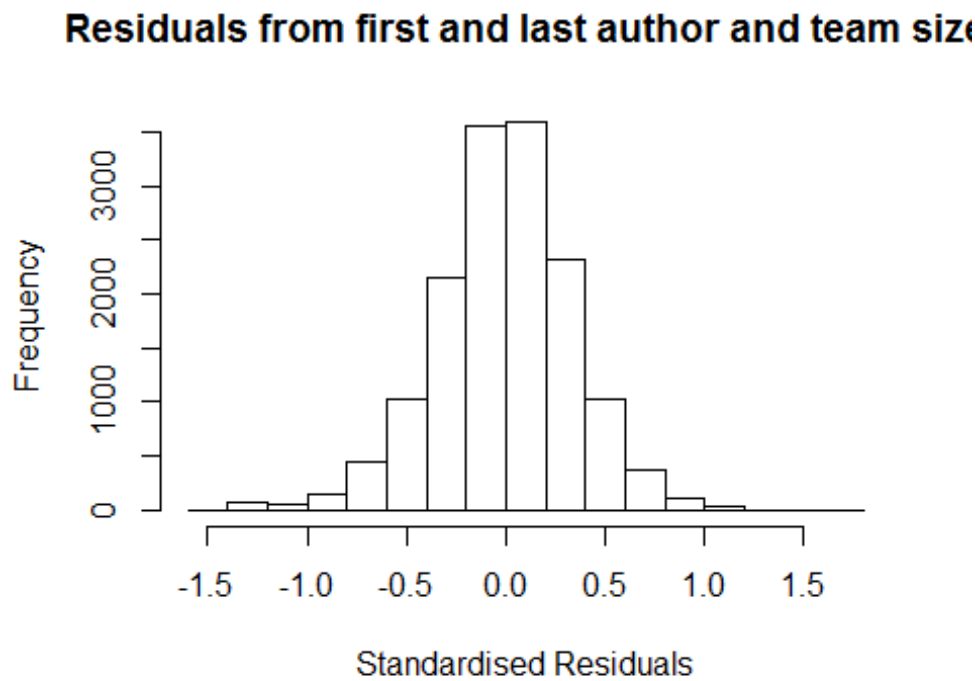
```
## 1132  880  871  924  670  329 1209 1197 1104 1137 1140 1056 1188 1384 1339
## 2011 2012
## 1507 1502
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  912  713  726  751  550  262  984  963  876  913  898  830  962 1138 1074
## 2011 2012
## 1173 1200
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
##  Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 210, df = 16, p-value <2e-16
```




```
##
##  Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 5.4, df = 1, p-value = 0.02
```

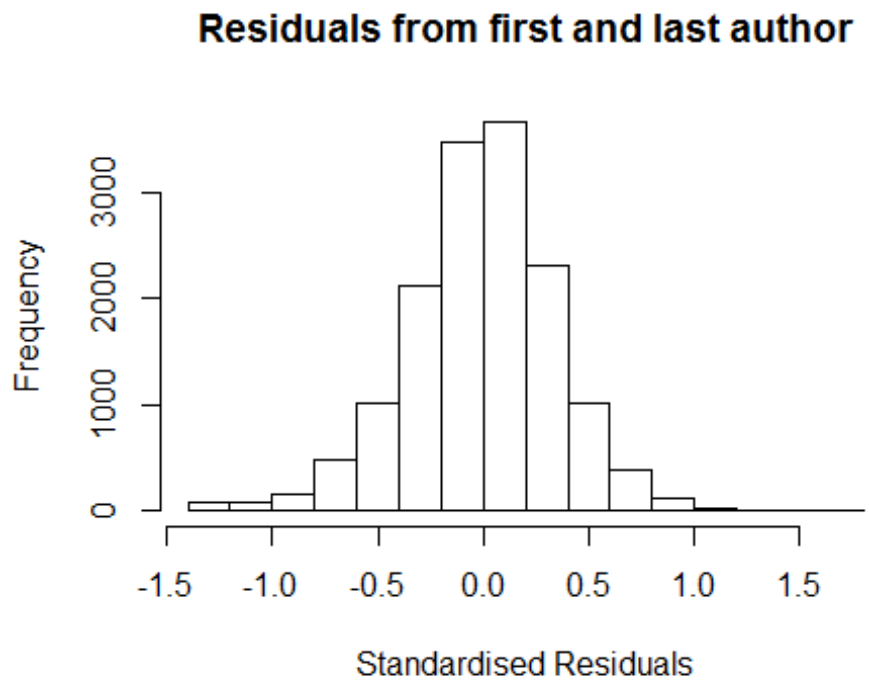



```
## [1] "Female first author team size 2018 geometric mean: 4.8099385692458"
## [1] "Male first author team size 2018 geometric mean: 4.34306258245784"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 130000, p-value = 0.01
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 4.53200758150402"
## [1] "Male last author team size 2018 geometric mean: 4.45004535570808"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 84000, p-value = 1
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.011  1           1.005
## LastAuthorFemale  1.011  1           1.006
## UniqueAuthors     1.056  4           1.007
## Year              1.074 16           1.002
```



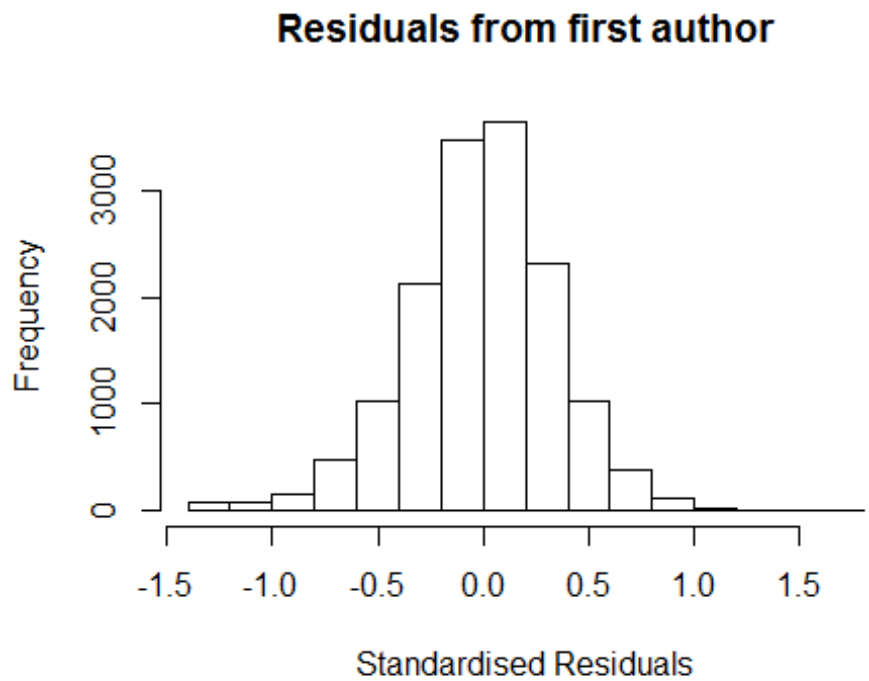

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
UniqueAuthors +
##     Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##     k.max = 1000))
##  \--> method = "MM"
## Residuals:
##       Min        1Q    Median        3Q       Max
## -1.408546 -0.212463  0.000578  0.209708  1.705842
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.21301    0.02534   47.87  < 2e-16 ***
## FirstAuthorFemale1 -0.00223    0.00663   -0.34  0.73692
## LastAuthorFemale1  -0.02126    0.00911   -2.33  0.01969 *
## UniqueAuthors2      0.12277    0.02397    5.12  3.0e-07 ***
## UniqueAuthors3      0.15715    0.02382    6.60  4.3e-11 ***
## UniqueAuthors4      0.15978    0.02404    6.65  3.1e-11 ***
## UniqueAuthors5      0.19554    0.02374    8.24  < 2e-16 ***
## Year1997           -0.02087    0.01807   -1.15  0.24818
## Year1998           -0.01131    0.01818   -0.62  0.53391
## Year1999           -0.05792    0.01816   -3.19  0.00143 **
```

```
## Year2000           -0.06774    0.01939   -3.49  0.00048 ***
## Year2001           -0.07224    0.03289   -2.20  0.02807 *
## Year2002           -0.08919    0.01602   -5.57  2.6e-08 ***
## Year2003           -0.10917    0.01570   -6.95  3.7e-12 ***
## Year2004           -0.09258    0.01598   -5.80  7.0e-09 ***
## Year2005           -0.12005    0.01596   -7.52  5.7e-14 ***
## Year2006           -0.11681    0.01553   -7.52  5.6e-14 ***
## Year2007           -0.09861    0.01650   -5.98  2.3e-09 ***
## Year2008           -0.10070    0.01542   -6.53  6.8e-11 ***
## Year2009           -0.10585    0.01540   -6.87  6.5e-12 ***
## Year2010           -0.08036    0.01568   -5.12  3.0e-07 ***
## Year2011           -0.10987    0.01545   -7.11  1.2e-12 ***
## Year2012           -0.08940    0.01554   -5.75  9.0e-09 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.315
## Multiple R-squared:  0.0222, Adjusted R-squared:  0.0208
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
##  3 observations c(141,3777,8625) are outliers with |weight| = 0 ( < 6.7e-
06);
##  1297 weights are ~= 1. The remaining 13625 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0084  0.8640  0.9510  0.8940  0.9860  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          6.70e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.009  1           1.005
## LastAuthorFemale  1.010  1           1.005
## Year              1.019 16           1.001
```



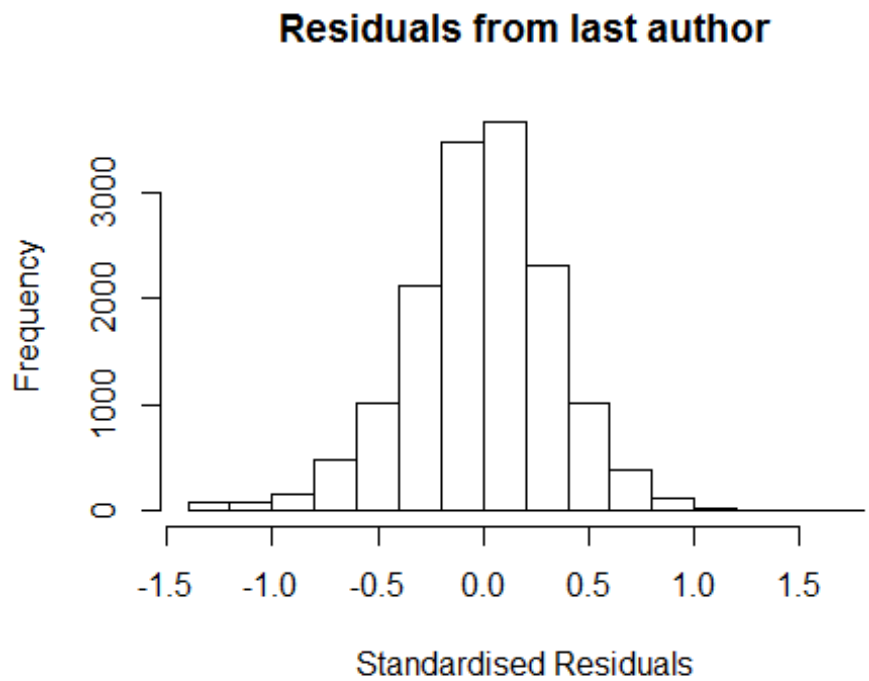

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.3593 -0.2146  0.0034  0.2083  1.7167
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)          1.35932    0.01214  112.01  < 2e-16 ***
## FirstAuthorFemale1  0.00114    0.00666    0.17   0.8639
## LastAuthorFemale1  -0.02118    0.00908   -2.33   0.0197 *
## Year1997           -0.01556    0.01823   -0.85   0.3933
## Year1998           -0.01276    0.01823   -0.70   0.4842
## Year1999           -0.05330    0.01835   -2.91   0.0037 **
## Year2000           -0.06258    0.01952   -3.21   0.0014 **
## Year2001           -0.07309    0.03327   -2.20   0.0280 *
## Year2002           -0.08020    0.01615   -4.97  6.9e-07 ***
## Year2003           -0.10073    0.01580   -6.38  1.9e-10 ***
## Year2004           -0.08318    0.01610   -5.17  2.4e-07 ***
## Year2005           -0.10920    0.01610   -6.78  1.2e-11 ***
```

```
## Year2006           -0.10375    0.01567   -6.62  3.7e-11 ***
## Year2007           -0.08370    0.01654   -5.06  4.2e-07 ***
## Year2008           -0.08589    0.01553   -5.53  3.2e-08 ***
## Year2009           -0.09288    0.01549   -5.99  2.1e-09 ***
## Year2010           -0.06641    0.01577   -4.21  2.5e-05 ***
## Year2011           -0.09377    0.01551   -6.05  1.5e-09 ***
## Year2012           -0.07150    0.01558   -4.59  4.5e-06 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.317
## Multiple R-squared:  0.0094, Adjusted R-squared:  0.0082
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  2 observations c(141,8625) are outliers with |weight| = 0 ( < 6.7e-06);
##  1299 weights are ~= 1. The remaining 13624 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0149  0.8630  0.9510  0.8930  0.9860  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          6.70e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.009  1           1.004
## Year              1.009 16           1.000
```




```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.35756 -0.21346  0.00345  0.20839  1.71844
##
## Coefficients:
##                      Estimate Std. Error t value Pr(>|t|)
## (Intercept)          1.357562   0.012112  112.08  < 2e-16 ***
## FirstAuthorFemale1   0.000325   0.006653    0.05   0.9610
## Year1997            -0.014972   0.018228   -0.82   0.4115
## Year1998            -0.012717   0.018228   -0.70   0.4854
## Year1999            -0.052897   0.018336   -2.88   0.0039 **
## Year2000            -0.062975   0.019523   -3.23   0.0013 **
## Year2001            -0.073097   0.033264   -2.20   0.0280 *
## Year2002            -0.080194   0.016161   -4.96  7.0e-07 ***
## Year2003            -0.100730   0.015805   -6.37  1.9e-10 ***
## Year2004            -0.083167   0.016110   -5.16  2.5e-07 ***
## Year2005            -0.109338   0.016111   -6.79  1.2e-11 ***
## Year2006            -0.103835   0.015675   -6.62  3.6e-11 ***
```

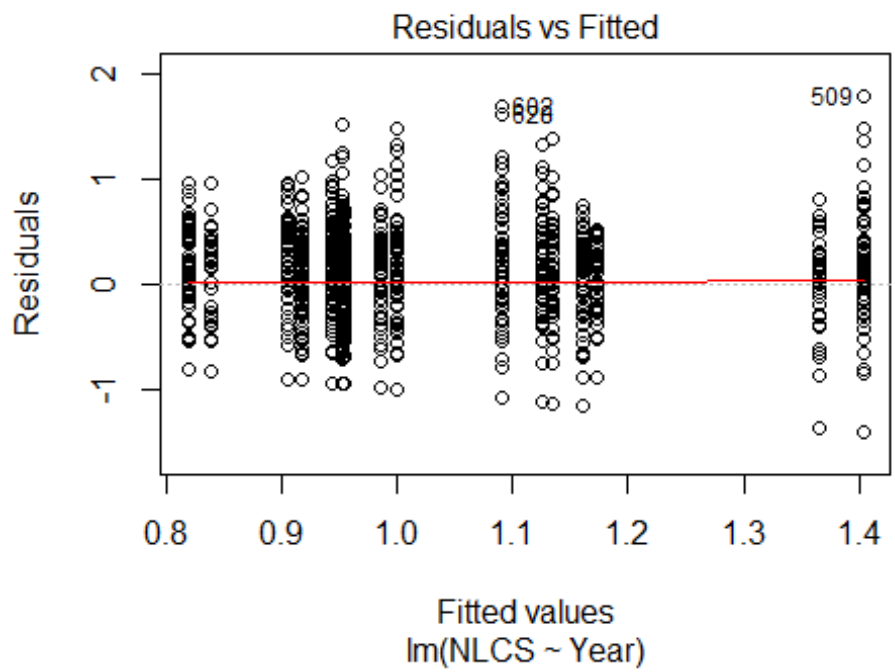
```
## Year2007           -0.084018   0.016543   -5.08  3.8e-07 ***
## Year2008           -0.086098   0.015532   -5.54  3.0e-08 ***
## Year2009           -0.093174   0.015494   -6.01  1.9e-09 ***
## Year2010           -0.066957   0.015765   -4.25  2.2e-05 ***
## Year2011           -0.095193   0.015488   -6.15  8.1e-10 ***
## Year2012           -0.072106   0.015585   -4.63  3.7e-06 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.317
## Multiple R-squared:  0.00903,    Adjusted R-squared:  0.00789
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  2 observations c(141,8625) are outliers with |weight| = 0 ( < 6.7e-06);
##  1308 weights are ~= 1. The remaining 13615 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0143  0.8630  0.9510  0.8930  0.9860  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          6.70e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##                  GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.01  1           1.005
## Year             1.01 16           1.000
```




```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.35951 -0.21472  0.00326  0.20818  1.71649
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.35951    0.01207  112.64  < 2e-16 ***
## LastAuthorFemale1 -0.02110    0.00909   -2.32   0.0202 *
## Year1997          -0.01560    0.01823   -0.86   0.3924
## Year1998          -0.01278    0.01823   -0.70   0.4835
## Year1999          -0.05329    0.01835   -2.90   0.0037 **
## Year2000          -0.06257    0.01952   -3.20   0.0014 **
## Year2001          -0.07310    0.03327   -2.20   0.0280 *
## Year2002          -0.08021    0.01615   -4.97  6.9e-07 ***
## Year2003          -0.10069    0.01580   -6.37  1.9e-10 ***
## Year2004          -0.08312    0.01610   -5.16  2.5e-07 ***
## Year2005          -0.10914    0.01610   -6.78  1.3e-11 ***
## Year2006          -0.10370    0.01567   -6.62  3.7e-11 ***
```

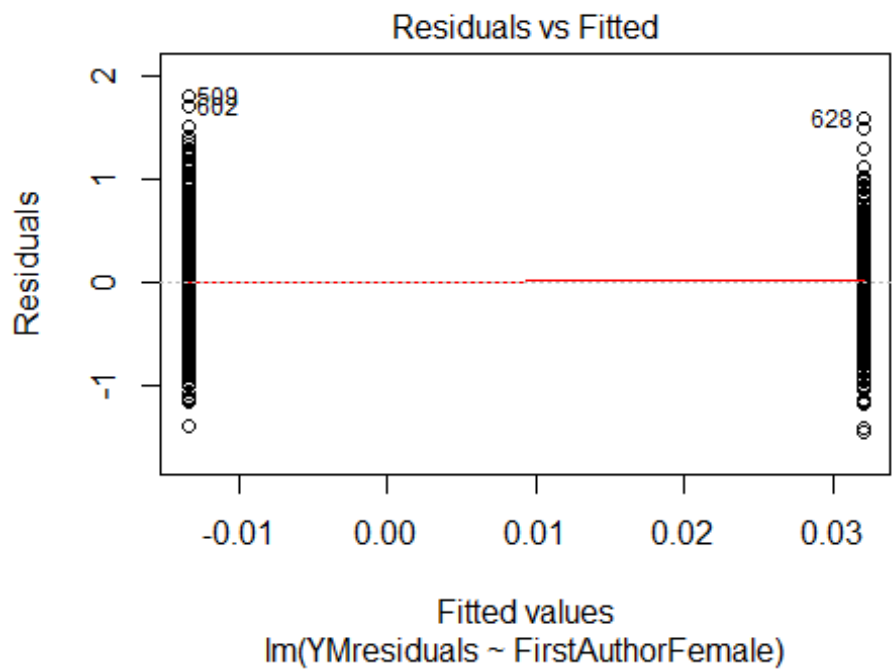
```
## Year2007          -0.08367    0.01654   -5.06  4.3e-07 ***
## Year2008          -0.08581    0.01553   -5.53  3.3e-08 ***
## Year2009          -0.09277    0.01549   -5.99  2.1e-09 ***
## Year2010          -0.06632    0.01576   -4.21  2.6e-05 ***
## Year2011          -0.09369    0.01551   -6.04  1.6e-09 ***
## Year2012          -0.07140    0.01558   -4.58  4.6e-06 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.317
## Multiple R-squared:  0.0094, Adjusted R-squared:  0.00827
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  2 observations c(141,8625) are outliers with |weight| = 0 ( < 6.7e-06);
##  1295 weights are ~= 1. The remaining 13628 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0149  0.8630  0.9510  0.8930  0.9860  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          6.70e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Sample size for the above analysis:  14925"
## [1] ""
## [1] ""
## [1] "####################################"
## [1] "Analysis of AJSC 1504"
## [1] "####################################"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##   58   64   67   91  132  134  118   81   99  113   96  115  161   62   59
## 2011 2012
##  100   90
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##   41   46   48   66   93   75   87   69   72   87   74   92  122   53   43
```

```
## 2011 2012
##   90   79
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##   38   41   45   56   86   69   77   64   61   78   69   80  113   50   41
## 2011 2012
##   83   74
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
##  Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 87, df = 16, p-value = 1e-11
```

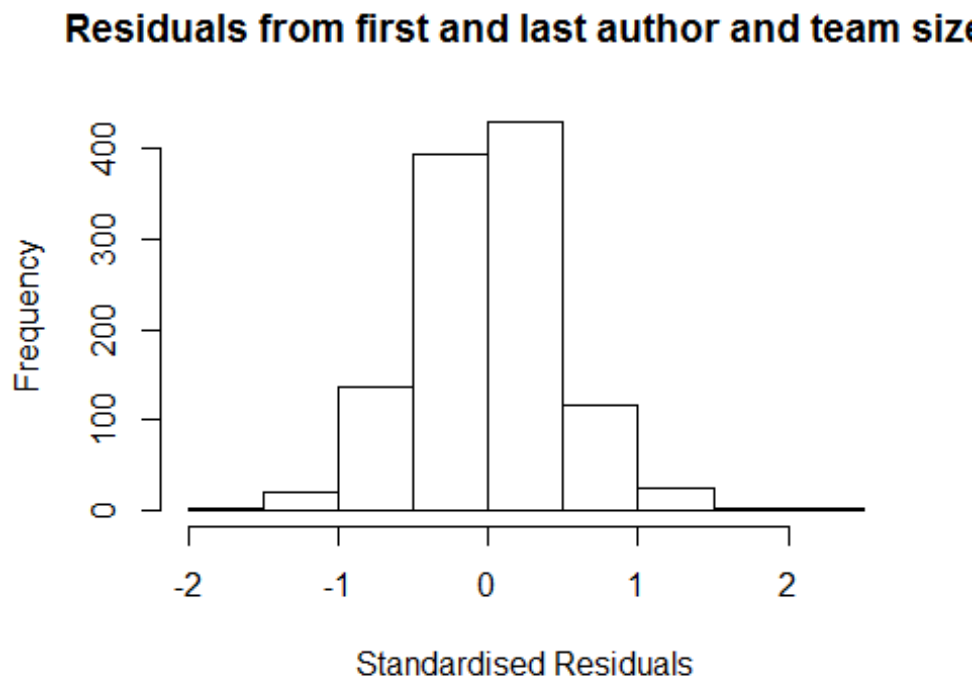



```
##
##  Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 0.0048, df = 1, p-value = 0.9

## [1] "Female first author team size 2018 geometric mean: 4.61868308194015"
## [1] "Male first author team size 2018 geometric mean: 3.5838545038605"

## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties
```

```
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 430, p-value = 0.1
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 3.8735421709082"
## [1] "Male last author team size 2018 geometric mean: 4.21746297180861"

## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties
```

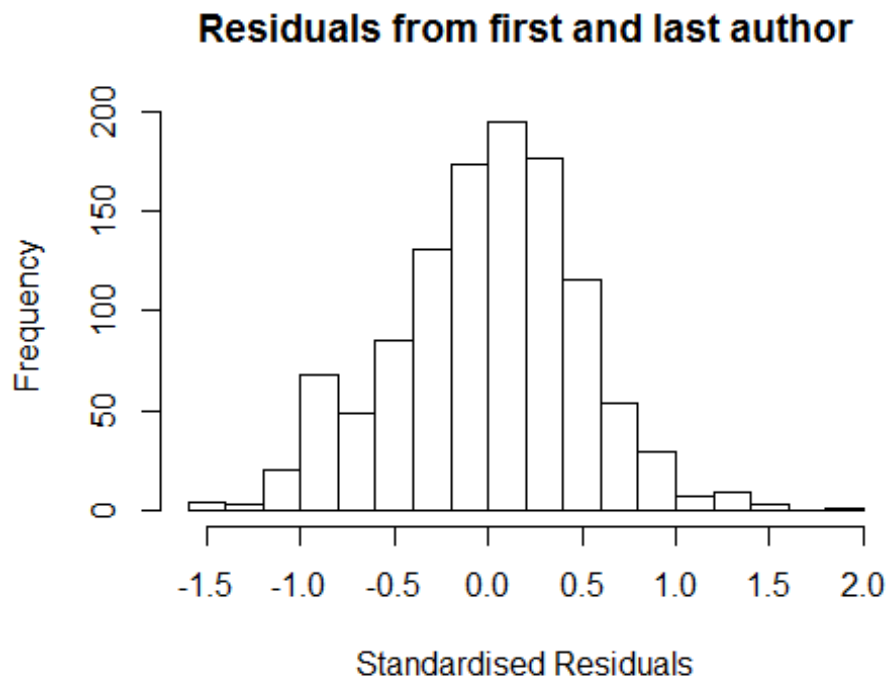



```
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 290, p-value = 0.6
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##                     GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.098  1           1.048
## LastAuthorFemale  1.134  1           1.065
## UniqueAuthors     1.437  4           1.046
## Year              1.579 16           1.014
```



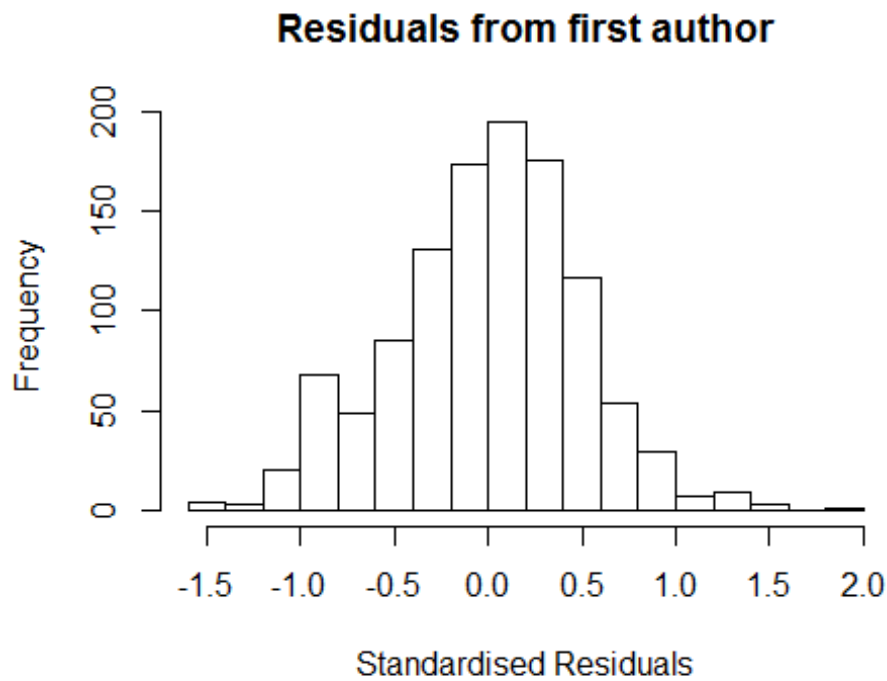

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
UniqueAuthors +
##     Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##     k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.6218 -0.3052  0.0156  0.3031  2.1006
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)          1.0485     0.0712   14.72  < 2e-16 ***
## FirstAuthorFemale1   0.0124     0.0317    0.39   0.6959
## LastAuthorFemale1   -0.0257     0.0336   -0.76   0.4449
## UniqueAuthors2       0.3389     0.0584    5.80  8.5e-09 ***
## UniqueAuthors3       0.3569     0.0524    6.81  1.6e-11 ***
## UniqueAuthors4       0.4100     0.0517    7.92  5.6e-15 ***
## UniqueAuthors5       0.4691     0.0492    9.54  < 2e-16 ***
## Year1997            -0.2415     0.0752   -3.21   0.0014 **
## Year1998            -0.2317     0.0872   -2.66   0.0080 **
## Year1999            -0.4346     0.0824   -5.27  1.6e-07 ***
```

```
## Year2000            -0.4374     0.0802   -5.45  6.2e-08 ***
## Year2001             0.0918     0.0951    0.97   0.3346
## Year2002            -0.3622     0.0911   -3.98  7.5e-05 ***
## Year2003            -0.2077     0.0944   -2.20   0.0280 *
## Year2004            -0.2453     0.0862   -2.85   0.0045 **
## Year2005            -0.3651     0.0864   -4.23  2.6e-05 ***
## Year2006            -0.4142     0.0776   -5.33  1.2e-07 ***
## Year2007            -0.4480     0.0734   -6.10  1.4e-09 ***
## Year2008            -0.4753     0.0686   -6.93  7.2e-12 ***
## Year2009            -0.4388     0.0790   -5.55  3.5e-08 ***
## Year2010            -0.5321     0.0911   -5.84  6.9e-09 ***
## Year2011            -0.5189     0.0778   -6.67  4.0e-11 ***
## Year2012            -0.4462     0.0848   -5.26  1.7e-07 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.45
## Multiple R-squared:  0.203,  Adjusted R-squared:  0.187
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
##  observation 341 is an outlier with |weight| <= 3.3e-05 ( < 8.9e-05);
##  94 weights are ~= 1. The remaining 1030 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.166   0.866   0.951   0.897   0.986   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          8.89e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.045  1           1.022
## LastAuthorFemale  1.065  1           1.032
## Year              1.102 16           1.003
```



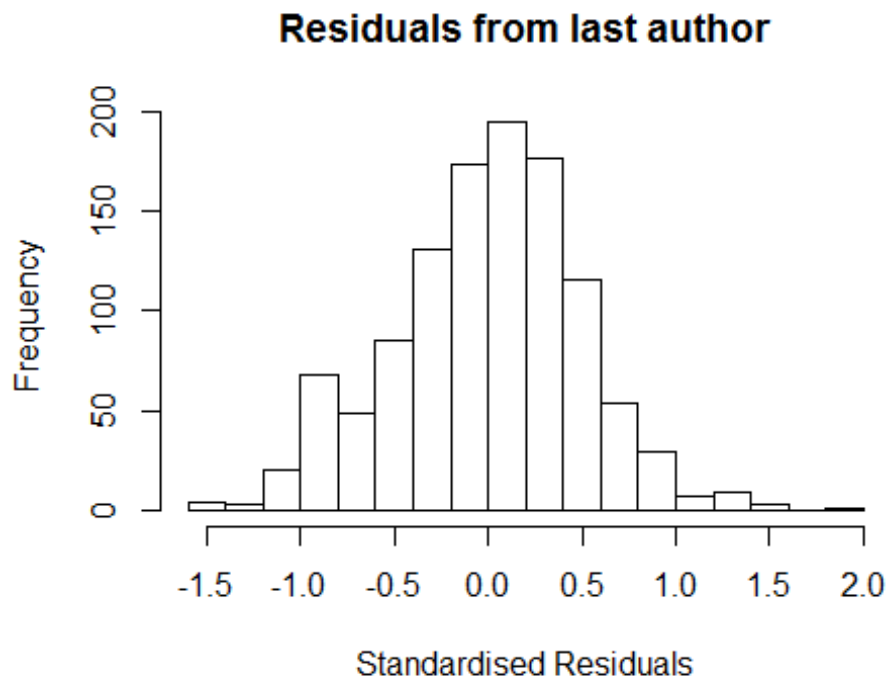

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.4652 -0.3223  0.0243  0.3305  1.8188
##
## Coefficients:
##                      Estimate Std. Error t value Pr(>|t|)
## (Intercept)          1.396851   0.058709   23.79  < 2e-16 ***
## FirstAuthorFemale1   0.068325   0.032837    2.08   0.0377 *
## LastAuthorFemale1   -0.000924   0.035482   -0.03   0.9792
## Year1997            -0.238642   0.076997   -3.10   0.0020 **
## Year1998            -0.252154   0.087099   -2.90   0.0039 **
## Year1999            -0.501578   0.094027   -5.33  1.2e-07 ***
## Year2000            -0.472362   0.084760   -5.57  3.1e-08 ***
## Year2001            -0.003408   0.098282   -0.03   0.9723
## Year2002            -0.428615   0.098856   -4.34  1.6e-05 ***
## Year2003            -0.267361   0.095098   -2.81   0.0050 **
## Year2004            -0.275862   0.092981   -2.97   0.0031 **
## Year2005            -0.419457   0.093996   -4.46  8.9e-06 ***
```

```
## Year2006           -0.422693   0.080454   -5.25  1.8e-07 ***
## Year2007           -0.485632   0.075262   -6.45  1.6e-10 ***
## Year2008           -0.508842   0.068285   -7.45  1.9e-13 ***
## Year2009           -0.469337   0.081399   -5.77  1.1e-08 ***
## Year2010           -0.591104   0.095783   -6.17  9.5e-10 ***
## Year2011           -0.606441   0.080138   -7.57  8.0e-14 ***
## Year2012           -0.504897   0.088642   -5.70  1.6e-08 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.472
## Multiple R-squared:  0.101,  Adjusted R-squared:  0.0863
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
##  97 weights are ~= 1. The remaining 1028 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.104   0.864   0.949   0.896   0.985   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          8.89e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.039  1           1.019
## Year              1.039 16           1.001
```




```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.4650 -0.3230  0.0246  0.3308  1.8190
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)          1.39683    0.05854   23.86  < 2e-16 ***
## FirstAuthorFemale1  0.06821    0.03297    2.07   0.0388 *
## Year1997           -0.23871    0.07700   -3.10   0.0020 **
## Year1998           -0.25223    0.08709   -2.90   0.0039 **
## Year1999           -0.50179    0.09401   -5.34  1.1e-07 ***
## Year2000           -0.47244    0.08476   -5.57  3.1e-08 ***
## Year2001           -0.00349    0.09780   -0.04   0.9715
## Year2002           -0.42880    0.09867   -4.35  1.5e-05 ***
## Year2003           -0.26739    0.09514   -2.81   0.0050 **
## Year2004           -0.27596    0.09305   -2.97   0.0031 **
## Year2005           -0.41953    0.09408   -4.46  9.1e-06 ***
## Year2006           -0.42279    0.08029   -5.27  1.7e-07 ***
```

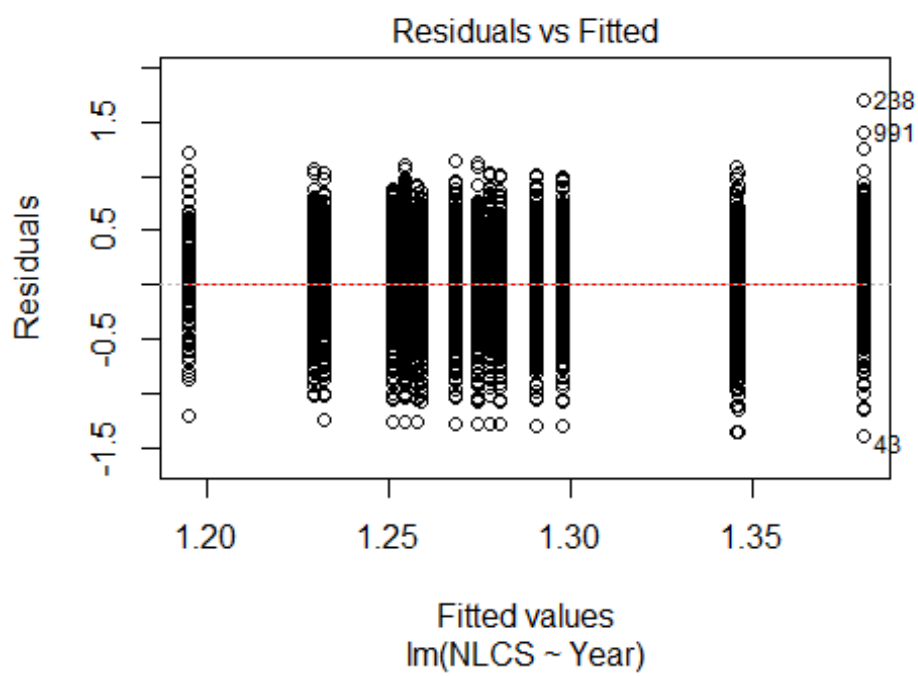
```
## Year2007           -0.48586    0.07509   -6.47  1.5e-10 ***
## Year2008           -0.50901    0.06806   -7.48  1.5e-13 ***
## Year2009           -0.46957    0.08112   -5.79  9.2e-09 ***
## Year2010           -0.59123    0.09579   -6.17  9.4e-10 ***
## Year2011           -0.60666    0.07984   -7.60  6.3e-14 ***
## Year2012           -0.50502    0.08874   -5.69  1.6e-08 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.471
## Multiple R-squared:  0.101,  Adjusted R-squared:  0.0872
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
##  97 weights are ~= 1. The remaining 1028 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.103   0.864   0.949   0.896   0.985   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          8.89e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##                  GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.06  1           1.030
## Year             1.06 16           1.002
```



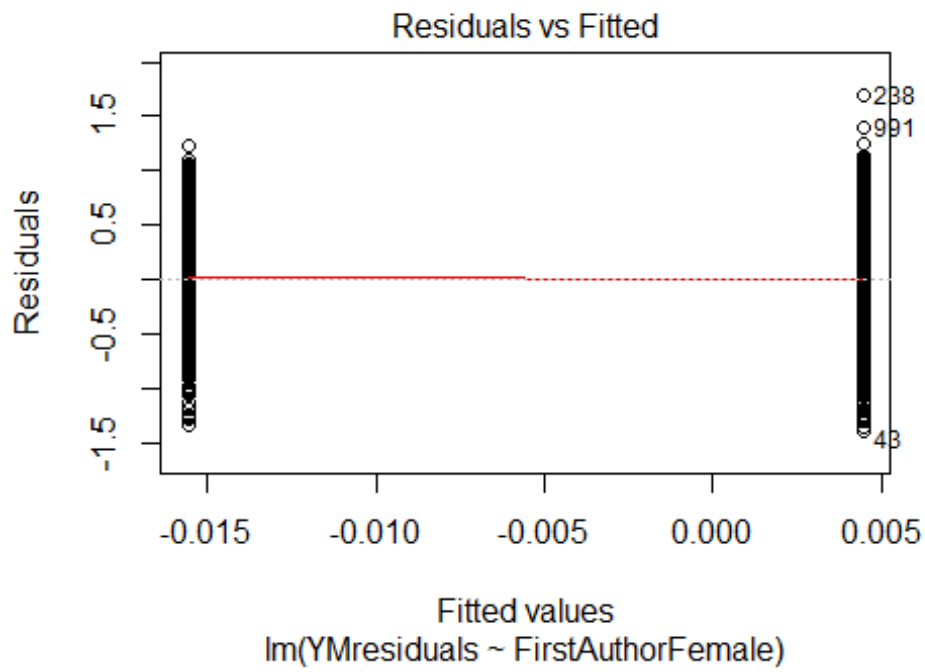

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.4162 -0.3209  0.0223  0.3276  1.8029
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)          1.4131     0.0575   24.56  < 2e-16 ***
## LastAuthorFemale1    0.0129     0.0355    0.36   0.7157
## Year1997            -0.2334     0.0765   -3.05   0.0023 **
## Year1998            -0.2465     0.0864   -2.85   0.0044 **
## Year1999            -0.4989     0.0944   -5.28  1.5e-07 ***
## Year2000            -0.4765     0.0849   -5.61  2.5e-08 ***
## Year2001            -0.0098     0.0976   -0.10   0.9200
## Year2002            -0.4290     0.0979   -4.38  1.3e-05 ***
## Year2003            -0.2703     0.0945   -2.86   0.0043 **
## Year2004            -0.2785     0.0926   -3.01   0.0027 **
## Year2005            -0.4216     0.0936   -4.50  7.4e-06 ***
## Year2006            -0.4167     0.0803   -5.19  2.5e-07 ***
```

```
## Year2007           -0.4867     0.0752   -6.47  1.5e-10 ***
## Year2008           -0.5054     0.0678   -7.45  1.9e-13 ***
## Year2009           -0.4663     0.0802   -5.81  8.0e-09 ***
## Year2010           -0.5825     0.0973   -5.99  2.9e-09 ***
## Year2011           -0.6096     0.0793   -7.69  3.2e-14 ***
## Year2012           -0.4972     0.0884   -5.63  2.3e-08 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.473
## Multiple R-squared:  0.0971, Adjusted R-squared:  0.0832
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
##  90 weights are ~= 1. The remaining 1035 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.114   0.862   0.950   0.897   0.986   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          8.89e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Sample size for the above analysis:  1125"
## [1] ""
## [1] ""
## [1] "####################################"
## [1] "Analysis of AJSC 1505"
## [1] "####################################"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 1358 1083  874  949  884  810 1377 1409 1443 1405 1467 1293 1375 1552 1488
## 2011 2012
## 1534 1432
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  842  572  530  641  469  154  884  870  863  819  893  773  868  917  902
## 2011 2012
```

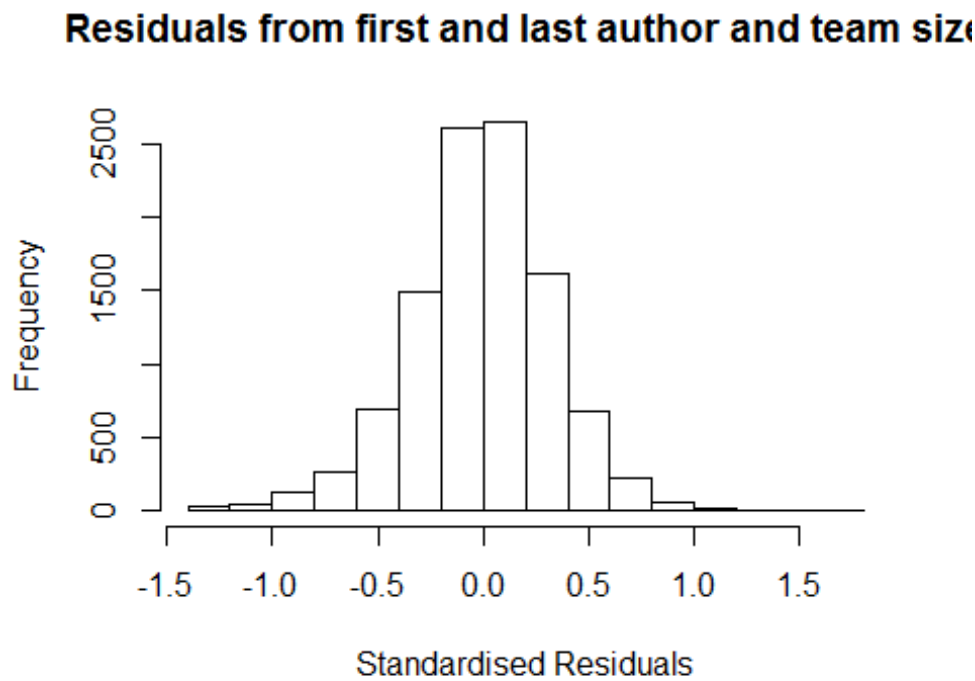
```
##  946  878
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  690  474  452  536  383  126  726  721  692  683  723  614  708  760  736
## 2011 2012
##  740  718
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
##  Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 160, df = 16, p-value <2e-16
```




```
##
##  Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 3, df = 1, p-value = 0.09
```

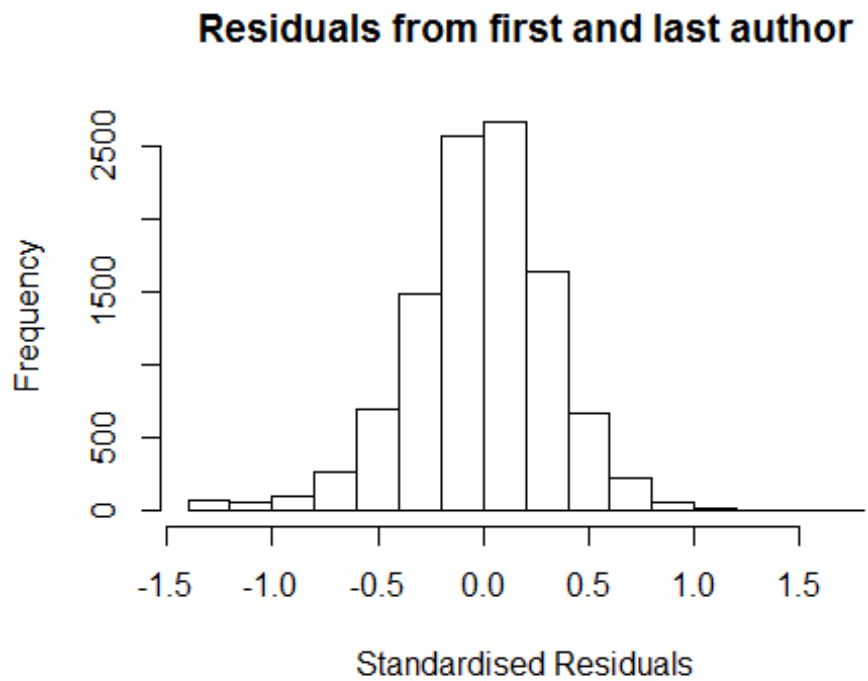



```
## [1] "Female first author team size 2018 geometric mean: 4.78737896158565"
## [1] "Male first author team size 2018 geometric mean: 4.56127945674712"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 21000, p-value = 0.5
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 4.86359764207168"
## [1] "Male last author team size 2018 geometric mean: 4.5724790069773"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 14000, p-value = 0.7
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.025  1           1.012
## LastAuthorFemale  1.022  1           1.011
## UniqueAuthors     1.133  4           1.016
## Year              1.121 16           1.004
```



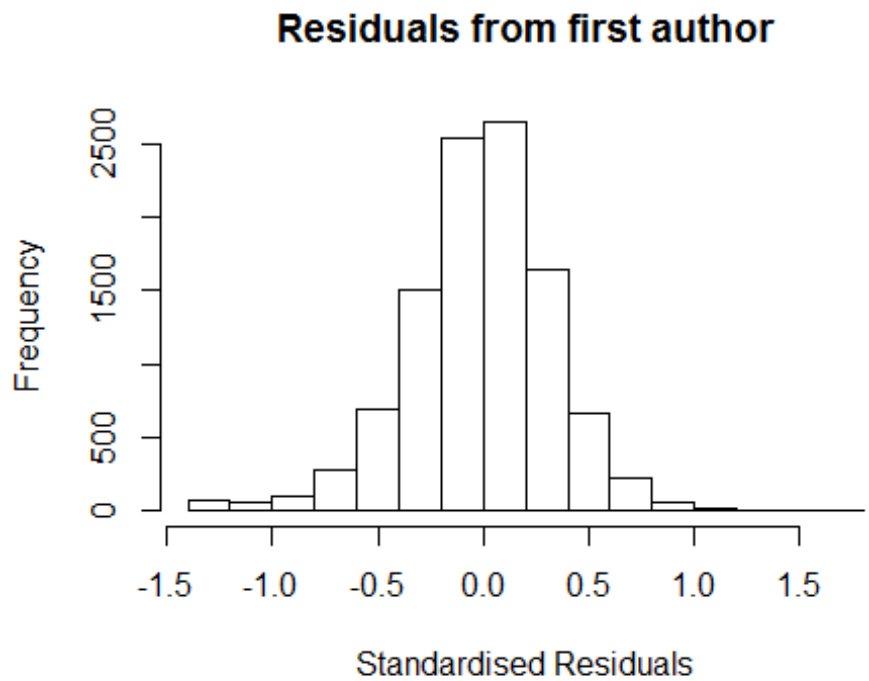

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
UniqueAuthors +
##     Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##     k.max = 1000))
##  \--> method = "MM"
## Residuals:
##       Min        1Q    Median        3Q       Max
## -1.393294 -0.200876  0.000751  0.197554  1.668668
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)           1.0836     0.0343   31.61  < 2e-16 ***
## FirstAuthorFemale1   -0.0172     0.0075   -2.29   0.0218 *
## LastAuthorFemale1    -0.0274     0.0101   -2.71   0.0067 **
## UniqueAuthors2        0.2763     0.0326    8.48  < 2e-16 ***
## UniqueAuthors3        0.3237     0.0325    9.97  < 2e-16 ***
## UniqueAuthors4        0.3245     0.0328    9.91  < 2e-16 ***
## UniqueAuthors5        0.3552     0.0325   10.94  < 2e-16 ***
## Year1997             -0.0215     0.0206   -1.05   0.2948
## Year1998             -0.0140     0.0209   -0.67   0.5024
## Year1999             -0.0601     0.0205   -2.93   0.0034 **
```

```
## Year2000            -0.0940     0.0226   -4.16  3.2e-05 ***
## Year2001            -0.1765     0.0391   -4.52  6.4e-06 ***
## Year2002            -0.0995     0.0176   -5.67  1.5e-08 ***
## Year2003            -0.1324     0.0171   -7.73  1.2e-14 ***
## Year2004            -0.1326     0.0178   -7.44  1.1e-13 ***
## Year2005            -0.1558     0.0179   -8.70  < 2e-16 ***
## Year2006            -0.1601     0.0173   -9.27  < 2e-16 ***
## Year2007            -0.1427     0.0181   -7.90  3.0e-15 ***
## Year2008            -0.1404     0.0174   -8.08  7.5e-16 ***
## Year2009            -0.1192     0.0177   -6.72  1.9e-11 ***
## Year2010            -0.1216     0.0180   -6.75  1.6e-11 ***
## Year2011            -0.1329     0.0185   -7.20  6.4e-13 ***
## Year2012            -0.0945     0.0180   -5.25  1.6e-07 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.301
## Multiple R-squared:  0.0547, Adjusted R-squared:  0.0527
## Convergence in 17 IRWLS iterations
##
## Robustness weights:
##  2 observations c(117,474) are outliers with |weight| = 0 ( < 9.5e-06);
##  882 weights are ~= 1. The remaining 9598 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0005  0.8630  0.9510  0.8930  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          9.54e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.010  1           1.005
## LastAuthorFemale  1.009  1           1.005
## Year              1.020 16           1.001
```



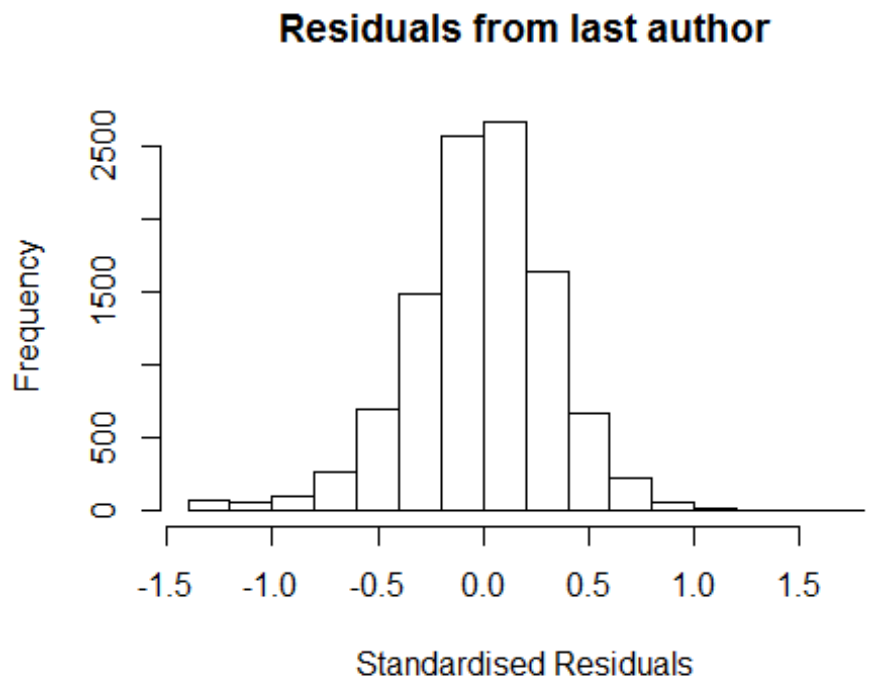

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.36927 -0.20531  0.00136  0.19795  1.68654
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.38946    0.01350  102.93  < 2e-16 ***
## FirstAuthorFemale1 -0.01007    0.00748   -1.35   0.1785
## LastAuthorFemale1  -0.02495    0.01005   -2.48   0.0130 *
## Year1997           -0.02019    0.02084   -0.97   0.3327
## Year1998           -0.02441    0.02100   -1.16   0.2450
## Year1999           -0.05996    0.02067   -2.90   0.0037 **
## Year2000           -0.10356    0.02289   -4.52  6.1e-06 ***
## Year2001           -0.20794    0.03788   -5.49  4.1e-08 ***
## Year2002           -0.09154    0.01765   -5.19  2.2e-07 ***
## Year2003           -0.12738    0.01727   -7.38  1.7e-13 ***
## Year2004           -0.12362    0.01796   -6.88  6.1e-12 ***
## Year2005           -0.14766    0.01797   -8.22  2.3e-16 ***
```

```
## Year2006           -0.15043    0.01741   -8.64  < 2e-16 ***
## Year2007           -0.12861    0.01816   -7.08  1.5e-12 ***
## Year2008           -0.12904    0.01743   -7.40  1.4e-13 ***
## Year2009           -0.10441    0.01769   -5.90  3.7e-09 ***
## Year2010           -0.10484    0.01799   -5.83  5.8e-09 ***
## Year2011           -0.11148    0.01830   -6.09  1.2e-09 ***
## Year2012           -0.07223    0.01783   -4.05  5.1e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.304
## Multiple R-squared:  0.0212, Adjusted R-squared:  0.0196
## Convergence in 10 IRWLS iterations
##
## Robustness weights:
##  observation 117 is an outlier with |weight| = 0 ( < 9.5e-06);
##  899 weights are ~= 1. The remaining 9582 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0025  0.8630  0.9520  0.8920  0.9860  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          9.54e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.011  1           1.005
## Year              1.011 16           1.000
```




```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##       Min        1Q    Median        3Q       Max
## -1.387324 -0.206225  0.000551  0.197763  1.688676
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.38732    0.01344  103.19  < 2e-16 ***
## FirstAuthorFemale1 -0.01112    0.00749   -1.48   0.1376
## Year1997           -0.01989    0.02083   -0.95   0.3396
## Year1998           -0.02410    0.02098   -1.15   0.2507
## Year1999           -0.05988    0.02065   -2.90   0.0037 **
## Year2000           -0.10371    0.02288   -4.53  5.9e-06 ***
## Year2001           -0.20862    0.03780   -5.52  3.5e-08 ***
## Year2002           -0.09164    0.01764   -5.20  2.1e-07 ***
## Year2003           -0.12770    0.01726   -7.40  1.5e-13 ***
## Year2004           -0.12344    0.01795   -6.88  6.4e-12 ***
## Year2005           -0.14823    0.01796   -8.25  < 2e-16 ***
## Year2006           -0.15080    0.01741   -8.66  < 2e-16 ***
```

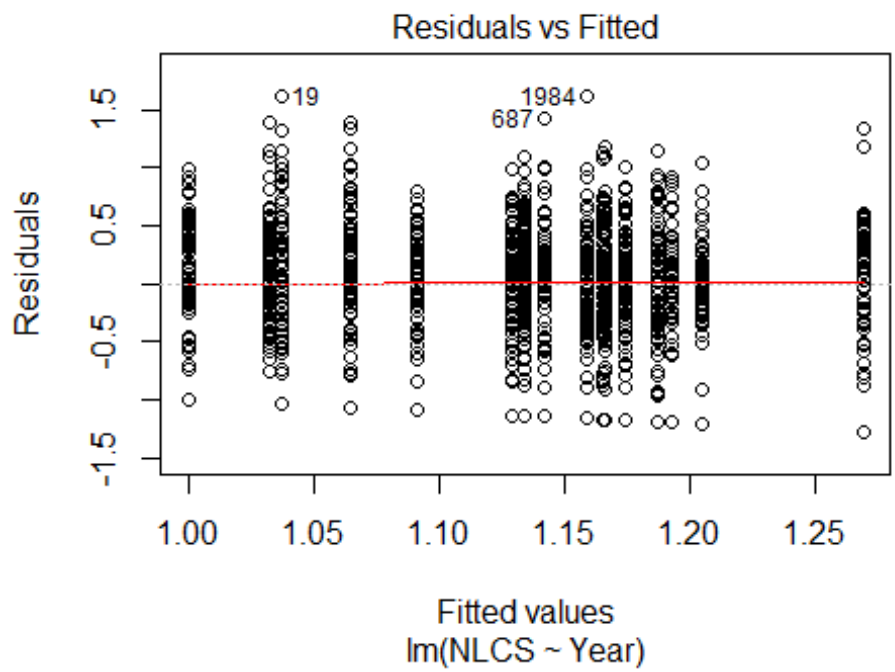
```
## Year2007           -0.12958    0.01815   -7.14  9.9e-13 ***
## Year2008           -0.12999    0.01742   -7.46  9.1e-14 ***
## Year2009           -0.10475    0.01767   -5.93  3.2e-09 ***
## Year2010           -0.10535    0.01798   -5.86  4.8e-09 ***
## Year2011           -0.11293    0.01828   -6.18  6.8e-10 ***
## Year2012           -0.07306    0.01781   -4.10  4.1e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.304
## Multiple R-squared:  0.0206, Adjusted R-squared:  0.0191
## Convergence in 10 IRWLS iterations
##
## Robustness weights:
##  observation 117 is an outlier with |weight| = 0 ( < 9.5e-06);
##  908 weights are ~= 1. The remaining 9573 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0022  0.8640  0.9520  0.8920  0.9860  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          9.54e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##                   GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.009  1           1.005
## Year             1.009 16           1.000
```




```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##       Min        1Q    Median        3Q       Max
## -1.367523 -0.204807  0.000861  0.196957  1.688223
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.3878     0.0134  103.33  < 2e-16 ***
## LastAuthorFemale1  -0.0257     0.0101   -2.56   0.0106 *
## Year1997           -0.0203     0.0208   -0.97   0.3313
## Year1998           -0.0246     0.0210   -1.17   0.2412
## Year1999           -0.0601     0.0207   -2.91   0.0036 **
## Year2000           -0.1039     0.0229   -4.54  5.7e-06 ***
## Year2001           -0.2082     0.0378   -5.50  3.9e-08 ***
## Year2002           -0.0916     0.0176   -5.19  2.1e-07 ***
## Year2003           -0.1278     0.0173   -7.40  1.4e-13 ***
## Year2004           -0.1243     0.0179   -6.93  4.6e-12 ***
## Year2005           -0.1485     0.0180   -8.26  < 2e-16 ***
## Year2006           -0.1509     0.0174   -8.68  < 2e-16 ***
```

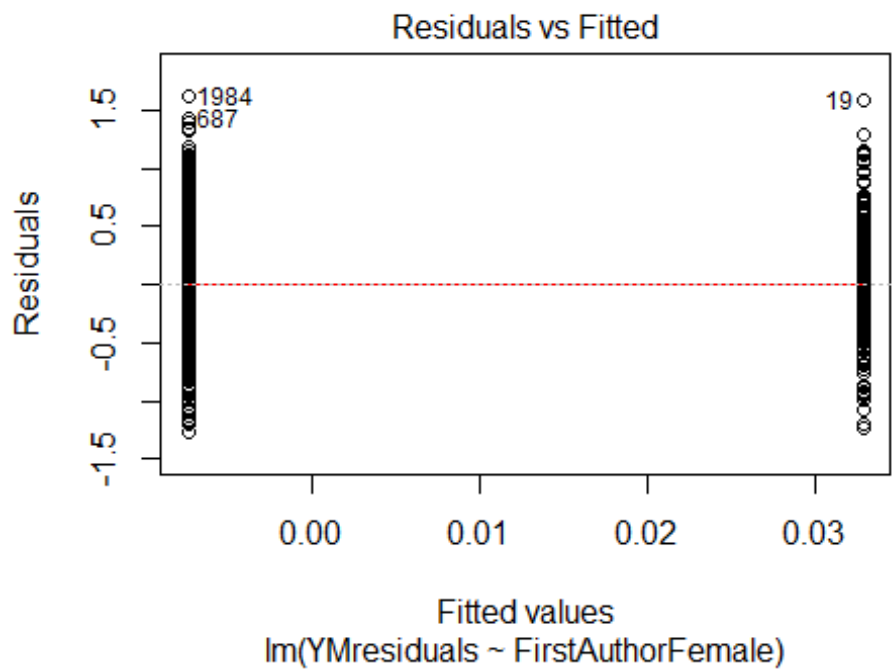
```
## Year2007           -0.1290     0.0181   -7.11  1.3e-12 ***
## Year2008           -0.1297     0.0174   -7.44  1.1e-13 ***
## Year2009           -0.1056     0.0177   -5.98  2.3e-09 ***
## Year2010           -0.1058     0.0180   -5.89  4.0e-09 ***
## Year2011           -0.1123     0.0183   -6.14  8.7e-10 ***
## Year2012           -0.0732     0.0178   -4.11  4.1e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.304
## Multiple R-squared:  0.0211, Adjusted R-squared:  0.0195
## Convergence in 10 IRWLS iterations
##
## Robustness weights:
##  observation 117 is an outlier with |weight| = 0 ( < 9.5e-06);
##  890 weights are ~= 1. The remaining 9591 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0023  0.8630  0.9520  0.8920  0.9860  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          9.54e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                    psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Sample size for the above analysis:  10482"
## [1] ""
## [1] ""
## [1] "####################################"
## [1] "Analysis of AJSC 1506"
## [1] "####################################"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  154  201  163  134  149  133  146  110  114  134  188  124  198  154  196
## 2011 2012
##  153  160
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##   89  107   97   61   72   63   78   68   55   77   97   72   99   80  111
```

```
## 2011 2012
##   73   85
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##   74   93   83   51   62   52   67   60   42   60   74   64   78   67   90
## 2011 2012
##   58   66
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
##  Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 39, df = 16, p-value = 0.001
```

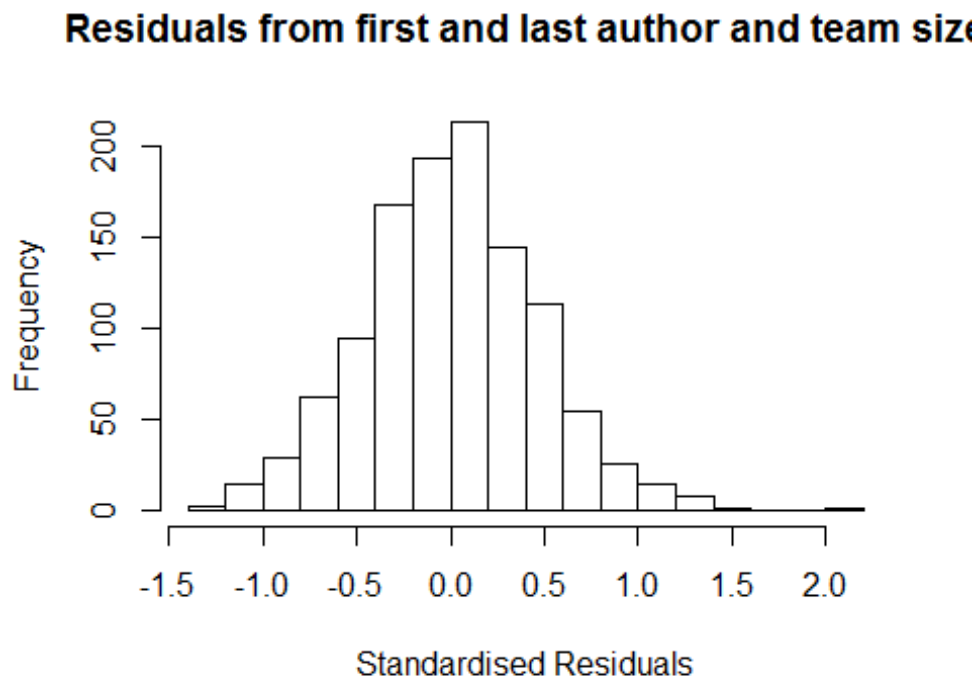



```
##
##  Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 0.57, df = 1, p-value = 0.4

## [1] "Female first author team size 2018 geometric mean: 5.49459858743711"
## [1] "Male first author team size 2018 geometric mean: 3.24153457581968"

## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties
```

```
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 240, p-value = 0.02
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 3.98742113447093"
## [1] "Male last author team size 2018 geometric mean: 3.43017598301107"

## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties
```

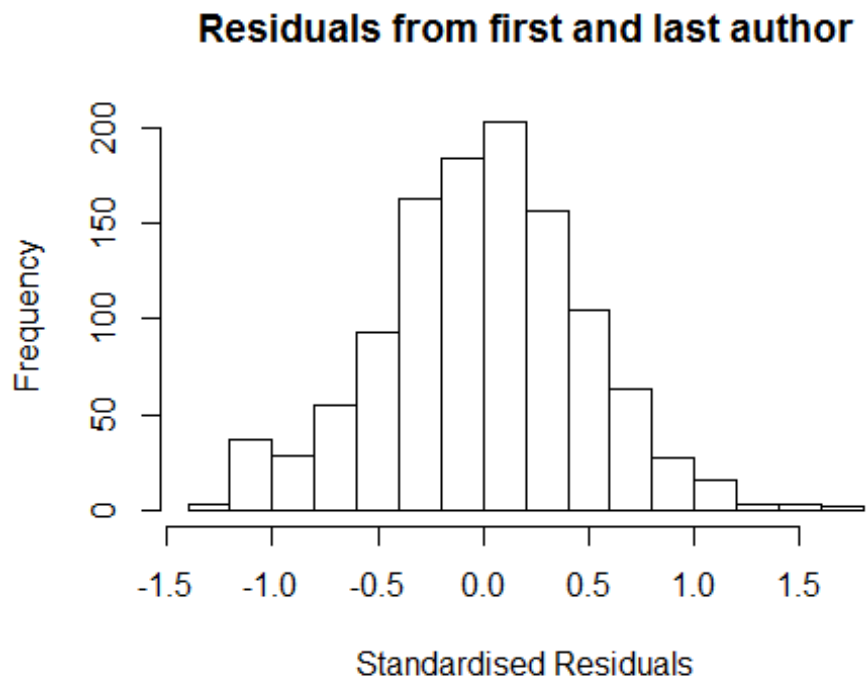



```
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 130, p-value = 0.7
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##                     GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.081  1           1.040
## LastAuthorFemale  1.121  1           1.059
## UniqueAuthors     1.682  4           1.067
## Year              1.932 16           1.021
```



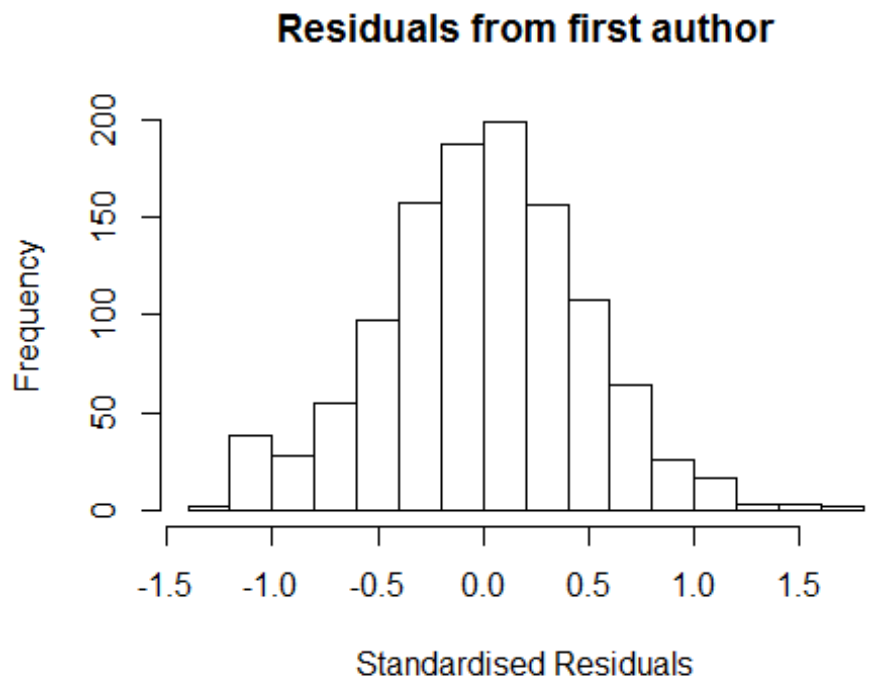

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
UniqueAuthors +
##     Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##     k.max = 1000))
##  \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.30717 -0.30196  0.00652  0.29883  2.04156
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         0.65803    0.06154   10.69  < 2e-16 ***
## FirstAuthorFemale1 -0.00291    0.03570   -0.08    0.935
## LastAuthorFemale1  -0.00753    0.04819   -0.16    0.876
## UniqueAuthors2      0.42552    0.05502    7.73  2.3e-14 ***
## UniqueAuthors3      0.46618    0.05531    8.43  < 2e-16 ***
## UniqueAuthors4      0.50777    0.05968    8.51  < 2e-16 ***
## UniqueAuthors5      0.52911    0.05924    8.93  < 2e-16 ***
## Year1997            0.07093    0.08222    0.86    0.388
## Year1998            0.02378    0.08443    0.28    0.778
## Year1999            0.06947    0.08978    0.77    0.439
```

```
## Year2000            0.10178    0.09469    1.07    0.283
## Year2001            0.22442    0.09130    2.46    0.014 *
## Year2002            0.03826    0.08590    0.45    0.656
## Year2003            0.00420    0.08486    0.05    0.960
## Year2004            0.12680    0.09008    1.41    0.160
## Year2005            0.14137    0.08587    1.65    0.100 .
## Year2006            0.06503    0.07836    0.83    0.407
## Year2007            0.10321    0.09379    1.10    0.271
## Year2008            0.06941    0.08203    0.85    0.398
## Year2009            0.03193    0.08533    0.37    0.708
## Year2010           -0.09330    0.07970   -1.17    0.242
## Year2011            0.05102    0.07991    0.64    0.523
## Year2012            0.00603    0.08295    0.07    0.942
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.445
## Multiple R-squared:  0.14,   Adjusted R-squared:  0.124
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
##  92 weights are ~= 1. The remaining 1049 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0016  0.8700  0.9500  0.9030  0.9860  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          8.76e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.063  1           1.031
## LastAuthorFemale  1.098  1           1.048
## Year              1.167 16           1.005
```



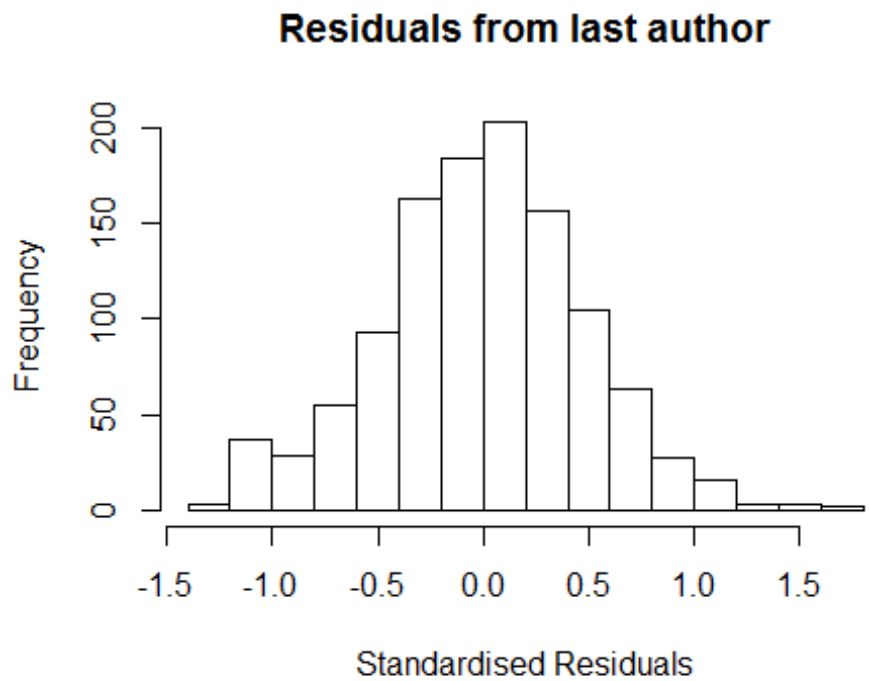

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.3258 -0.3125  0.0111  0.3216  1.6941
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)           0.9242     0.0744   12.41   <2e-16 ***
## FirstAuthorFemale1    0.0338     0.0362    0.93   0.3511
## LastAuthorFemale1     0.0303     0.0486    0.62   0.5323
## Year1997              0.1077     0.1025    1.05   0.2938
## Year1998              0.0869     0.0971    0.90   0.3709
## Year1999              0.1687     0.1044    1.62   0.1064
## Year2000              0.2498     0.1004    2.49   0.0130 *
## Year2001              0.4016     0.0958    4.19    3e-05 ***
## Year2002              0.1829     0.0974    1.88   0.0605 .
## Year2003              0.1598     0.0946    1.69   0.0912 .
## Year2004              0.2665     0.0989    2.69   0.0072 **
## Year2005              0.2633     0.0939    2.80   0.0051 **
```

```
## Year2006             0.2236     0.0887    2.52   0.0118 *
## Year2007             0.2722     0.0977    2.79   0.0054 **
## Year2008             0.2174     0.0903    2.41   0.0162 *
## Year2009             0.2067     0.0933    2.22   0.0269 *
## Year2010             0.0783     0.0867    0.90   0.3667
## Year2011             0.2411     0.0862    2.80   0.0053 **
## Year2012             0.1979     0.0929    2.13   0.0333 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.459
## Multiple R-squared:  0.0356, Adjusted R-squared:  0.0201
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
##  95 weights are ~= 1. The remaining 1046 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.145   0.861   0.950   0.899   0.985   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          8.76e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.063  1           1.031
## Year              1.063 16           1.002
```




```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.32871 -0.31218  0.00772  0.31937  1.68651
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)           0.9293     0.0734   12.67  < 2e-16 ***
## FirstAuthorFemale1    0.0362     0.0365    0.99   0.3216
## Year1997              0.1041     0.1023    1.02   0.3091
## Year1998              0.0834     0.0967    0.86   0.3883
## Year1999              0.1670     0.1044    1.60   0.1100
## Year2000              0.2453     0.0996    2.46   0.0140 *
## Year2001              0.3994     0.0957    4.17  3.3e-05 ***
## Year2002              0.1806     0.0975    1.85   0.0642 .
## Year2003              0.1592     0.0946    1.68   0.0927 .
## Year2004              0.2629     0.0983    2.68   0.0076 **
## Year2005              0.2629     0.0941    2.79   0.0053 **
## Year2006              0.2219     0.0884    2.51   0.0122 *
```

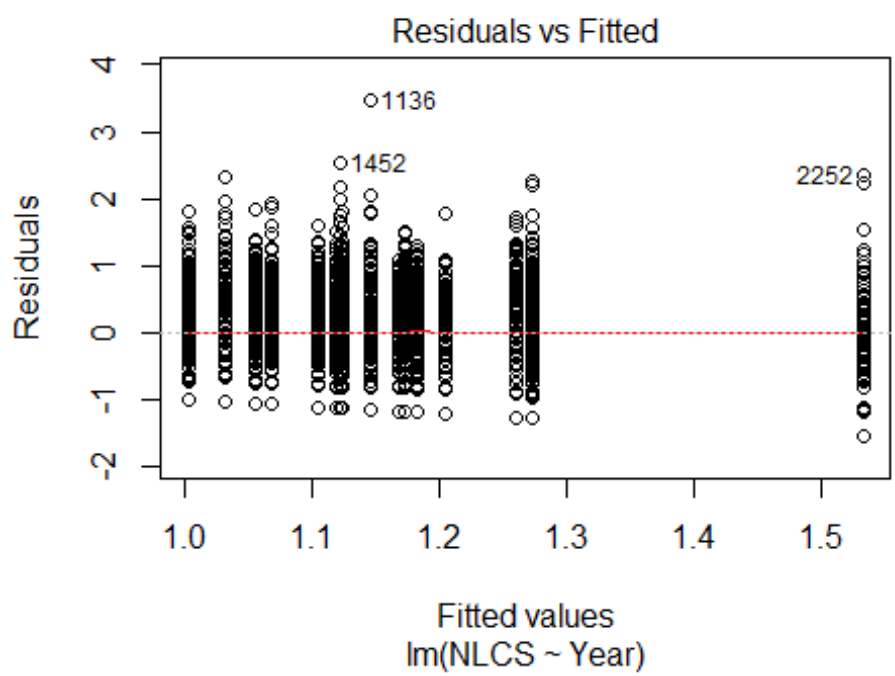
```
## Year2007             0.2693     0.0972    2.77   0.0057 **
## Year2008             0.2153     0.0903    2.38   0.0172 *
## Year2009             0.2029     0.0926    2.19   0.0286 *
## Year2010             0.0778     0.0869    0.90   0.3706
## Year2011             0.2377     0.0860    2.77   0.0058 **
## Year2012             0.1940     0.0923    2.10   0.0357 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.459
## Multiple R-squared:  0.035,  Adjusted R-squared:  0.0204
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
##  93 weights are ~= 1. The remaining 1048 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.149   0.864   0.950   0.899   0.985   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          8.76e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##                   GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.097  1           1.047
## Year             1.097 16           1.003
```



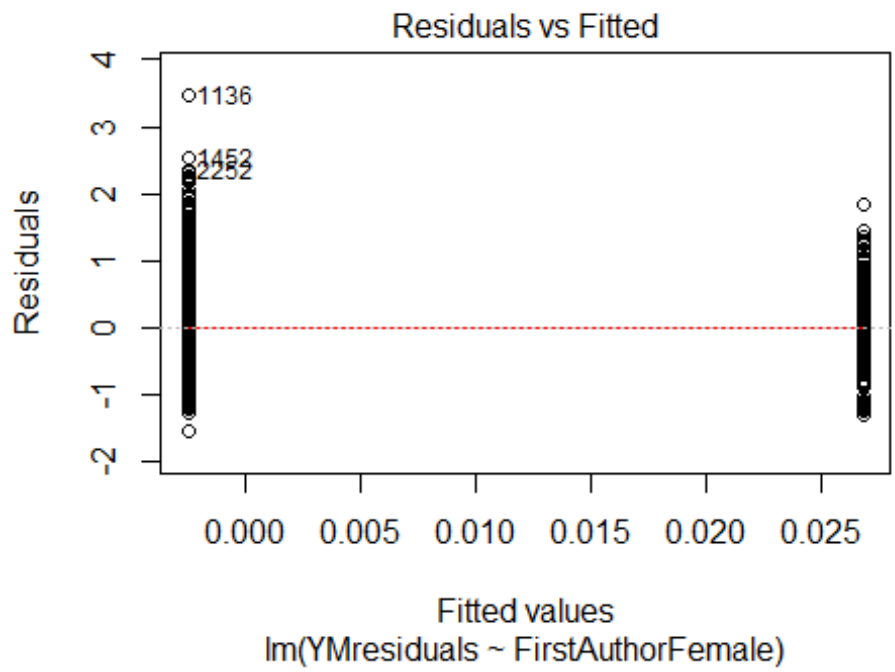

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.3281 -0.3117  0.0126  0.3187  1.7263
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)          0.9257     0.0744   12.45   <2e-16 ***
## LastAuthorFemale1    0.0345     0.0490    0.70   0.4810
## Year1997             0.1095     0.1024    1.07   0.2850
## Year1998             0.0919     0.0964    0.95   0.3403
## Year1999             0.1728     0.1043    1.66   0.0979 .
## Year2000             0.2540     0.1004    2.53   0.0115 *
## Year2001             0.4024     0.0960    4.19    3e-05 ***
## Year2002             0.1888     0.0965    1.96   0.0508 .
## Year2003             0.1652     0.0941    1.76   0.0794 .
## Year2004             0.2724     0.0981    2.78   0.0056 **
## Year2005             0.2692     0.0930    2.89   0.0039 **
## Year2006             0.2257     0.0888    2.54   0.0112 *
```

```
## Year2007            0.2736     0.0976    2.80   0.0052 **
## Year2008            0.2242     0.0899    2.49   0.0128 *
## Year2009            0.2136     0.0925    2.31   0.0211 *
## Year2010            0.0839     0.0862    0.97   0.3308
## Year2011            0.2493     0.0856    2.91   0.0037 **
## Year2012            0.2018     0.0926    2.18   0.0295 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.459
## Multiple R-squared:  0.035,  Adjusted R-squared:  0.0204
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
##  93 weights are ~= 1. The remaining 1048 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.126   0.862   0.950   0.899   0.985   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          8.76e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Sample size for the above analysis:  1141"
## [1] ""
## [1] ""
## [1] "####################################"
## [1] "Analysis of AJSC 1507"
## [1] "####################################"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  497  339  386  350  331  370  365  345  444  429  418  412  379  278  262
## 2011 2012
##  253  278
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  183  142  158  149   98   87  155  145  212  218  194  217  184  133  138
## 2011 2012
```

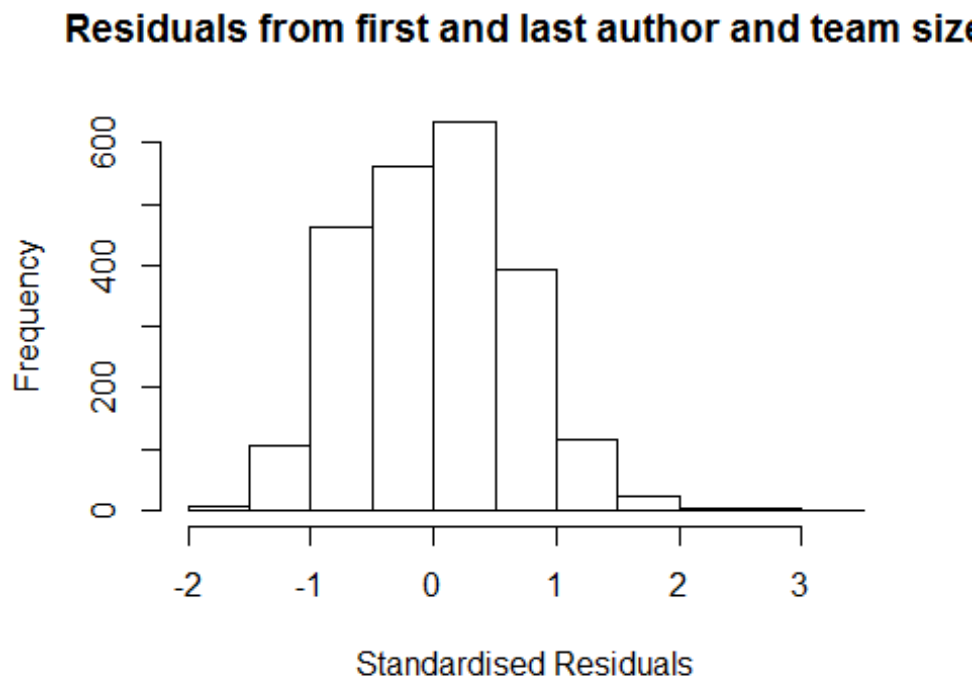
```
##  129  153
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  155  130  142  122   87   74  130  128  185  190  168  184  158  110  117
## 2011 2012
##  104  124
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
##  Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 81, df = 16, p-value = 1e-10
```




```
##
##  Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 7.2, df = 1, p-value = 0.007
```

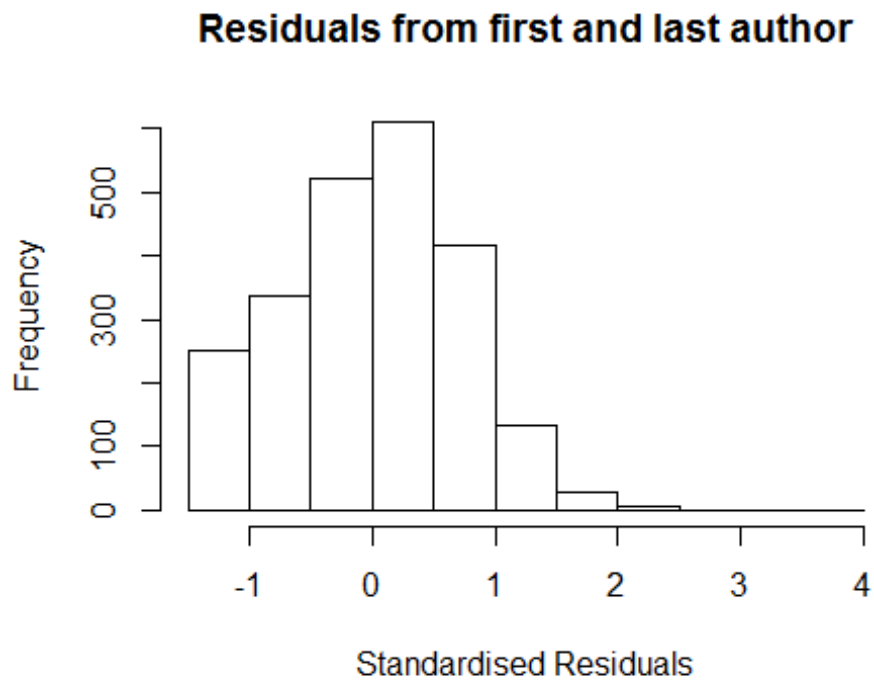



```
## [1] "Female first author team size 2018 geometric mean: 3.14914912692746"
## [1] "Male first author team size 2018 geometric mean: 2.73547895415557"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 9200, p-value = 0.03
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 2.53783676407889"
## [1] "Male last author team size 2018 geometric mean: 2.84806850328446"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 4900, p-value = 0.3
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.116  1           1.056
## LastAuthorFemale  1.110  1           1.053
## UniqueAuthors     1.272  4           1.031
## Year              1.302 16           1.008
```



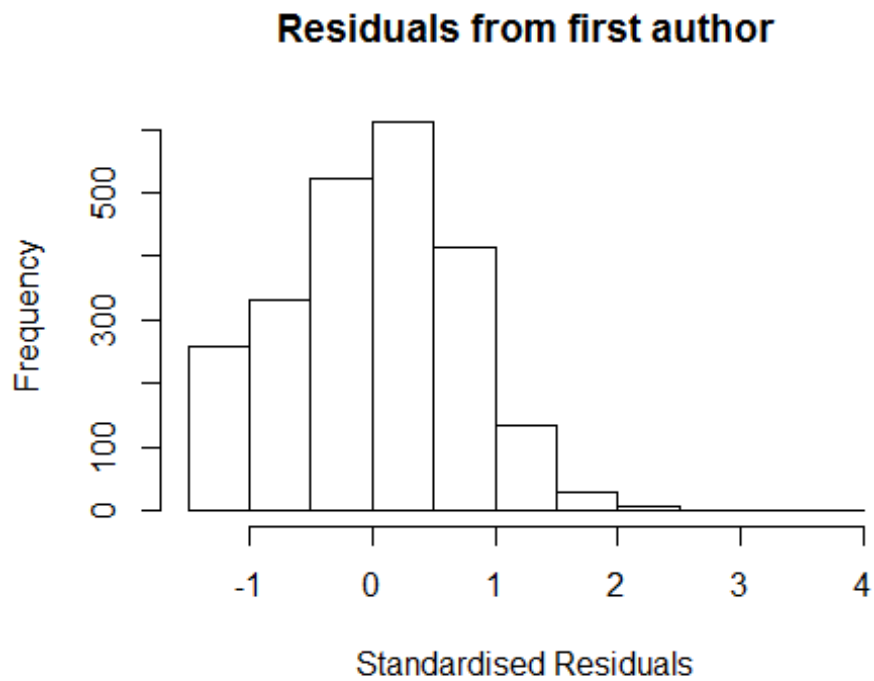

```
## [1] "List of  3 outliers with residuals above  2.5"
##        ScopusId  NLCS Year OneField Fields residuals
## 728  0031104450 3.357 1997     1507      3     2.755
## 1136 0032023438 4.610 1998     1507      1     3.268
## 2628 0036591982 3.537 2002     1500      5     2.634
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
UniqueAuthors +
##     Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##     k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.6077 -0.4992  0.0248  0.4571  3.2682
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)          0.7741     0.0604   12.81   <2e-16 ***
## FirstAuthorFemale1   0.0156     0.0459    0.34   0.7332
## LastAuthorFemale1   -0.0792     0.0505   -1.57   0.1168
## UniqueAuthors2       0.5553     0.0371   14.99   <2e-16 ***
## UniqueAuthors3       0.6020     0.0413   14.57   <2e-16 ***
## UniqueAuthors4       0.7049     0.0593   11.88   <2e-16 ***
## UniqueAuthors5       0.8499     0.0654   12.99   <2e-16 ***
## Year1997            -0.1722     0.0908   -1.90   0.0581 .
```

```
## Year1998            -0.0343     0.0899   -0.38   0.7027
## Year1999            -0.0283     0.0903   -0.31   0.7538
## Year2000             0.0999     0.1103    0.91   0.3652
## Year2001             0.2240     0.0968    2.31   0.0207 *
## Year2002             0.1293     0.0886    1.46   0.1446
## Year2003            -0.0163     0.0814   -0.20   0.8410
## Year2004            -0.1155     0.0698   -1.66   0.0981 .
## Year2005            -0.1259     0.0722   -1.74   0.0813 .
## Year2006            -0.1516     0.0729   -2.08   0.0377 *
## Year2007            -0.2074     0.0723   -2.87   0.0042 **
## Year2008            -0.1330     0.0724   -1.84   0.0663 .
## Year2009             0.0180     0.0795    0.23   0.8213
## Year2010            -0.1119     0.0794   -1.41   0.1589
## Year2011            -0.1054     0.0830   -1.27   0.2041
## Year2012            -0.1365     0.0770   -1.77   0.0764 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.656
## Multiple R-squared:  0.18,   Adjusted R-squared:  0.172
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
##  observation 388 is an outlier with |weight| = 0 ( < 4.3e-05);
##  170 weights are ~= 1. The remaining 2137 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0386  0.8820  0.9430  0.9100  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          4.33e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.081  1           1.040
## LastAuthorFemale  1.087  1           1.042
## Year              1.058 16           1.002
```



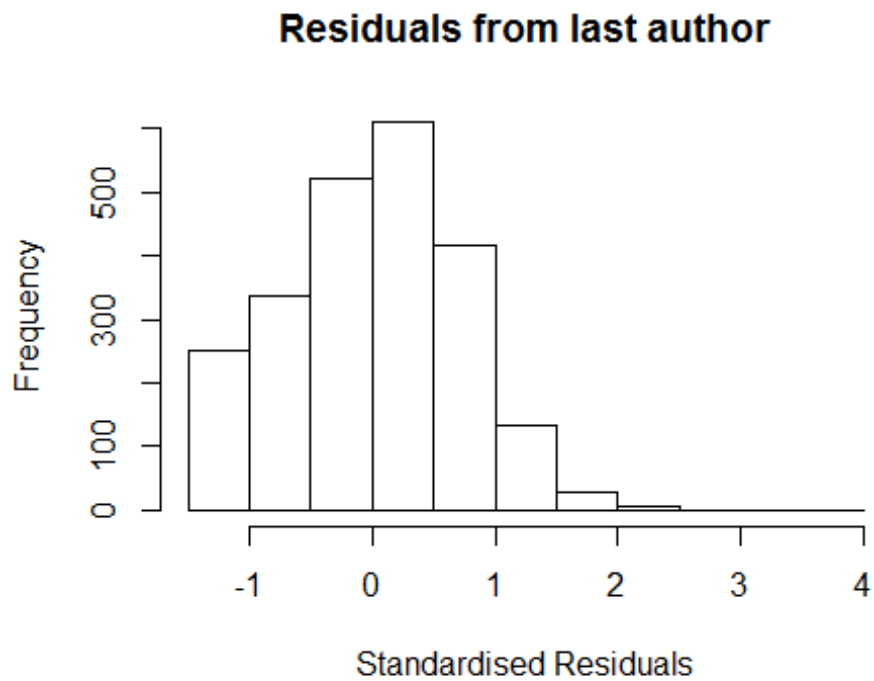

```
## [1] "List of  1 outliers with residuals above  2.5"
##        ScopusId NLCS Year OneField Fields residuals
## 1136 0032023438 4.61 1998     1507      1      3.55
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.4827 -0.5122  0.0351  0.5066  3.5504
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)           1.1391     0.0688   16.55   <2e-16 ***
## FirstAuthorFemale1    0.0640     0.0476    1.34   0.1789
## LastAuthorFemale1    -0.0493     0.0516   -0.95   0.3403
## Year1997             -0.2254     0.1053   -2.14   0.0323 *
## Year1998             -0.0796     0.1017   -0.78   0.4343
## Year1999             -0.1240     0.1062   -1.17   0.2431
## Year2000              0.1001     0.1255    0.80   0.4250
## Year2001              0.3436     0.1074    3.20   0.0014 **
## Year2002              0.0838     0.0981    0.85   0.3929
## Year2003             -0.0672     0.0955   -0.70   0.4818
## Year2004             -0.1416     0.0855   -1.66   0.0977 .
## Year2005             -0.1204     0.0867   -1.39   0.1650
```

```
## Year2006            -0.0752     0.0852   -0.88   0.3772
## Year2007            -0.1757     0.0857   -2.05   0.0406 *
## Year2008            -0.0601     0.0876   -0.69   0.4923
## Year2009             0.1402     0.0907    1.55   0.1224
## Year2010             0.0446     0.0912    0.49   0.6248
## Year2011             0.0623     0.0912    0.68   0.4942
## Year2012             0.0334     0.0873    0.38   0.7026
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.742
## Multiple R-squared:  0.0275, Adjusted R-squared:  0.0199
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
##  observation 388 is an outlier with |weight| = 0 ( < 4.3e-05);
##  197 weights are ~= 1. The remaining 2110 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.257   0.859   0.949   0.916   0.984   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          4.33e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.029  1           1.014
## Year              1.029 16           1.001
```




```
## [1] "List of  1 outliers with residuals above  2.5"
##        ScopusId NLCS Year OneField Fields residuals
## 1136 0032023438 4.61 1998     1507      1      3.55
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.4824 -0.5099  0.0307  0.5076  3.5521
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)           1.1352     0.0684   16.59   <2e-16 ***
## FirstAuthorFemale1    0.0489     0.0472    1.04   0.3001
## Year1997             -0.2223     0.1049   -2.12   0.0342 *
## Year1998             -0.0773     0.1017   -0.76   0.4474
## Year1999             -0.1223     0.1063   -1.15   0.2501
## Year2000              0.1001     0.1253    0.80   0.4244
## Year2001              0.3473     0.1072    3.24   0.0012 **
## Year2002              0.0866     0.0979    0.88   0.3763
## Year2003             -0.0669     0.0955   -0.70   0.4837
## Year2004             -0.1400     0.0853   -1.64   0.1012
## Year2005             -0.1193     0.0868   -1.38   0.1692
## Year2006             -0.0737     0.0851   -0.87   0.3861
```

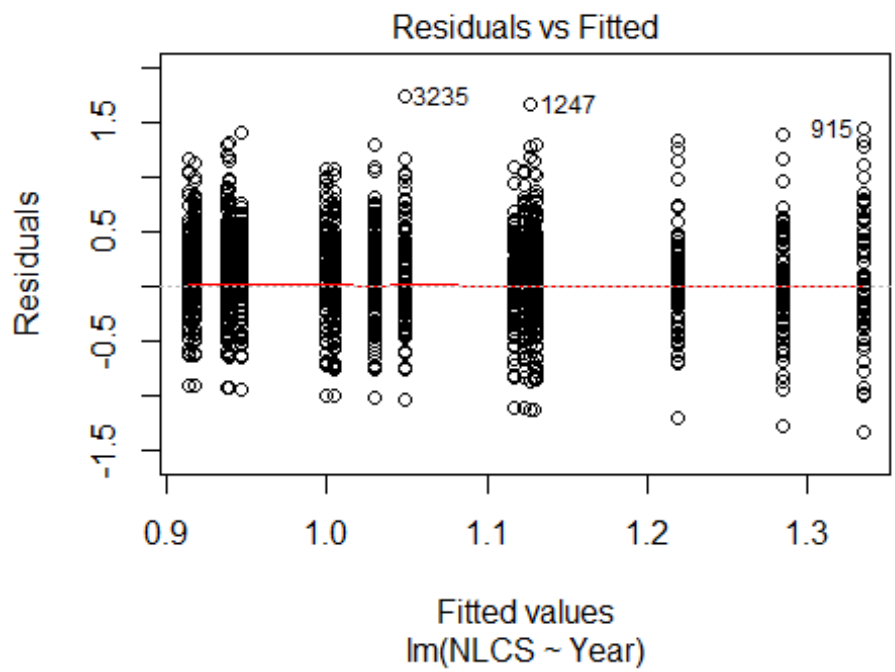
```
## Year2007            -0.1741     0.0857   -2.03   0.0423 *
## Year2008            -0.0600     0.0876   -0.68   0.4935
## Year2009             0.1403     0.0909    1.54   0.1230
## Year2010             0.0443     0.0911    0.49   0.6268
## Year2011             0.0626     0.0912    0.69   0.4920
## Year2012             0.0332     0.0872    0.38   0.7037
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.743
## Multiple R-squared:  0.0272, Adjusted R-squared:  0.02
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
##  observation 388 is an outlier with |weight| = 0 ( < 4.3e-05);
##  193 weights are ~= 1. The remaining 2114 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.257   0.859   0.949   0.916   0.984   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          4.33e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##                   GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.033  1           1.017
## Year             1.033 16           1.001
```



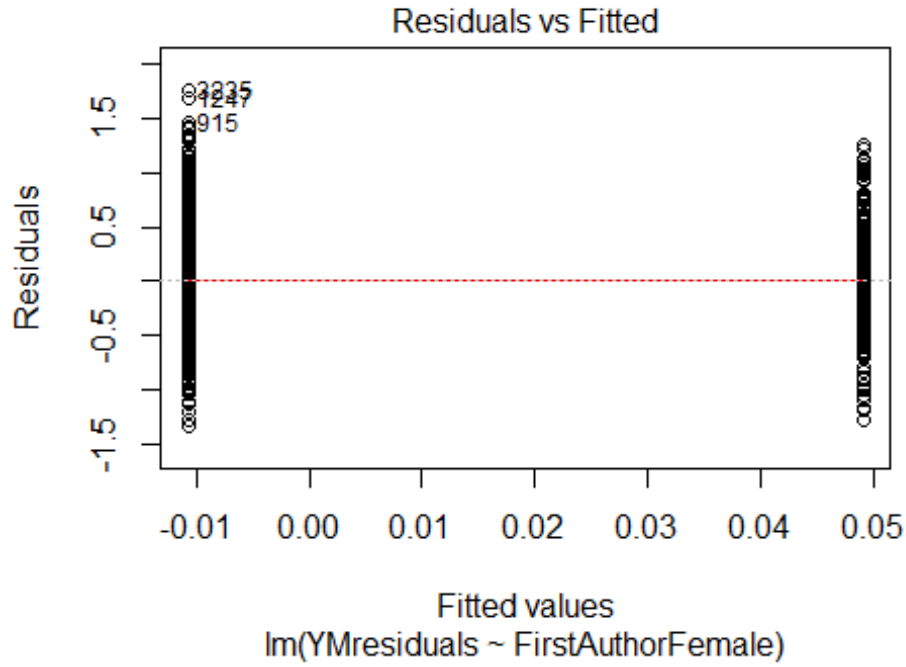

```
## [1] "List of  1 outliers with residuals above  2.5"
##        ScopusId NLCS Year OneField Fields residuals
## 1136 0032023438 4.61 1998     1507      1      3.55
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.4846 -0.5096  0.0375  0.5054  3.5466
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.1404     0.0690   16.54   <2e-16 ***
## LastAuthorFemale1  -0.0284     0.0500   -0.57   0.5695
## Year1997           -0.2255     0.1055   -2.14   0.0327 *
## Year1998           -0.0771     0.1018   -0.76   0.4494
## Year1999           -0.1213     0.1065   -1.14   0.2548
## Year2000            0.1034     0.1255    0.82   0.4101
## Year2001            0.3442     0.1074    3.20   0.0014 **
## Year2002            0.0875     0.0980    0.89   0.3720
## Year2003           -0.0662     0.0955   -0.69   0.4884
## Year2004           -0.1383     0.0855   -1.62   0.1059
## Year2005           -0.1176     0.0868   -1.35   0.1759
## Year2006           -0.0722     0.0851   -0.85   0.3965
```

```
## Year2007           -0.1734     0.0857   -2.02   0.0433 *
## Year2008           -0.0538     0.0874   -0.62   0.5386
## Year2009            0.1427     0.0909    1.57   0.1164
## Year2010            0.0491     0.0913    0.54   0.5909
## Year2011            0.0652     0.0913    0.71   0.4752
## Year2012            0.0374     0.0873    0.43   0.6682
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.742
## Multiple R-squared:  0.0271, Adjusted R-squared:  0.0198
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
##  observation 388 is an outlier with |weight| = 0 ( < 4.3e-05);
##  192 weights are ~= 1. The remaining 2115 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.257   0.860   0.949   0.916   0.984   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          4.33e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                    psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Sample size for the above analysis:  2308"
## [1] ""
## [1] ""
## [1] "####################################"
## [1] "Analysis of AJSC 1508"
## [1] "####################################"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  178  206  153  174  169  190  181  203  214  201  240  195  248  158  251
## 2011 2012
##  252  250
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##   91  105   91   97   76   83   80  122  119  112  141  114  133   91  144
```

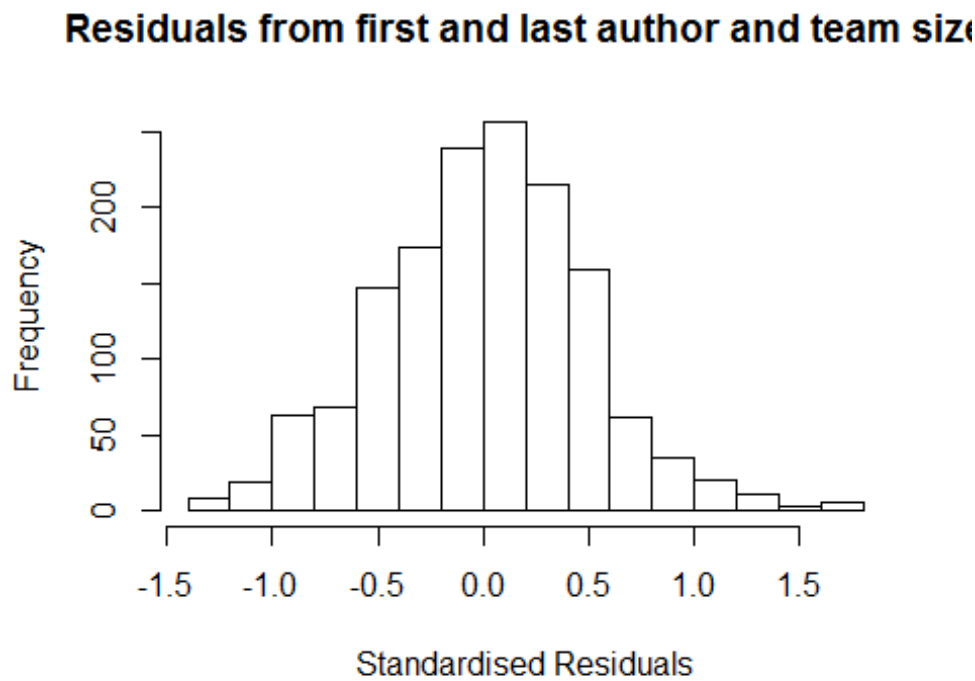
```
## 2011 2012
##  131  133
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##   69   85   77   82   64   66   68   98   93   93  115   85  105   67  110
## 2011 2012
##  107   99
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
##  Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 53, df = 16, p-value = 6e-06
```




```
##
##  Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 3.1, df = 1, p-value = 0.08
```

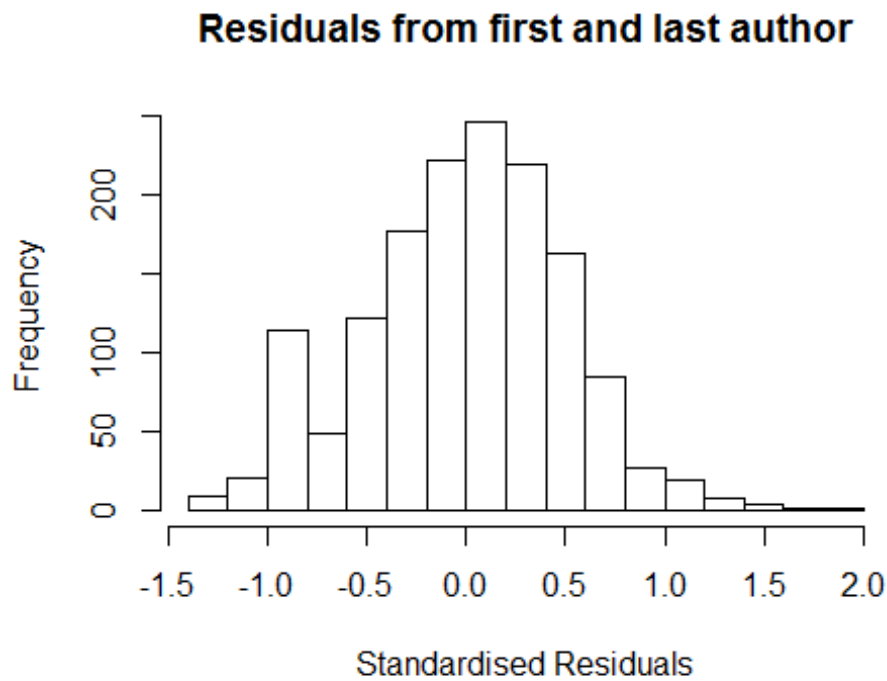



```
## [1] "Female first author team size 2018 geometric mean: 3.90190132671725"
## [1] "Male first author team size 2018 geometric mean: 3.53972842273337"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 6000, p-value = 0.4
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 3.27635450663732"
## [1] "Male last author team size 2018 geometric mean: 3.68677460024484"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 3400, p-value = 0.2
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.067  1           1.033
## LastAuthorFemale  1.055  1           1.027
## UniqueAuthors     1.415  4           1.044
## Year              1.529 16           1.013
```



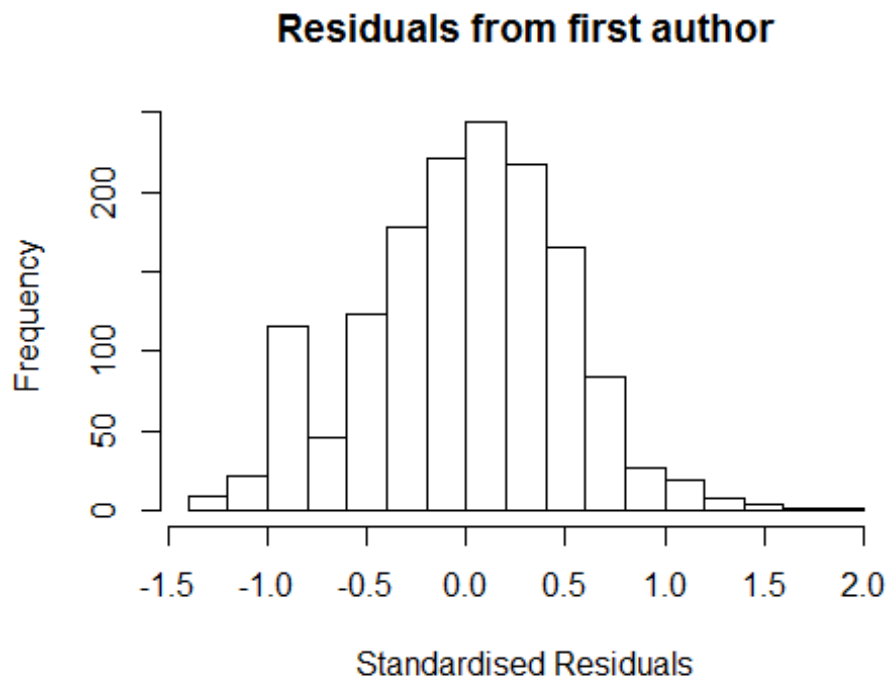

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
UniqueAuthors +
##     Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##     k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.3058 -0.3181  0.0155  0.3165  1.7520
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         0.79395    0.06764   11.74  < 2e-16 ***
## FirstAuthorFemale1  0.05096    0.03156    1.61  0.10662
## LastAuthorFemale1   0.00870    0.03788    0.23  0.81849
## UniqueAuthors2      0.34822    0.05118    6.80  1.5e-11 ***
## UniqueAuthors3      0.41854    0.05096    8.21  4.7e-16 ***
## UniqueAuthors4      0.43429    0.05433    7.99  2.6e-15 ***
## UniqueAuthors5      0.46769    0.05068    9.23  < 2e-16 ***
## Year1997            0.00374    0.07051    0.05  0.95769
## Year1998            0.00769    0.08474    0.09  0.92767
## Year1999            0.04047    0.07731    0.52  0.60069
```

```
## Year2000            0.16365    0.10099    1.62  0.10533
## Year2001            0.13107    0.08311    1.58  0.11502
## Year2002           -0.06696    0.08739   -0.77  0.44369
## Year2003           -0.23048    0.07845   -2.94  0.00336 **
## Year2004           -0.26425    0.08367   -3.16  0.00162 **
## Year2005           -0.14552    0.08164   -1.78  0.07489 .
## Year2006           -0.24379    0.07385   -3.30  0.00099 ***
## Year2007           -0.15415    0.07267   -2.12  0.03406 *
## Year2008           -0.18339    0.07185   -2.55  0.01080 *
## Year2009           -0.09064    0.07452   -1.22  0.22410
## Year2010           -0.20529    0.06934   -2.96  0.00312 **
## Year2011           -0.24726    0.06987   -3.54  0.00041 ***
## Year2012           -0.32817    0.06983   -4.70  2.9e-06 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.454
## Multiple R-squared:  0.157,  Adjusted R-squared:  0.145
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
##  118 weights are ~= 1. The remaining 1365 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.104   0.867   0.946   0.896   0.987   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          6.74e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.072  1           1.035
## LastAuthorFemale  1.043  1           1.021
## Year              1.115 16           1.003
```



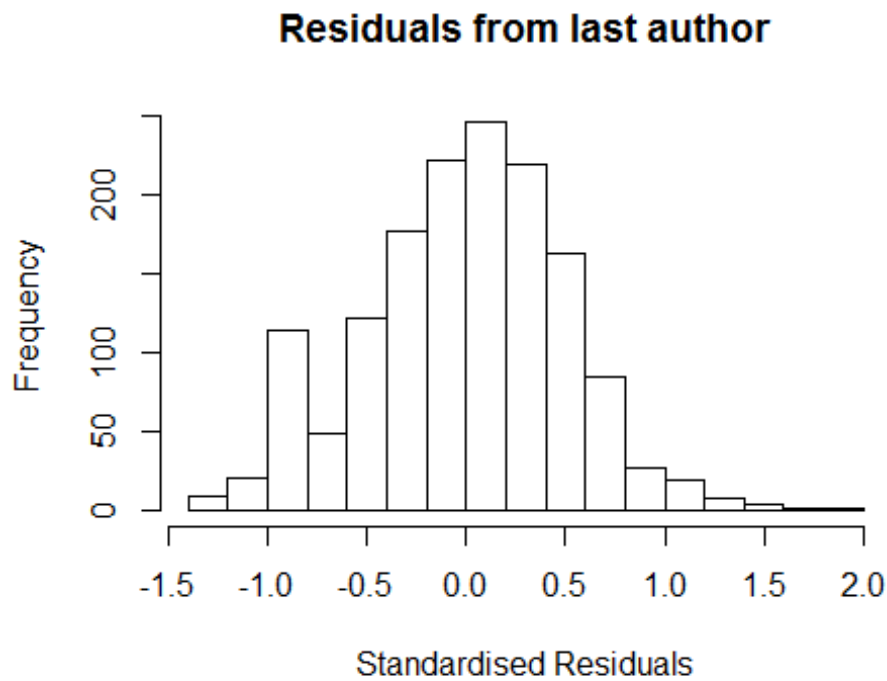

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.2957 -0.3309  0.0199  0.3329  1.8154
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)          1.10132    0.05886   18.71   <2e-16 ***
## FirstAuthorFemale1   0.07863    0.03264    2.41   0.0161 *
## LastAuthorFemale1    0.00276    0.03906    0.07   0.9437
## Year1997             0.01678    0.07267    0.23   0.8174
## Year1998             0.05037    0.08706    0.58   0.5629
## Year1999             0.11051    0.07931    1.39   0.1637
## Year2000             0.19443    0.10200    1.91   0.0568 .
## Year2001             0.16327    0.08016    2.04   0.0419 *
## Year2002            -0.01666    0.08906   -0.19   0.8516
## Year2003            -0.19128    0.08536   -2.24   0.0252 *
## Year2004            -0.26311    0.09358   -2.81   0.0050 **
## Year2005            -0.14231    0.09382   -1.52   0.1295
```

```
## Year2006           -0.22757    0.08296   -2.74   0.0062 **
## Year2007           -0.12326    0.08070   -1.53   0.1269
## Year2008           -0.15525    0.07688   -2.02   0.0436 *
## Year2009           -0.02335    0.07410   -0.32   0.7527
## Year2010           -0.12470    0.06938   -1.80   0.0725 .
## Year2011           -0.18799    0.07053   -2.67   0.0078 **
## Year2012           -0.23647    0.07225   -3.27   0.0011 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.471
## Multiple R-squared:  0.0692, Adjusted R-squared:  0.0577
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
##  124 weights are ~= 1. The remaining 1359 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.104   0.858   0.946   0.895   0.984   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          6.74e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##                   GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.07  1           1.034
## Year              1.07 16           1.002
```




```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.2962 -0.3309  0.0196  0.3325  1.8150
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)           1.1018     0.0581   18.95   <2e-16 ***
## FirstAuthorFemale1    0.0789     0.0327    2.41   0.0160 *
## Year1997              0.0166     0.0726    0.23   0.8187
## Year1998              0.0501     0.0868    0.58   0.5640
## Year1999              0.1104     0.0793    1.39   0.1637
## Year2000              0.1944     0.1020    1.91   0.0568 .
## Year2001              0.1632     0.0801    2.04   0.0418 *
## Year2002             -0.0170     0.0887   -0.19   0.8484
## Year2003             -0.1914     0.0853   -2.24   0.0249 *
## Year2004             -0.2632     0.0935   -2.81   0.0049 **
## Year2005             -0.1426     0.0935   -1.53   0.1274
## Year2006             -0.2276     0.0829   -2.74   0.0061 **
```

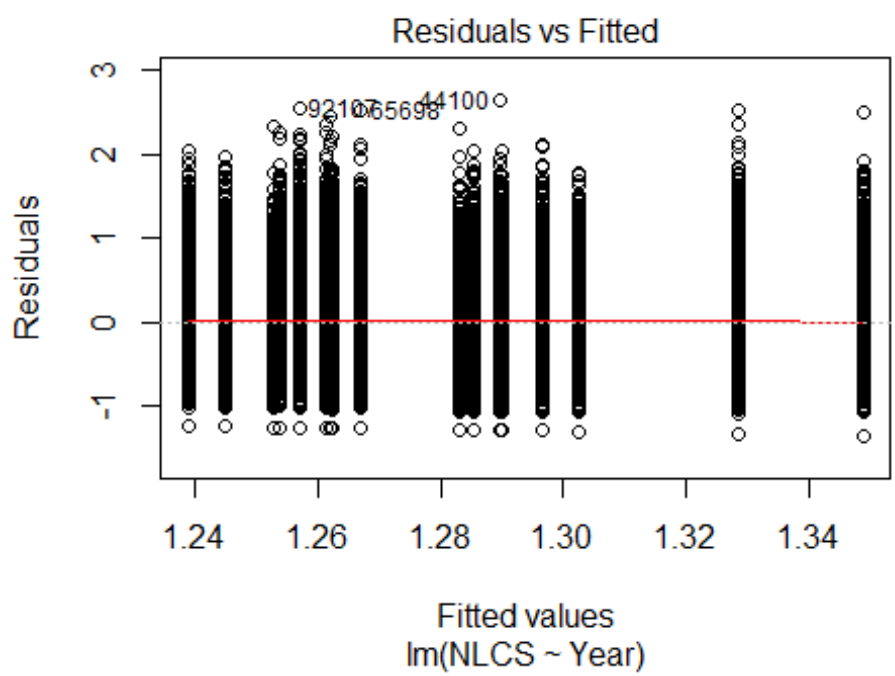
```
## Year2007            -0.1234     0.0805   -1.53   0.1256
## Year2008            -0.1555     0.0766   -2.03   0.0425 *
## Year2009            -0.0235     0.0739   -0.32   0.7505
## Year2010            -0.1247     0.0693   -1.80   0.0723 .
## Year2011            -0.1882     0.0702   -2.68   0.0074 **
## Year2012            -0.2366     0.0721   -3.28   0.0011 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.471
## Multiple R-squared:  0.0692, Adjusted R-squared:  0.0584
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
##  124 weights are ~= 1. The remaining 1359 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.104   0.858   0.946   0.895   0.984   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          6.74e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##                   GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.042  1           1.021
## Year             1.042 16           1.001
```



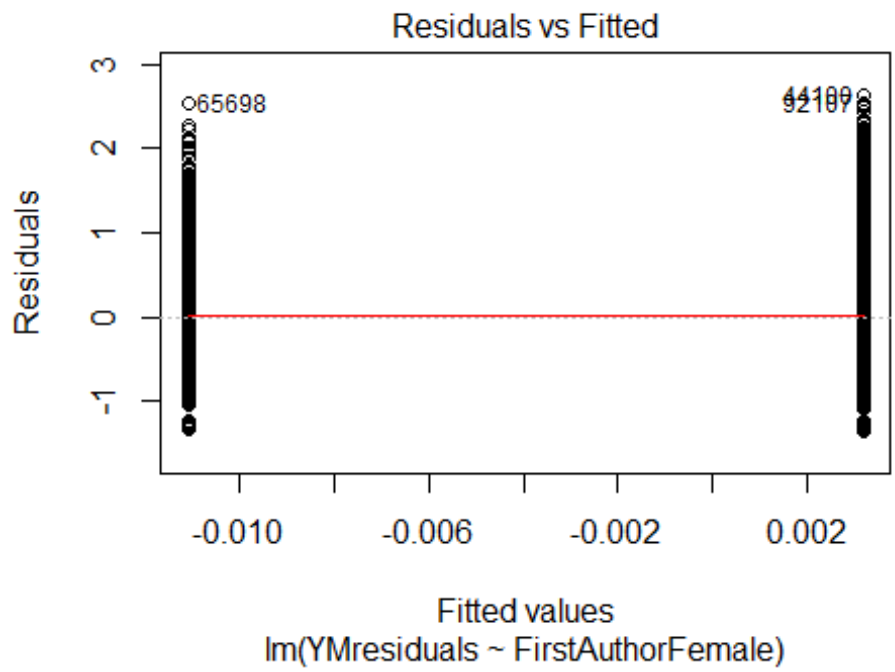

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.3136 -0.3378  0.0227  0.3378  1.7971
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)          1.1132     0.0592   18.80   <2e-16 ***
## LastAuthorFemale1    0.0130     0.0385    0.34   0.7361
## Year1997             0.0107     0.0733    0.15   0.8836
## Year1998             0.0488     0.0880    0.55   0.5795
## Year1999             0.1062     0.0800    1.33   0.1846
## Year2000             0.2003     0.1020    1.96   0.0498 *
## Year2001             0.1613     0.0803    2.01   0.0449 *
## Year2002            -0.0165     0.0889   -0.19   0.8529
## Year2003            -0.1910     0.0862   -2.22   0.0268 *
## Year2004            -0.2571     0.0939   -2.74   0.0063 **
## Year2005            -0.1427     0.0944   -1.51   0.1308
## Year2006            -0.2258     0.0839   -2.69   0.0072 **
```

```
## Year2007           -0.1173     0.0811   -1.45   0.1481
## Year2008           -0.1476     0.0773   -1.91   0.0562 .
## Year2009           -0.0142     0.0742   -0.19   0.8487
## Year2010           -0.1183     0.0698   -1.70   0.0902 .
## Year2011           -0.1890     0.0708   -2.67   0.0077 **
## Year2012           -0.2373     0.0728   -3.26   0.0011 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.47
## Multiple R-squared:  0.0657, Adjusted R-squared:  0.0548
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
##  130 weights are ~= 1. The remaining 1353 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.112   0.858   0.945   0.894   0.984   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          6.74e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Sample size for the above analysis:  1483"
## [1] ""
## [1] ""
## [1] "####################################"
## [1] "Analysis of AJSC 1600"
## [1] "####################################"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 5400 5012 4745 4372 4432 4416 5133 5145 5295 5733 6011 6034 6529 7219 7288
## 2011 2012
## 7697 7599
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 3107 2652 2625 2575 2279 1543 3068 3058 3009 3328 3488 3548 3827 4184 4264
## 2011 2012
```

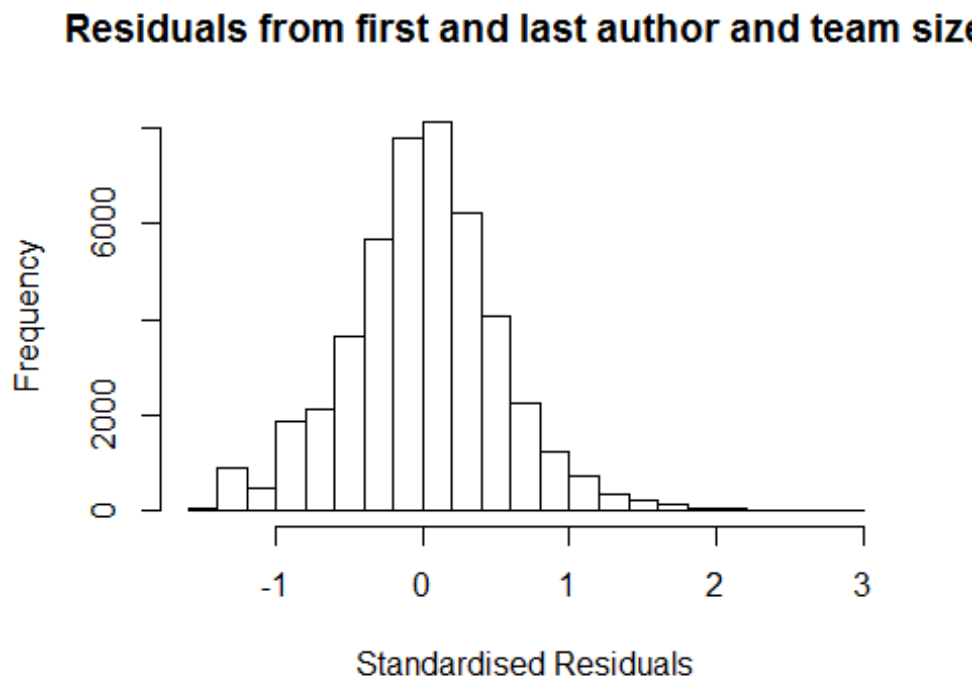
```
## 4475 4405
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 2578 2211 2220 2176 1928 1306 2578 2527 2486 2735 2855 2900 3171 3416 3509
## 2011 2012
## 3607 3574
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
##  Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 210, df = 16, p-value <2e-16
```




```
##
##  Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 23, df = 1, p-value = 1e-06
```



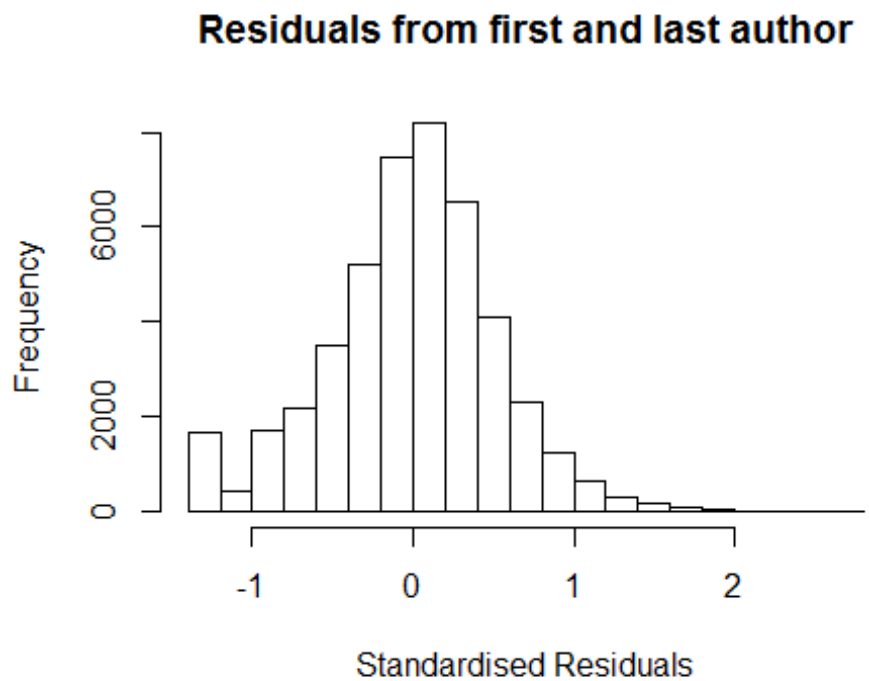

```
## [1] "Female first author team size 2018 geometric mean: 4.65309390891214"
## [1] "Male first author team size 2018 geometric mean: 4.24100701646579"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 1400000, p-value = 2e-04
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 4.1990434527604"
## [1] "Male last author team size 2018 geometric mean: 4.39484591264503"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 970000, p-value = 0.04
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.017  1           1.009
## LastAuthorFemale  1.008  1           1.004
## UniqueAuthors     1.044  4           1.005
## Year              1.050 16           1.002
```




```
## [1] "List of  4 outliers with residuals above  2.5"
##          ScopusId  NLCS Year OneField Fields residuals
## 3052   0345979435 3.843 1996     1600      1     2.869
## 3791   0037571112 3.672 1996     1600      2     2.698
## 44100  4444221565 3.927 2004     1600      2     2.529
## 92107 76149120388 3.793 2010     1600      2     2.558
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
UniqueAuthors +
##     Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##     k.max = 1000))
##  \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.49626 -0.30297  0.00979  0.30966  2.86907
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)          0.97393    0.01480   65.79  < 2e-16 ***
## FirstAuthorFemale1 -0.01729    0.00546   -3.17  0.00155 **
## LastAuthorFemale1  -0.05427    0.00711   -7.63  2.4e-14 ***
## UniqueAuthors2       0.38661    0.01186   32.61  < 2e-16 ***
## UniqueAuthors3       0.39622    0.01176   33.70  < 2e-16 ***
## UniqueAuthors4       0.43007    0.01195   35.99  < 2e-16 ***
## UniqueAuthors5       0.49382    0.01155   42.76  < 2e-16 ***
```

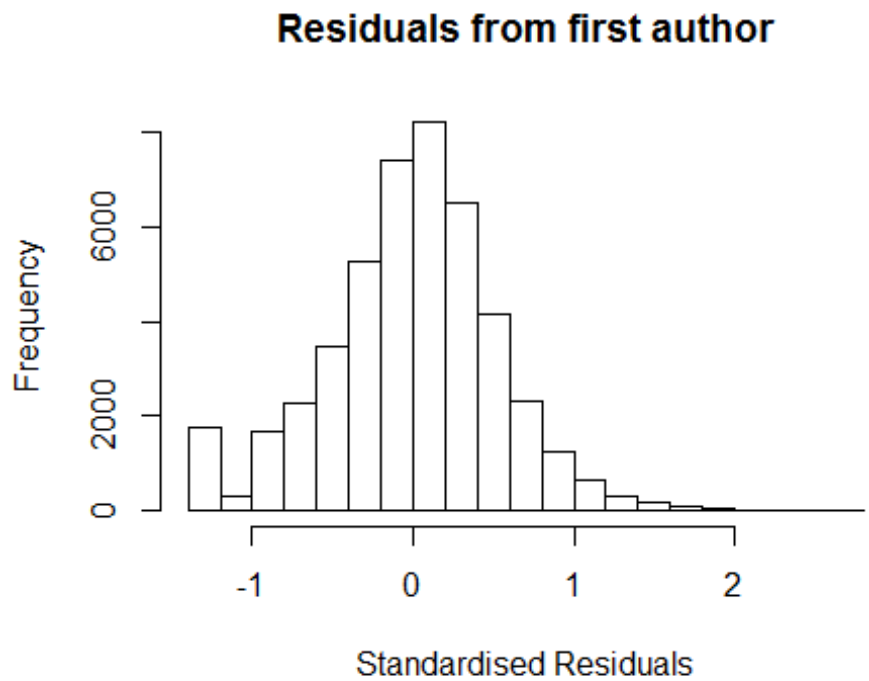
```
## Year1997           -0.02571    0.01580   -1.63  0.10368
## Year1998            0.02851    0.01608    1.77  0.07627 .
## Year1999           -0.04401    0.01580   -2.79  0.00534 **
## Year2000           -0.09268    0.01612   -5.75  8.9e-09 ***
## Year2001           -0.06953    0.01915   -3.63  0.00028 ***
## Year2002           -0.06259    0.01441   -4.34  1.4e-05 ***
## Year2003           -0.07281    0.01431   -5.09  3.6e-07 ***
## Year2004           -0.06942    0.01453   -4.78  1.8e-06 ***
## Year2005           -0.08332    0.01404   -5.93  3.0e-09 ***
## Year2006           -0.09648    0.01424   -6.77  1.3e-11 ***
## Year2007           -0.09713    0.01429   -6.80  1.1e-11 ***
## Year2008           -0.10305    0.01427   -7.22  5.1e-13 ***
## Year2009           -0.10742    0.01408   -7.63  2.5e-14 ***
## Year2010           -0.12550    0.01395   -9.00  < 2e-16 ***
## Year2011           -0.13079    0.01363   -9.59  < 2e-16 ***
## Year2012           -0.11606    0.01359   -8.54  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.461
## Multiple R-squared:  0.078,  Adjusted R-squared:  0.0776
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
##  22 observations
c(376,663,1461,1545,1811,1996,3211,3868,5041,6336,18146,24416,27166,29349,318
60,35483,37659,38172,38252,38254,41728,42188)
##   are outliers with |weight| <= 1.4e-06 ( < 2.2e-06);
##  4072 weights are ~= 1. The remaining 41683 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.000   0.855   0.950   0.887   0.986   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          2.18e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                    psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##                     GVIF Df GVIF^(1/(2*Df))
```

```
## FirstAuthorFemale 1.009  1           1.004
## LastAuthorFemale  1.005  1           1.002
## Year              1.014 16           1.000
```



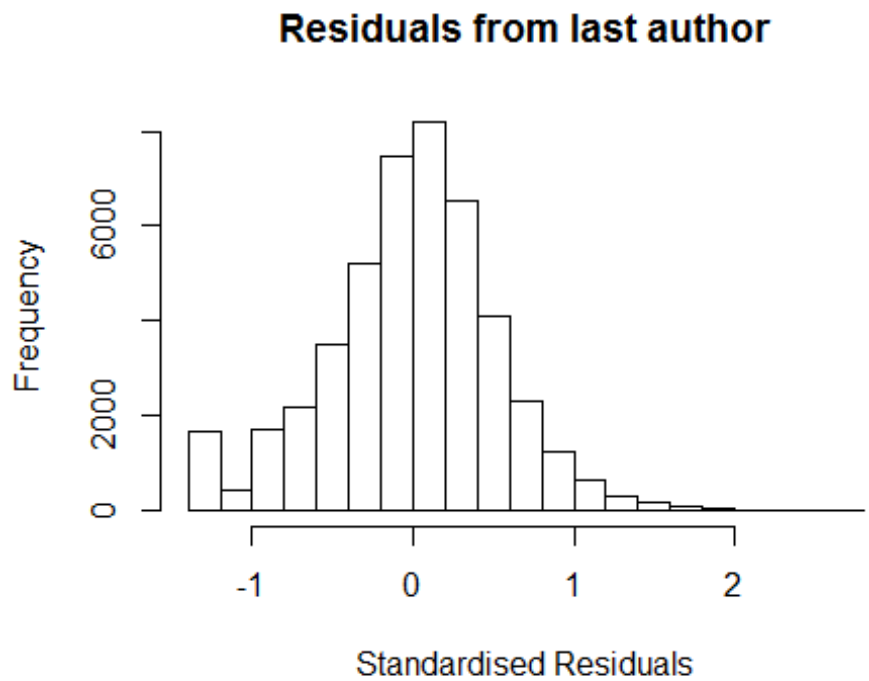

```
## [1] "List of  5 outliers with residuals above  2.5"
##          ScopusId  NLCS Year OneField Fields residuals
## 1421   0030232761 3.847 1996     1502      6     2.514
## 3052   0345979435 3.843 1996     1600      1     2.510
## 44100  4444221565 3.927 2004     1600      2     2.629
## 65698 34249742469 3.790 2007     1600      2     2.513
## 92107 76149120388 3.793 2010     1600      2     2.544
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.3669 -0.3127  0.0189  0.3159  2.6285
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)          1.33286    0.01114  119.61  < 2e-16 ***
## FirstAuthorFemale1  0.00307    0.00562    0.55   0.5855
## LastAuthorFemale1  -0.05866    0.00756   -7.76  9.0e-15 ***
## Year1997           -0.02433    0.01612   -1.51   0.1312
```

```
## Year1998            0.03100    0.01619    1.92   0.0554 .
## Year1999           -0.03841    0.01610   -2.39   0.0170 *
## Year2000           -0.08881    0.01654   -5.37  7.9e-08 ***
## Year2001           -0.07746    0.01935   -4.00  6.2e-05 ***
## Year2002           -0.02777    0.01460   -1.90   0.0572 .
## Year2003           -0.04339    0.01443   -3.01   0.0026 **
## Year2004           -0.03439    0.01491   -2.31   0.0211 *
## Year2005           -0.04631    0.01446   -3.20   0.0014 **
## Year2006           -0.06131    0.01465   -4.18  2.9e-05 ***
## Year2007           -0.05912    0.01468   -4.03  5.7e-05 ***
## Year2008           -0.06216    0.01467   -4.24  2.3e-05 ***
## Year2009           -0.06617    0.01442   -4.59  4.5e-06 ***
## Year2010           -0.08362    0.01438   -5.82  6.1e-09 ***
## Year2011           -0.08618    0.01392   -6.19  6.0e-10 ***
## Year2012           -0.06375    0.01389   -4.59  4.5e-06 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.469
## Multiple R-squared:  0.00519,    Adjusted R-squared:  0.0048
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  12 observations
c(663,1461,1811,5041,18146,27166,29349,31860,35483,37659,38172,45221)
##   are outliers with |weight| = 0 ( < 2.2e-06);
##  4041 weights are ~= 1. The remaining 41724 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0001  0.8510  0.9500  0.8850  0.9860  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          2.18e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.009  1           1.004
## Year              1.009 16           1.000
```




```
## [1] "List of  5 outliers with residuals above  2.5"
##          ScopusId  NLCS Year OneField Fields residuals
## 1421   0030232761 3.847 1996     1502      6     2.514
## 3052   0345979435 3.843 1996     1600      1     2.510
## 44100  4444221565 3.927 2004     1600      2     2.629
## 65698 34249742469 3.790 2007     1600      2     2.513
## 92107 76149120388 3.793 2010     1600      2     2.544
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.3581 -0.3120  0.0177  0.3152  2.6347
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.32801    0.01113  119.33  < 2e-16 ***
## FirstAuthorFemale1 -0.00229    0.00569   -0.40  0.68770
## Year1997           -0.02468    0.01614   -1.53  0.12609
## Year1998            0.03005    0.01619    1.86  0.06349 .
## Year1999           -0.03970    0.01611   -2.46  0.01372 *
## Year2000           -0.09018    0.01654   -5.45  5.0e-08 ***
## Year2001           -0.08014    0.01936   -4.14  3.5e-05 ***
## Year2002           -0.02816    0.01460   -1.93  0.05383 .
```

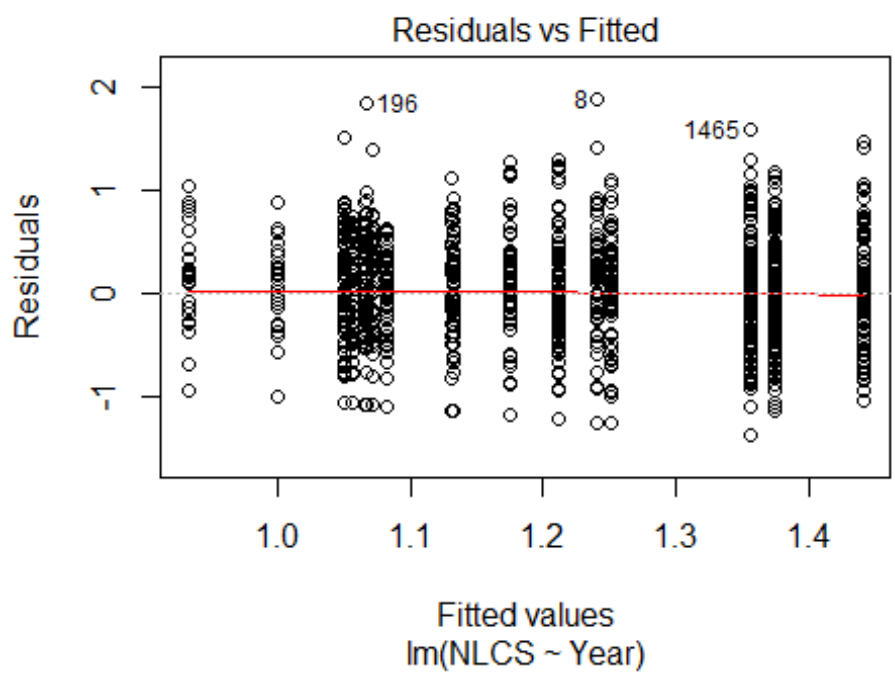
```
## Year2003           -0.04441    0.01443   -3.08  0.00210 **
## Year2004           -0.03570    0.01492   -2.39  0.01671 *
## Year2005           -0.04776    0.01447   -3.30  0.00097 ***
## Year2006           -0.06274    0.01465   -4.28  1.9e-05 ***
## Year2007           -0.06097    0.01469   -4.15  3.3e-05 ***
## Year2008           -0.06397    0.01469   -4.36  1.3e-05 ***
## Year2009           -0.06744    0.01443   -4.67  3.0e-06 ***
## Year2010           -0.08620    0.01439   -5.99  2.1e-09 ***
## Year2011           -0.08924    0.01392   -6.41  1.4e-10 ***
## Year2012           -0.06734    0.01387   -4.85  1.2e-06 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.469
## Multiple R-squared:  0.00361,    Adjusted R-squared:  0.00324
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  12 observations
c(663,1461,1811,5041,18146,27166,29349,31860,35483,37659,38172,45221)
##   are outliers with |weight| = 0 ( < 2.2e-06);
##  4076 weights are ~= 1. The remaining 41689 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.000   0.851   0.950   0.885   0.986   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          2.18e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##                   GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.005  1           1.003
## Year             1.005 16           1.000
```



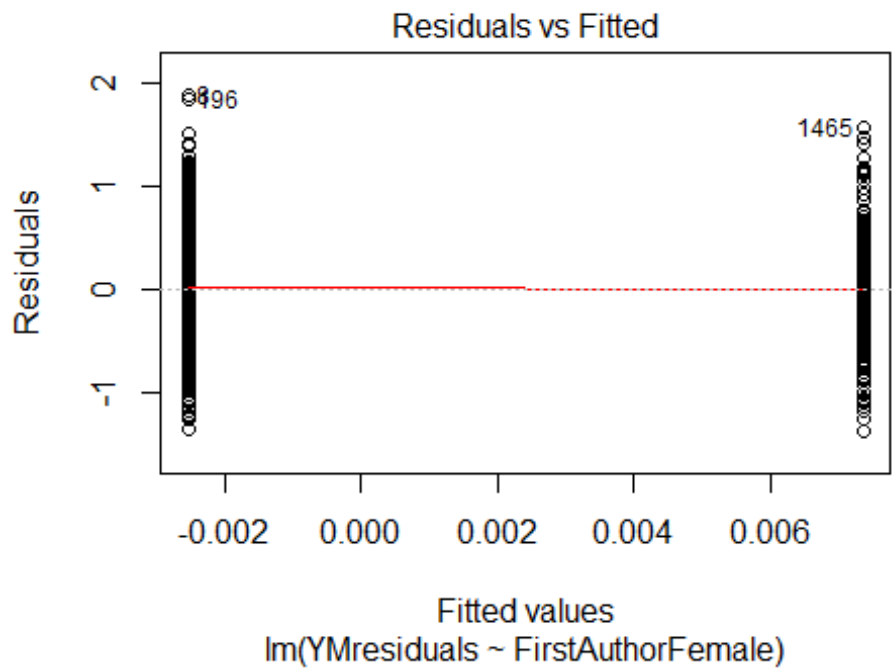

```
## [1] "List of  5 outliers with residuals above  2.5"
##          ScopusId  NLCS Year OneField Fields residuals
## 1421   0030232761 3.847 1996     1502      6     2.514
## 3052   0345979435 3.843 1996     1600      1     2.510
## 44100  4444221565 3.927 2004     1600      2     2.629
## 65698 34249742469 3.790 2007     1600      2     2.513
## 92107 76149120388 3.793 2010     1600      2     2.544
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.3644 -0.3124  0.0186  0.3160  2.6279
##
## Coefficients:
##                   Estimate Std. Error t value Pr(>|t|)
## (Intercept)        1.33337    0.01109  120.26  < 2e-16 ***
## LastAuthorFemale1 -0.05821    0.00761   -7.65  2.1e-14 ***
## Year1997          -0.02432    0.01612   -1.51   0.1315
## Year1998           0.03104    0.01619    1.92   0.0551 .
## Year1999          -0.03838    0.01610   -2.38   0.0171 *
## Year2000          -0.08873    0.01654   -5.37  8.1e-08 ***
## Year2001          -0.07736    0.01934   -4.00  6.4e-05 ***
## Year2002          -0.02770    0.01460   -1.90   0.0579 .
```

```
## Year2003          -0.04323    0.01443   -3.00   0.0027 **
## Year2004          -0.03423    0.01491   -2.30   0.0217 *
## Year2005          -0.04614    0.01447   -3.19   0.0014 **
## Year2006          -0.06116    0.01465   -4.17  3.0e-05 ***
## Year2007          -0.05898    0.01469   -4.02  5.9e-05 ***
## Year2008          -0.06191    0.01466   -4.22  2.4e-05 ***
## Year2009          -0.06592    0.01441   -4.57  4.8e-06 ***
## Year2010          -0.08338    0.01437   -5.80  6.6e-09 ***
## Year2011          -0.08596    0.01392   -6.18  6.6e-10 ***
## Year2012          -0.06347    0.01388   -4.57  4.8e-06 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.469
## Multiple R-squared:  0.00519,    Adjusted R-squared:  0.00482
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  12 observations
c(663,1461,1811,5041,18146,27166,29349,31860,35483,37659,38172,45221)
##   are outliers with |weight| = 0 ( < 2.2e-06);
##  4036 weights are ~= 1. The remaining 41729 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0001  0.8510  0.9500  0.8850  0.9860  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          2.18e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Sample size for the above analysis:  45777"
## [1] ""
## [1] ""
## [1] "####################################"
## [1] "Analysis of AJSC 1601"
## [1] "####################################"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
```

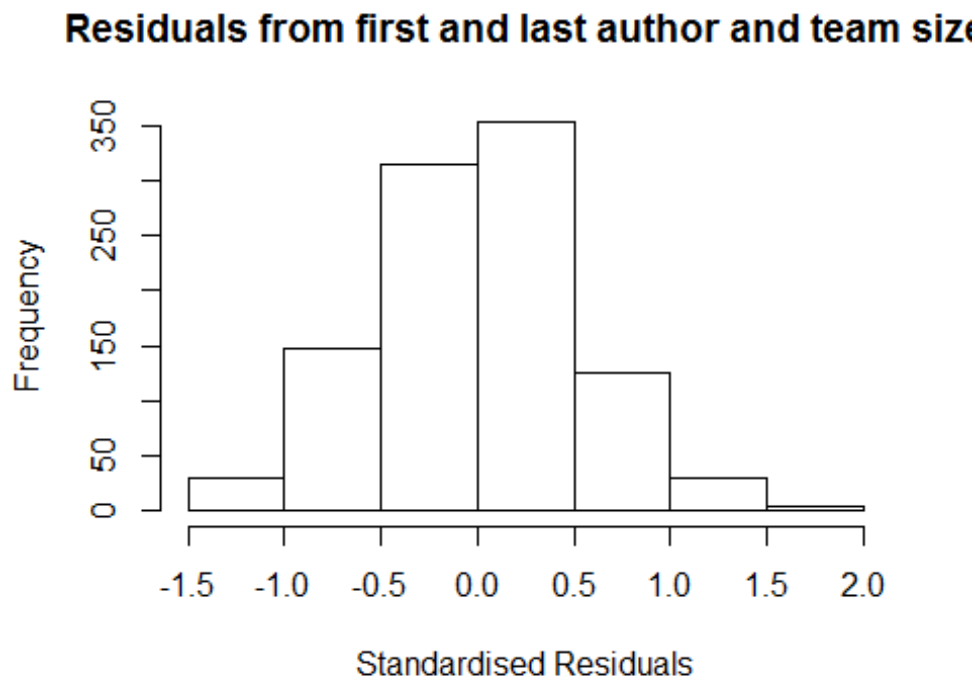
```
##  133   90   78   96  116  120  121  132  159  226  231  212  135   59   71
## 2011 2012
##   87   86
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##   65   41   26   58   60   37   60   74   93  120  144  136   78   32   48
## 2011 2012
##   61   52
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##   61   37   24   53   53   29   50   65   68   95  118  109   68   29   43
## 2011 2012
##   54   47
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
##  Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 33, df = 16, p-value = 0.008
```




```
##
##  Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 1.2, df = 1, p-value = 0.3
```

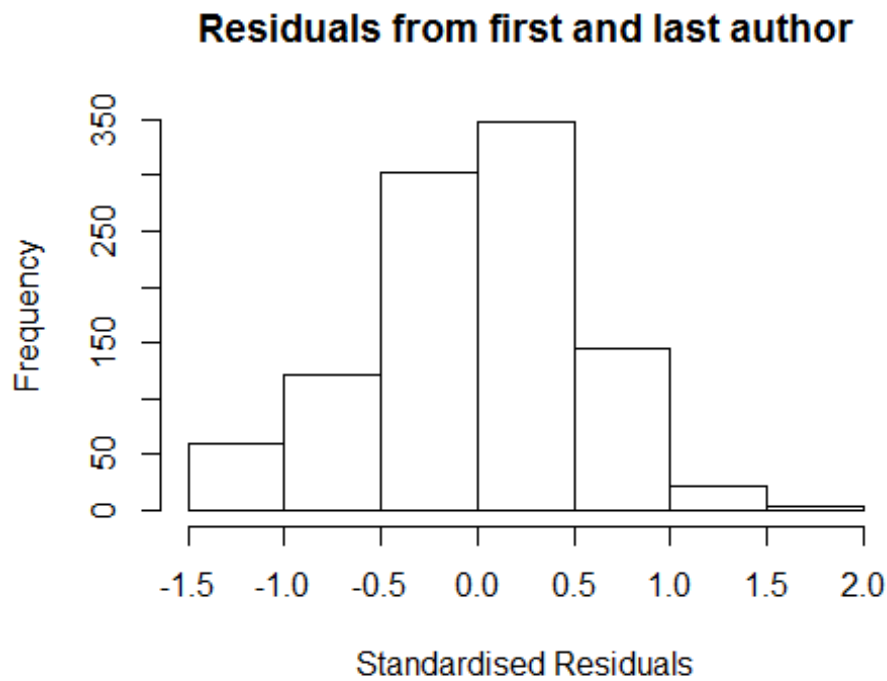



```
## [1] "Female first author team size 2018 geometric mean: 3.48305325532075"
## [1] "Male first author team size 2018 geometric mean: 3.49753670839528"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 3500, p-value = 0.8
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 2.77659065344802"
## [1] "Male last author team size 2018 geometric mean: 3.72454554052923"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 1900, p-value = 0.006
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.138  1           1.067
## LastAuthorFemale  1.094  1           1.046
## UniqueAuthors     1.537  4           1.055
## Year              1.594 16           1.015
```



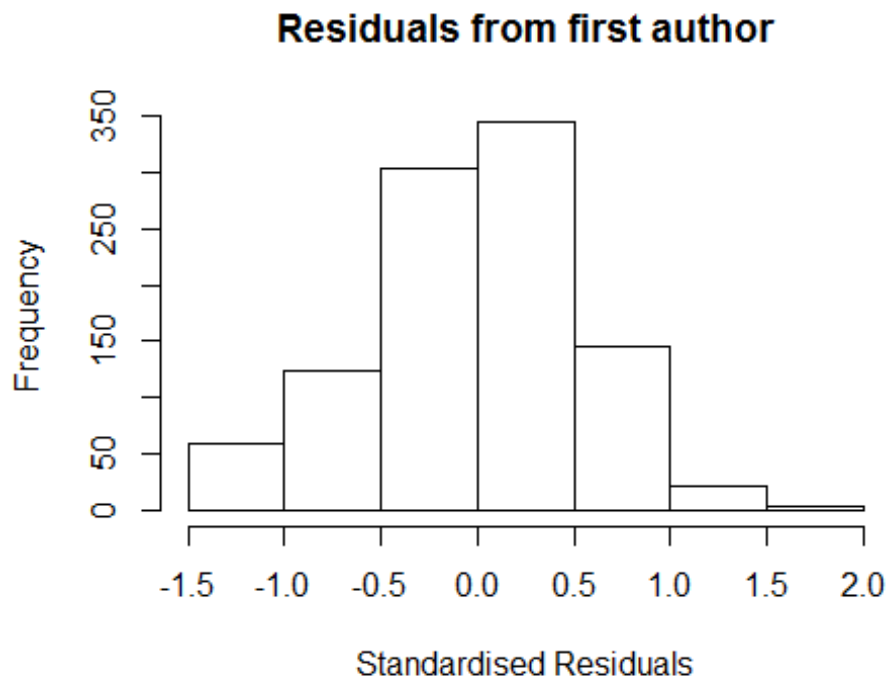

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
UniqueAuthors +
##     Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##     k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.3991 -0.3273  0.0108  0.3515  1.7200
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)           0.8843     0.0888    9.96  < 2e-16 ***
## FirstAuthorFemale1   -0.0446     0.0395   -1.13    0.258
## LastAuthorFemale1     0.0353     0.0411    0.86    0.392
## UniqueAuthors2        0.4290     0.0599    7.16  1.5e-12 ***
## UniqueAuthors3        0.4115     0.0596    6.90  9.3e-12 ***
## UniqueAuthors4        0.5362     0.0603    8.89  < 2e-16 ***
## UniqueAuthors5        0.5148     0.0612    8.41  < 2e-16 ***
## Year1997             -0.1173     0.1457   -0.81    0.421
## Year1998             -0.0745     0.1601   -0.47    0.642
## Year1999             -0.0639     0.1171   -0.55    0.585
```

```
## Year2000            -0.1358     0.1136   -1.20    0.232
## Year2001            -0.3443     0.1456   -2.37    0.018 *
## Year2002            -0.1903     0.0951   -2.00    0.046 *
## Year2003            -0.1668     0.1106   -1.51    0.132
## Year2004            -0.1018     0.1003   -1.01    0.310
## Year2005             0.0687     0.0911    0.75    0.451
## Year2006             0.0135     0.0923    0.15    0.884
## Year2007             0.0283     0.0909    0.31    0.756
## Year2008            -0.0615     0.1007   -0.61    0.542
## Year2009            -0.2045     0.1086   -1.88    0.060 .
## Year2010            -0.1899     0.1082   -1.76    0.080 .
## Year2011            -0.2031     0.0980   -2.07    0.038 *
## Year2012            -0.1158     0.1066   -1.09    0.278
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.505
## Multiple R-squared:  0.154,  Adjusted R-squared:  0.135
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
##  86 weights are ~= 1. The remaining 917 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.223   0.864   0.950   0.902   0.987   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          9.97e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.080  1           1.039
## LastAuthorFemale  1.059  1           1.029
## Year              1.130 16           1.004
```



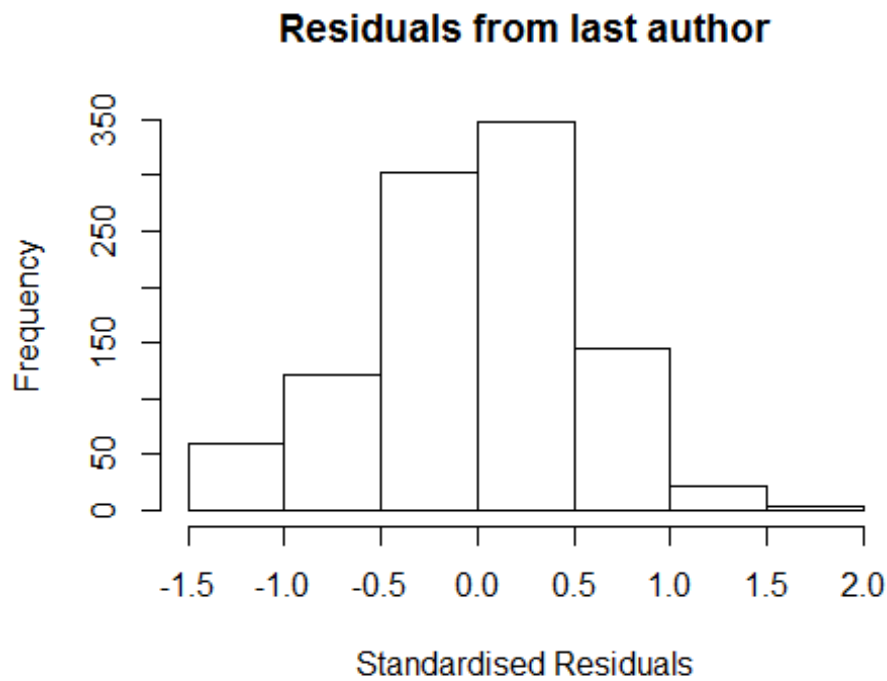

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.3229 -0.3607  0.0172  0.3650  1.8922
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)          1.22576    0.08747   14.01   <2e-16 ***
## FirstAuthorFemale1   0.01840    0.04038    0.46    0.649
## LastAuthorFemale1    0.05298    0.04343    1.22    0.223
## Year1997            -0.15052    0.14295   -1.05    0.293
## Year1998            -0.12826    0.17020   -0.75    0.451
## Year1999            -0.04068    0.12160   -0.33    0.738
## Year2000            -0.19600    0.11997   -1.63    0.103
## Year2001            -0.33075    0.15901   -2.08    0.038 *
## Year2002            -0.15351    0.10348   -1.48    0.138
## Year2003            -0.18240    0.12146   -1.50    0.133
## Year2004            -0.04004    0.10876   -0.37    0.713
## Year2005             0.14639    0.09822    1.49    0.136
```

```
## Year2006            0.07871    0.10006    0.79    0.432
## Year2007            0.10035    0.09999    1.00    0.316
## Year2008            0.00401    0.11152    0.04    0.971
## Year2009           -0.20425    0.12117   -1.69    0.092 .
## Year2010           -0.18541    0.12130   -1.53    0.127
## Year2011           -0.18323    0.10933   -1.68    0.094 .
## Year2012           -0.09798    0.11648   -0.84    0.400
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.536
## Multiple R-squared:  0.0534, Adjusted R-squared:  0.0361
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
##  91 weights are ~= 1. The remaining 912 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.187   0.867   0.952   0.903   0.986   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          9.97e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.071  1           1.035
## Year              1.071 16           1.002
```




```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.3378 -0.3619  0.0132  0.3602  1.8830
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.23502    0.08675   14.24   <2e-16 ***
## FirstAuthorFemale1  0.02501    0.04029    0.62    0.535
## Year1997           -0.15300    0.14271   -1.07    0.284
## Year1998           -0.13558    0.16877   -0.80    0.422
## Year1999           -0.04506    0.12126   -0.37    0.710
## Year2000           -0.19604    0.12043   -1.63    0.104
## Year2001           -0.33139    0.15914   -2.08    0.038 *
## Year2002           -0.15369    0.10336   -1.49    0.137
## Year2003           -0.18352    0.12196   -1.50    0.133
## Year2004           -0.04496    0.10853   -0.41    0.679
## Year2005            0.14201    0.09825    1.45    0.149
## Year2006            0.07782    0.10012    0.78    0.437
```

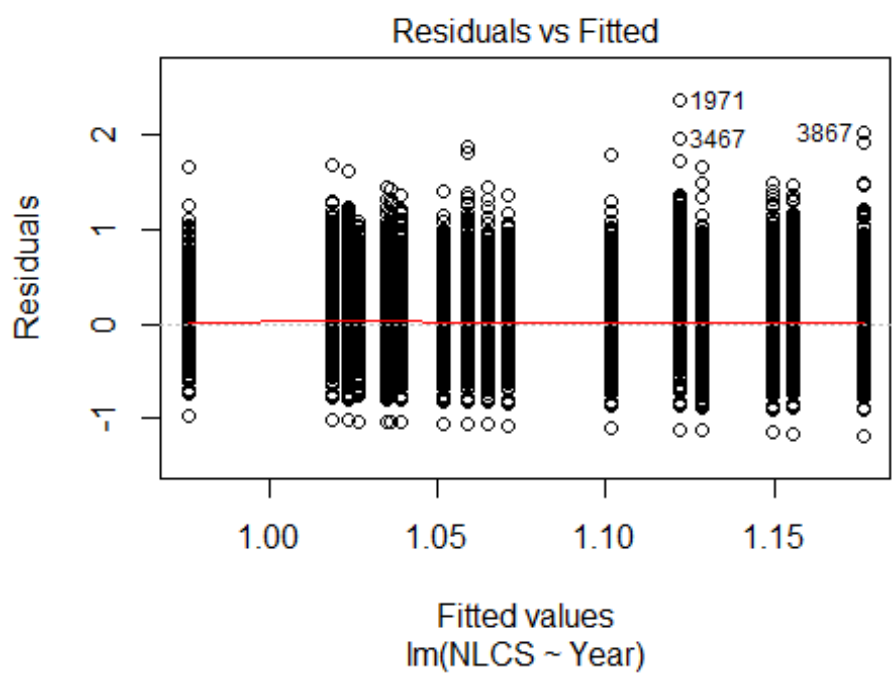
```
## Year2007            0.09877    0.10006    0.99    0.324
## Year2008            0.00805    0.11173    0.07    0.943
## Year2009           -0.20557    0.12064   -1.70    0.089 .
## Year2010           -0.18470    0.12197   -1.51    0.130
## Year2011           -0.18130    0.10965   -1.65    0.099 .
## Year2012           -0.09296    0.11653   -0.80    0.425
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.536
## Multiple R-squared:  0.0523, Adjusted R-squared:  0.036
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
##  104 weights are ~= 1. The remaining 899 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.192   0.866   0.948   0.902   0.985   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          9.97e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##                   GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.053  1           1.026
## Year             1.053 16           1.002
```



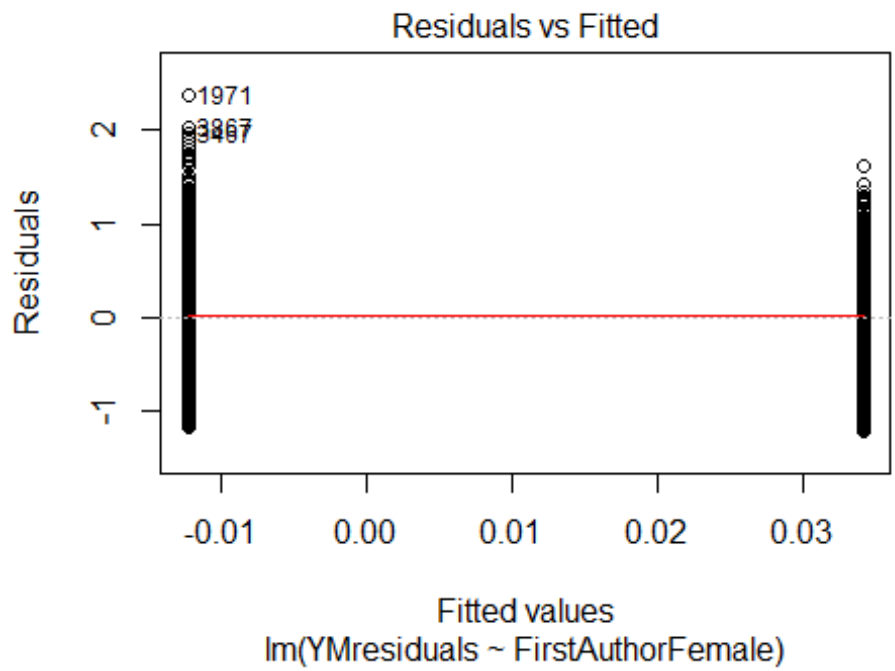

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.3081 -0.3549  0.0194  0.3622  1.8892
##
## Coefficients:
##                   Estimate Std. Error t value Pr(>|t|)
## (Intercept)        1.22884    0.08727   14.08   <2e-16 ***
## LastAuthorFemale1  0.05584    0.04332    1.29    0.198
## Year1997          -0.14873    0.14322   -1.04    0.299
## Year1998          -0.12823    0.17107   -0.75    0.454
## Year1999          -0.03975    0.12182   -0.33    0.744
## Year2000          -0.19374    0.11979   -1.62    0.106
## Year2001          -0.33116    0.15963   -2.07    0.038 *
## Year2002          -0.15223    0.10369   -1.47    0.142
## Year2003          -0.18108    0.12110   -1.50    0.135
## Year2004          -0.04020    0.10877   -0.37    0.712
## Year2005           0.14697    0.09831    1.50    0.135
## Year2006           0.07929    0.10006    0.79    0.428
```

```
## Year2007           0.10102    0.10002    1.01    0.313
## Year2008           0.00617    0.11180    0.06    0.956
## Year2009          -0.20315    0.12088   -1.68    0.093 .
## Year2010          -0.18286    0.12096   -1.51    0.131
## Year2011          -0.17869    0.10885   -1.64    0.101
## Year2012          -0.09634    0.11660   -0.83    0.409
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.536
## Multiple R-squared:  0.0531, Adjusted R-squared:  0.0368
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
##  93 weights are ~= 1. The remaining 910 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.189   0.867   0.952   0.903   0.986   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          9.97e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Sample size for the above analysis:  1003"
## [1] ""
## [1] ""
## [1] "####################################"
## [1] "Analysis of AJSC 1602"
## [1] "####################################"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 1663 1878 1632 1738 1474 1534 1795 1748 1680 2015 1996 1987 1987 2054 1961
## 2011 2012
## 2014 1964
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 1036 1196 1035 1062  762  585 1134 1111 1027 1169 1243 1266 1217 1291 1241
## 2011 2012
```

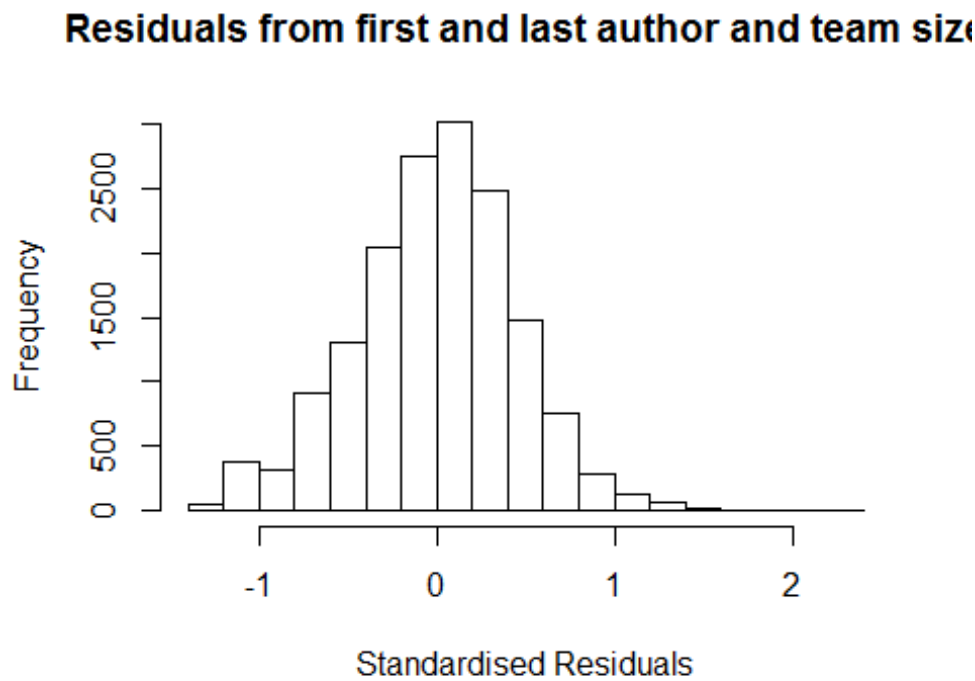
```
## 1252 1221
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  904 1044  884  899  654  494  955  939  851  994 1039 1055 1027 1084 1033
## 2011 2012
## 1054 1032
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
##  Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 170, df = 16, p-value <2e-16
```




```
##
##  Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 45, df = 1, p-value = 2e-11
```

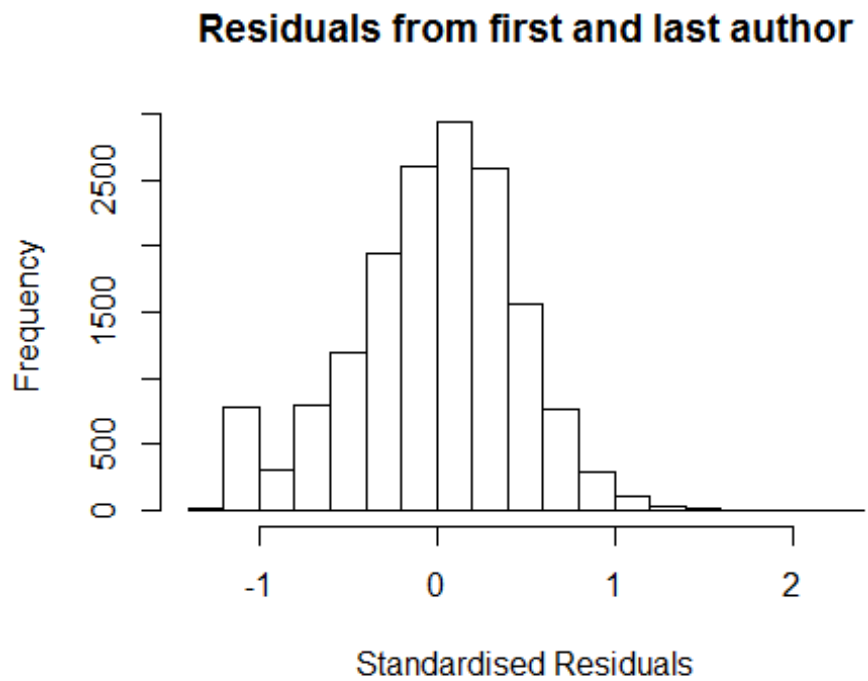



```
## [1] "Female first author team size 2018 geometric mean: 4.15670586587764"
## [1] "Male first author team size 2018 geometric mean: 3.49457284827817"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 92000, p-value = 9e-04
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 3.9424813519408"
## [1] "Male last author team size 2018 geometric mean: 3.64128845443974"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 65000, p-value = 0.2
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.030  1           1.015
## LastAuthorFemale  1.015  1           1.007
## UniqueAuthors     1.084  4           1.010
## Year              1.089 16           1.003
```



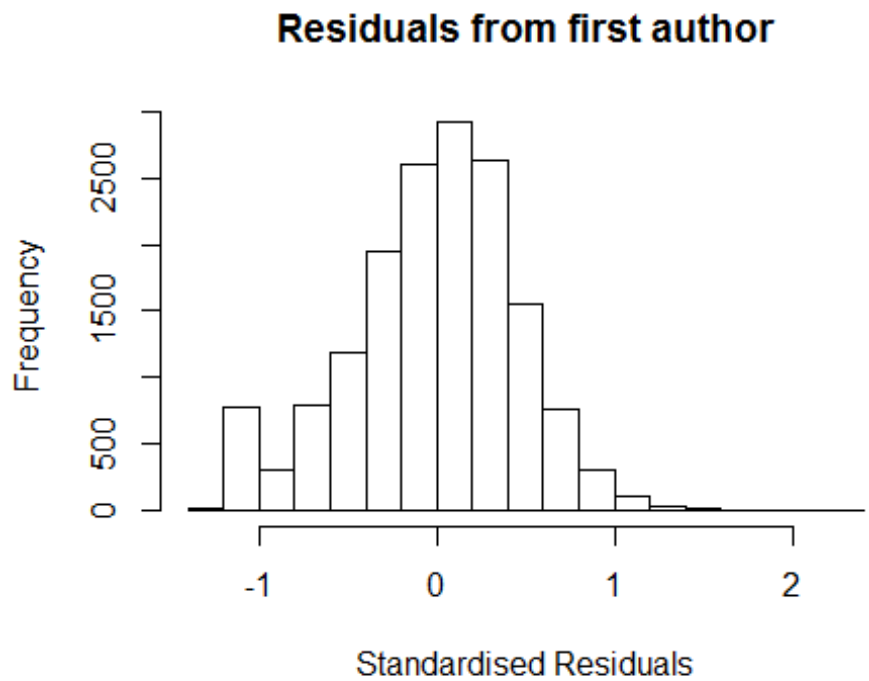

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
UniqueAuthors +
##     Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##     k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.3325 -0.2888  0.0159  0.2893  2.3554
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         0.78810    0.02128   37.04  < 2e-16 ***
## FirstAuthorFemale1  0.02309    0.00792    2.91   0.0036 **
## LastAuthorFemale1  -0.02102    0.00977   -2.15   0.0314 *
## UniqueAuthors2      0.39188    0.01689   23.21  < 2e-16 ***
## UniqueAuthors3      0.43967    0.01682   26.13  < 2e-16 ***
## UniqueAuthors4      0.47313    0.01731   27.33  < 2e-16 ***
## UniqueAuthors5      0.52127    0.01652   31.55  < 2e-16 ***
## Year1997           -0.05137    0.02323   -2.21   0.0270 *
## Year1998            0.00364    0.02377    0.15   0.8783
## Year1999           -0.02081    0.02370   -0.88   0.3801
```

```
## Year2000           -0.04176    0.02474   -1.69   0.0914 .
## Year2001           -0.20421    0.02565   -7.96  1.8e-15 ***
## Year2002           -0.10537    0.02162   -4.87  1.1e-06 ***
## Year2003           -0.13659    0.02286   -5.98  2.3e-09 ***
## Year2004           -0.13818    0.02224   -6.21  5.3e-10 ***
## Year2005           -0.17402    0.02240   -7.77  8.4e-15 ***
## Year2006           -0.17127    0.02202   -7.78  7.8e-15 ***
## Year2007           -0.16478    0.02129   -7.74  1.1e-14 ***
## Year2008           -0.17439    0.02115   -8.25  < 2e-16 ***
## Year2009           -0.18554    0.02075   -8.94  < 2e-16 ***
## Year2010           -0.15671    0.02125   -7.38  1.7e-13 ***
## Year2011           -0.20348    0.02135   -9.53  < 2e-16 ***
## Year2012           -0.19153    0.02141   -8.95  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.427
## Multiple R-squared:  0.112,  Adjusted R-squared:  0.11
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  2 observations c(1008,1790) are outliers with |weight| = 0 ( < 6.3e-06);
##  1360 weights are ~= 1. The remaining 14580 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0052  0.8630  0.9500  0.8990  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          6.27e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.018  1           1.009
## LastAuthorFemale  1.010  1           1.005
## Year              1.021 16           1.001
```



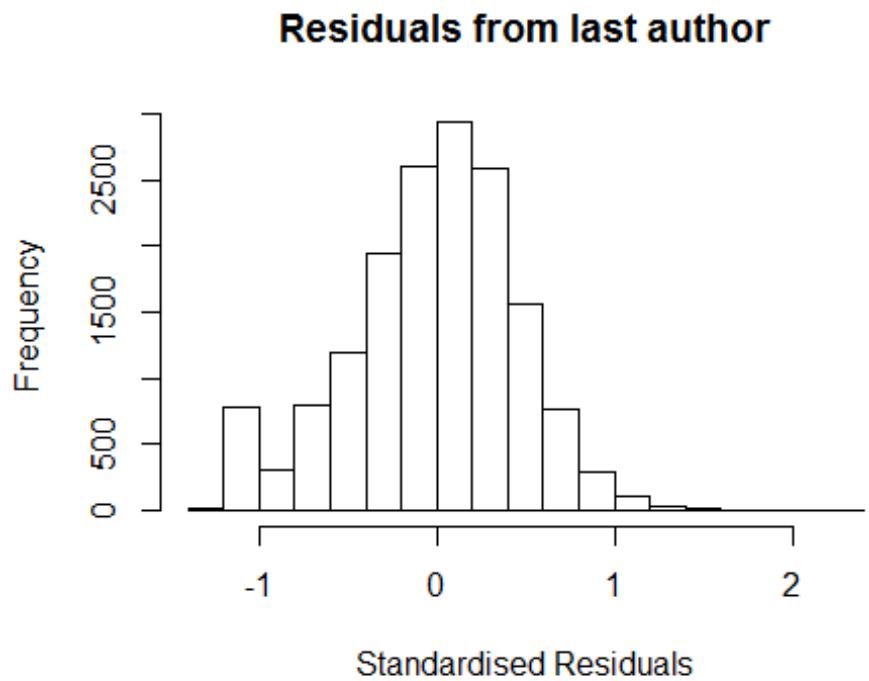

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.2284 -0.2946  0.0215  0.3021  2.3763
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)          1.15684    0.01774   65.22  < 2e-16 ***
## FirstAuthorFemale1   0.05156    0.00813    6.34  2.3e-10 ***
## LastAuthorFemale1   -0.01958    0.01009   -1.94  0.05238 .
## Year1997            -0.04918    0.02406   -2.04  0.04098 *
## Year1998             0.02004    0.02472    0.81  0.41764
## Year1999            -0.01455    0.02466   -0.59  0.55515
## Year2000            -0.02738    0.02625   -1.04  0.29680
## Year2001            -0.19383    0.02648   -7.32  2.6e-13 ***
## Year2002            -0.06298    0.02268   -2.78  0.00550 **
## Year2003            -0.08970    0.02414   -3.72  0.00020 ***
## Year2004            -0.08233    0.02343   -3.51  0.00044 ***
## Year2005            -0.12777    0.02398   -5.33  1.0e-07 ***
```

```
## Year2006           -0.12200    0.02349   -5.19  2.1e-07 ***
## Year2007           -0.10633    0.02232   -4.76  1.9e-06 ***
## Year2008           -0.10922    0.02208   -4.95  7.7e-07 ***
## Year2009           -0.13325    0.02213   -6.02  1.8e-09 ***
## Year2010           -0.09941    0.02226   -4.46  8.1e-06 ***
## Year2011           -0.13868    0.02263   -6.13  9.0e-10 ***
## Year2012           -0.12323    0.02265   -5.44  5.4e-08 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.439
## Multiple R-squared:  0.0144, Adjusted R-squared:  0.0132
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
##  observation 1008 is an outlier with |weight| = 0 ( < 6.3e-06);
##  1349 weights are ~= 1. The remaining 14592 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0016  0.8620  0.9490  0.8960  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          6.27e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.014  1           1.007
## Year              1.014 16           1.000
```




```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.2242 -0.2939  0.0221  0.3015  2.3785
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)          1.15474    0.01770   65.23  < 2e-16 ***
## FirstAuthorFemale1   0.04970    0.00813    6.11  9.9e-10 ***
## Year1997            -0.04928    0.02406   -2.05  0.04055 *
## Year1998             0.01976    0.02472    0.80  0.42419
## Year1999            -0.01467    0.02464   -0.60  0.55172
## Year2000            -0.02760    0.02624   -1.05  0.29290
## Year2001            -0.19443    0.02650   -7.34  2.3e-13 ***
## Year2002            -0.06317    0.02269   -2.78  0.00537 **
## Year2003            -0.09034    0.02415   -3.74  0.00018 ***
## Year2004            -0.08280    0.02345   -3.53  0.00041 ***
## Year2005            -0.12815    0.02401   -5.34  9.5e-08 ***
## Year2006            -0.12263    0.02348   -5.22  1.8e-07 ***
```

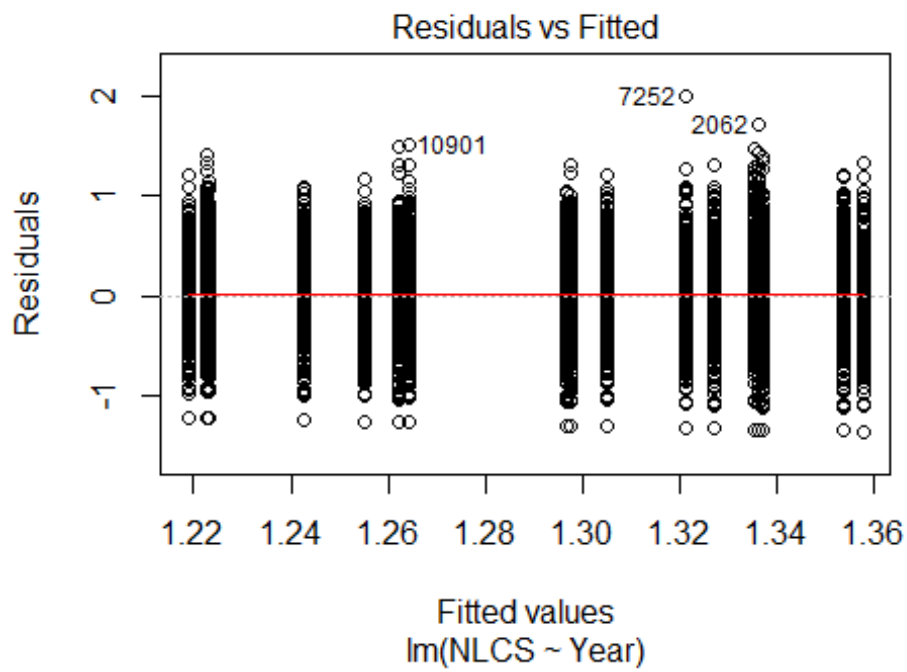
```
## Year2007           -0.10685    0.02233   -4.79  1.7e-06 ***
## Year2008           -0.10972    0.02209   -4.97  6.8e-07 ***
## Year2009           -0.13401    0.02213   -6.06  1.4e-09 ***
## Year2010           -0.10054    0.02225   -4.52  6.3e-06 ***
## Year2011           -0.13989    0.02261   -6.19  6.3e-10 ***
## Year2012           -0.12441    0.02264   -5.49  4.0e-08 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.439
## Multiple R-squared:  0.0141, Adjusted R-squared:  0.0131
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
##  observation 1008 is an outlier with |weight| = 0 ( < 6.3e-06);
##  1345 weights are ~= 1. The remaining 14596 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0014  0.8620  0.9490  0.8960  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          6.27e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##                   GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.006  1           1.003
## Year             1.006 16           1.000
```



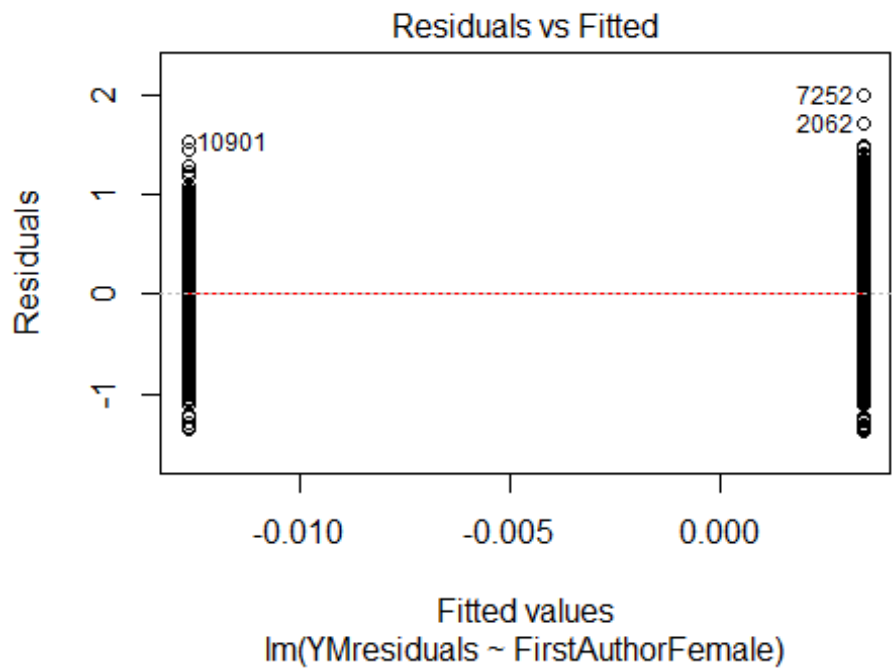

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.1881 -0.2981  0.0237  0.3011  2.3657
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)          1.1662     0.0177   65.94  < 2e-16 ***
## LastAuthorFemale1   -0.0123     0.0101   -1.23  0.22015
## Year1997            -0.0479     0.0241   -1.99  0.04708 *
## Year1998             0.0219     0.0247    0.88  0.37649
## Year1999            -0.0118     0.0247   -0.48  0.63305
## Year2000            -0.0279     0.0262   -1.06  0.28740
## Year2001            -0.1924     0.0265   -7.26  4.1e-13 ***
## Year2002            -0.0589     0.0227   -2.60  0.00947 **
## Year2003            -0.0841     0.0241   -3.49  0.00049 ***
## Year2004            -0.0769     0.0234   -3.28  0.00103 **
## Year2005            -0.1239     0.0240   -5.16  2.6e-07 ***
## Year2006            -0.1188     0.0235   -5.05  4.4e-07 ***
```

```
## Year2007           -0.1006     0.0223   -4.50  6.7e-06 ***
## Year2008           -0.1038     0.0221   -4.70  2.6e-06 ***
## Year2009           -0.1293     0.0222   -5.83  5.5e-09 ***
## Year2010           -0.0940     0.0223   -4.22  2.5e-05 ***
## Year2011           -0.1336     0.0226   -5.90  3.6e-09 ***
## Year2012           -0.1180     0.0227   -5.21  1.9e-07 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.439
## Multiple R-squared:  0.012,  Adjusted R-squared:  0.011
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
##  observation 1008 is an outlier with |weight| = 0 ( < 6.3e-06);
##  1342 weights are ~= 1. The remaining 14599 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0026  0.8630  0.9490  0.8960  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          6.27e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Sample size for the above analysis:  15942"
## [1] ""
## [1] ""
## [1] "####################################"
## [1] "Analysis of AJSC 1603"
## [1] "####################################"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  735  712  775  789  762  713  801  900  998 1000 1077 1152 1076 1072 1331
## 2011 2012
## 1243 1258
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  332  293  381  359  348  291  390  421  463  487  538  567  552  549  694
```

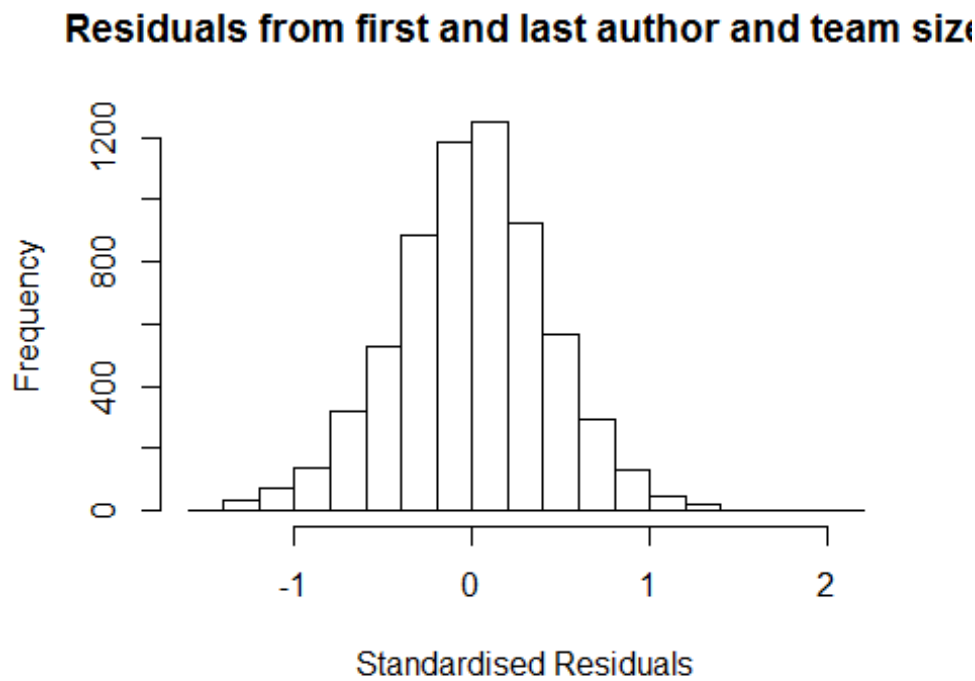
```
## 2011 2012
##  630  660
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  266  248  318  313  299  237  313  328  381  396  438  441  446  430  544
## 2011 2012
##  483  509
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
##  Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 95, df = 16, p-value = 3e-13
```




```
##
##  Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 17, df = 1, p-value = 5e-05
```

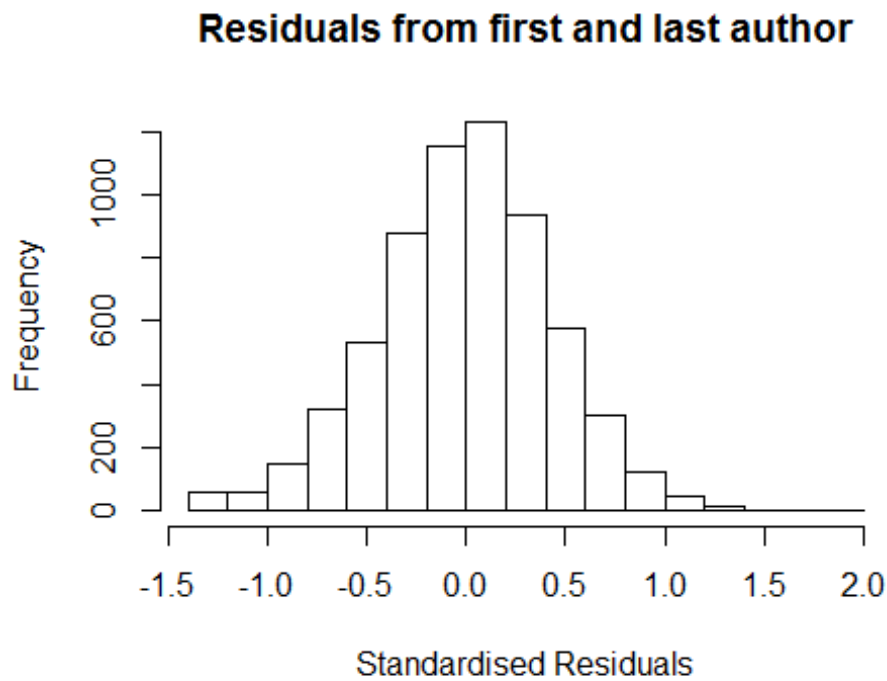



```
## [1] "Female first author team size 2018 geometric mean: 4.37916718363068"
## [1] "Male first author team size 2018 geometric mean: 3.9376784229724"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 8700, p-value = 0.2
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 4.26600189217586"
## [1] "Male last author team size 2018 geometric mean: 4.01583353896821"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 6200, p-value = 0.5
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.026  1           1.013
## LastAuthorFemale  1.020  1           1.010
## UniqueAuthors     1.137  4           1.016
## Year              1.161 16           1.005
```



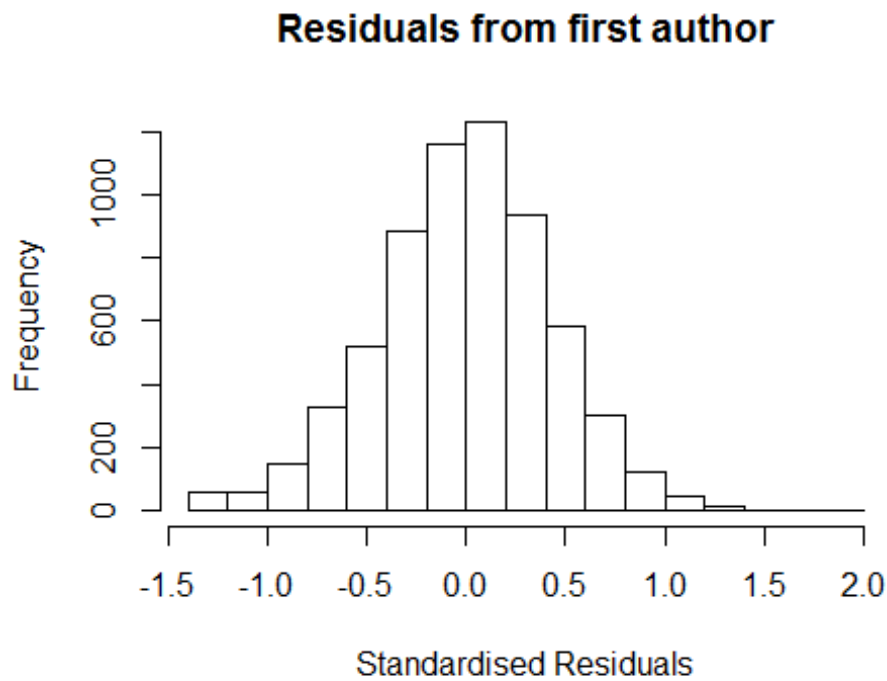

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
UniqueAuthors +
##     Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##     k.max = 1000))
##  \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.41613 -0.27482  0.00573  0.27540  2.19878
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.12e+00   4.89e-02   22.91  < 2e-16 ***
## FirstAuthorFemale1 -2.43e-02   1.26e-02   -1.93  0.05395 .
## LastAuthorFemale1   1.53e-02   1.52e-02    1.01  0.31473
## UniqueAuthors2      2.02e-01   3.87e-02    5.23  1.8e-07 ***
## UniqueAuthors3      2.25e-01   3.89e-02    5.78  7.9e-09 ***
## UniqueAuthors4      2.54e-01   3.93e-02    6.46  1.1e-10 ***
## UniqueAuthors5      2.72e-01   3.90e-02    6.97  3.5e-12 ***
## Year1997            2.15e-03   4.61e-02    0.05  0.96276
## Year1998            2.41e-02   4.30e-02    0.56  0.57556
## Year1999            3.05e-02   4.36e-02    0.70  0.48353
```

```
## Year2000           -3.11e-02   4.34e-02   -0.72  0.47383
## Year2001            4.50e-02   4.40e-02    1.02  0.30679
## Year2002           -1.11e-02   4.10e-02   -0.27  0.78604
## Year2003            7.44e-05   4.05e-02    0.00  0.99854
## Year2004           -6.72e-03   3.99e-02   -0.17  0.86626
## Year2005           -3.51e-02   3.85e-02   -0.91  0.36142
## Year2006           -5.89e-02   3.76e-02   -1.57  0.11714
## Year2007           -5.88e-02   3.82e-02   -1.54  0.12452
## Year2008           -9.98e-02   3.80e-02   -2.63  0.00860 **
## Year2009           -1.17e-01   3.83e-02   -3.05  0.00234 **
## Year2010           -9.03e-02   3.70e-02   -2.44  0.01463 *
## Year2011           -1.35e-01   3.69e-02   -3.66  0.00026 ***
## Year2012           -1.34e-01   3.83e-02   -3.49  0.00048 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.411
## Multiple R-squared:  0.031,  Adjusted R-squared:  0.0277
## Convergence in 15 IRWLS iterations
##
## Robustness weights:
##  observation 2477 is an outlier with |weight| = 0 ( < 1.6e-05);
##  546 weights are ~= 1. The remaining 5843 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0816  0.8660  0.9510  0.8990  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.56e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.020  1           1.010
## LastAuthorFemale  1.017  1           1.009
## Year              1.030 16           1.001
```



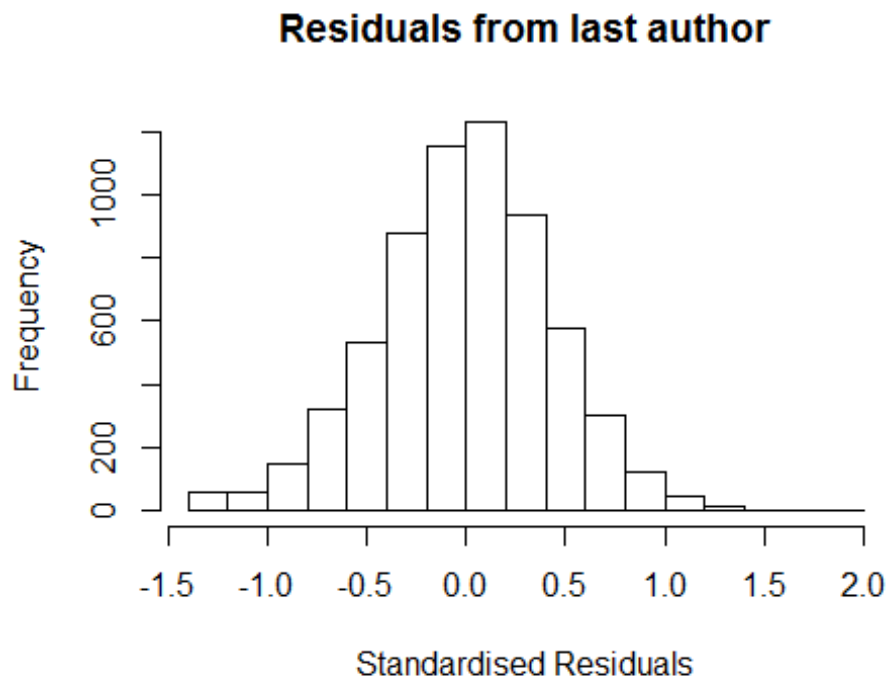

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.36606 -0.27844  0.00766  0.27177  1.97233
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.32544    0.03275   40.47   <2e-16 ***
## FirstAuthorFemale1 -0.01529    0.01272   -1.20   0.2294
## LastAuthorFemale1   0.01693    0.01535    1.10   0.2700
## Year1997           -0.00602    0.04588   -0.13   0.8956
## Year1998            0.02517    0.04313    0.58   0.5595
## Year1999            0.03896    0.04374    0.89   0.3732
## Year2000           -0.01518    0.04349   -0.35   0.7270
## Year2001            0.05591    0.04413    1.27   0.2053
## Year2002            0.00303    0.04120    0.07   0.9414
## Year2003            0.02207    0.04049    0.55   0.5857
## Year2004            0.01423    0.03992    0.36   0.7214
## Year2005           -0.01667    0.03866   -0.43   0.6664
```

```
## Year2006           -0.03599    0.03772   -0.95   0.3401
## Year2007           -0.03284    0.03821   -0.86   0.3901
## Year2008           -0.07742    0.03796   -2.04   0.0414 *
## Year2009           -0.09544    0.03818   -2.50   0.0125 *
## Year2010           -0.06398    0.03695   -1.73   0.0834 .
## Year2011           -0.10404    0.03684   -2.82   0.0048 **
## Year2012           -0.10138    0.03808   -2.66   0.0078 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.415
## Multiple R-squared:  0.0135, Adjusted R-squared:  0.0107
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  observation 2477 is an outlier with |weight| = 0 ( < 1.6e-05);
##  540 weights are ~= 1. The remaining 5849 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0577  0.8680  0.9510  0.8990  0.9860  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.56e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.016  1           1.008
## Year              1.016 16           1.001
```




```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.36902 -0.27827  0.00582  0.27122  1.97077
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.32765    0.03266   40.65   <2e-16 ***
## FirstAuthorFemale1 -0.01394    0.01271   -1.10   0.2729
## Year1997           -0.00735    0.04586   -0.16   0.8727
## Year1998            0.02417    0.04316    0.56   0.5755
## Year1999            0.03857    0.04377    0.88   0.3782
## Year2000           -0.01502    0.04352   -0.35   0.7300
## Year2001            0.05532    0.04414    1.25   0.2101
## Year2002            0.00250    0.04120    0.06   0.9515
## Year2003            0.02131    0.04051    0.53   0.5990
## Year2004            0.01359    0.03991    0.34   0.7335
## Year2005           -0.01640    0.03868   -0.42   0.6717
## Year2006           -0.03611    0.03775   -0.96   0.3389
```

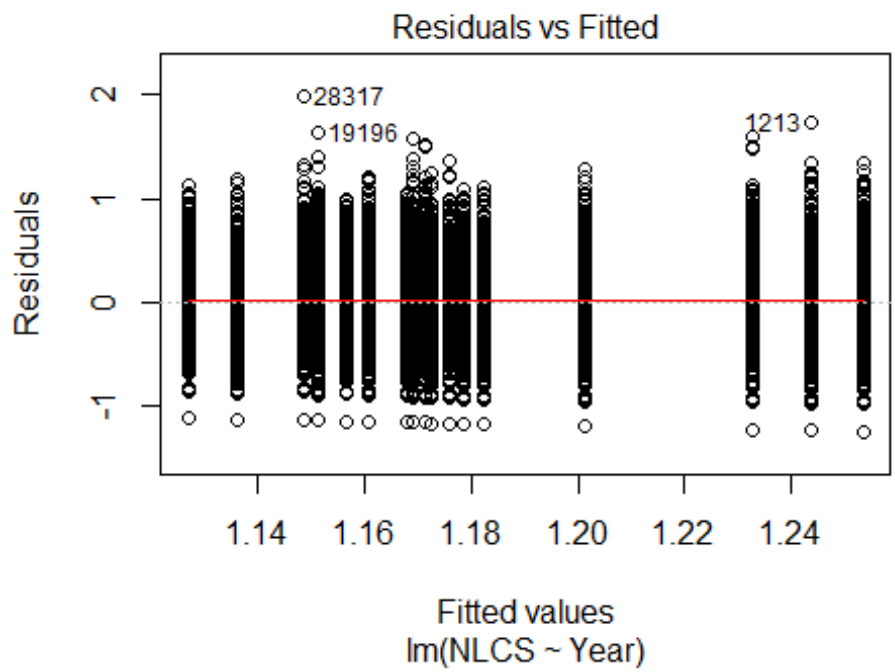
```
## Year2007           -0.03322    0.03823   -0.87   0.3849
## Year2008           -0.07737    0.03799   -2.04   0.0417 *
## Year2009           -0.09561    0.03819   -2.50   0.0123 *
## Year2010           -0.06354    0.03697   -1.72   0.0857 .
## Year2011           -0.10401    0.03686   -2.82   0.0048 **
## Year2012           -0.10091    0.03811   -2.65   0.0081 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.415
## Multiple R-squared:  0.0133, Adjusted R-squared:  0.0106
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  observation 2477 is an outlier with |weight| = 0 ( < 1.6e-05);
##  535 weights are ~= 1. The remaining 5854 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0585  0.8680  0.9510  0.8990  0.9860  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.56e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##                   GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.014  1           1.007
## Year             1.014 16           1.000
```



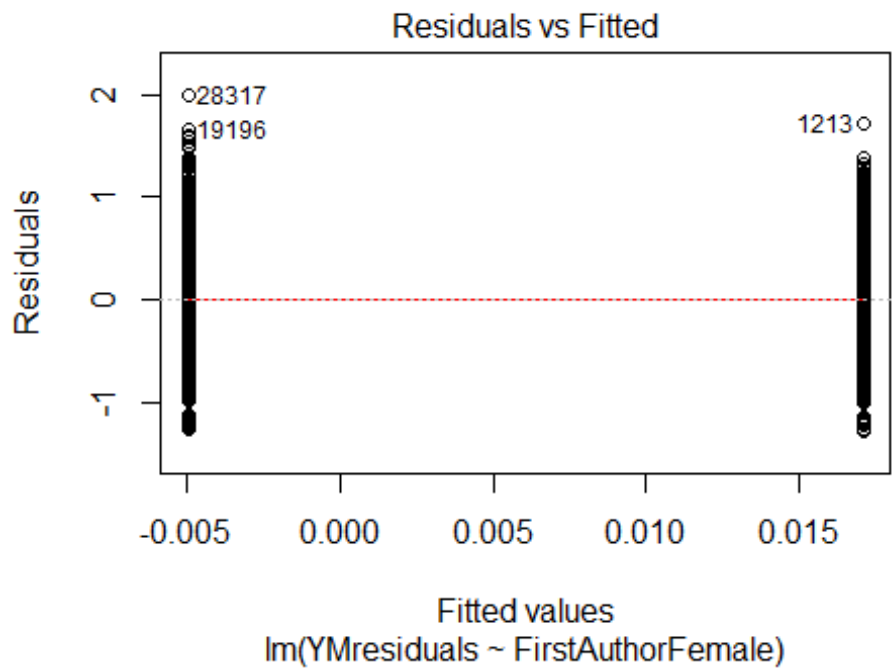

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.37830 -0.27956  0.00742  0.27347  1.97555
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.32319    0.03269   40.48   <2e-16 ***
## LastAuthorFemale1   0.01516    0.01533    0.99   0.3228
## Year1997           -0.00641    0.04582   -0.14   0.8887
## Year1998            0.02408    0.04307    0.56   0.5761
## Year1999            0.03843    0.04374    0.88   0.3797
## Year2000           -0.01570    0.04346   -0.36   0.7180
## Year2001            0.05511    0.04406    1.25   0.2110
## Year2002            0.00221    0.04121    0.05   0.9572
## Year2003            0.02071    0.04042    0.51   0.6083
## Year2004            0.01326    0.03988    0.33   0.7395
## Year2005           -0.01737    0.03863   -0.45   0.6530
## Year2006           -0.03659    0.03770   -0.97   0.3317
```

```
## Year2007          -0.03397    0.03815   -0.89   0.3732
## Year2008          -0.07849    0.03790   -2.07   0.0384 *
## Year2009          -0.09563    0.03816   -2.51   0.0122 *
## Year2010          -0.06524    0.03689   -1.77   0.0771 .
## Year2011          -0.10519    0.03679   -2.86   0.0043 **
## Year2012          -0.10330    0.03797   -2.72   0.0065 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.415
## Multiple R-squared:  0.0132, Adjusted R-squared:  0.0106
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  observation 2477 is an outlier with |weight| = 0 ( < 1.6e-05);
##  532 weights are ~= 1. The remaining 5857 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0566  0.8690  0.9510  0.8990  0.9860  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.56e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Sample size for the above analysis:  6390"
## [1] ""
## [1] ""
## [1] "####################################"
## [1] "Analysis of AJSC 1604"
## [1] "####################################"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 1945 1771 1780 1628 1631 1424 1394 1445 1292 1326 1425 1275 1414 1433 1453
## 2011 2012
## 1386 1565
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 1162  981 1033  975  874  550  866  943  793  794  879  809  860  935  971
```

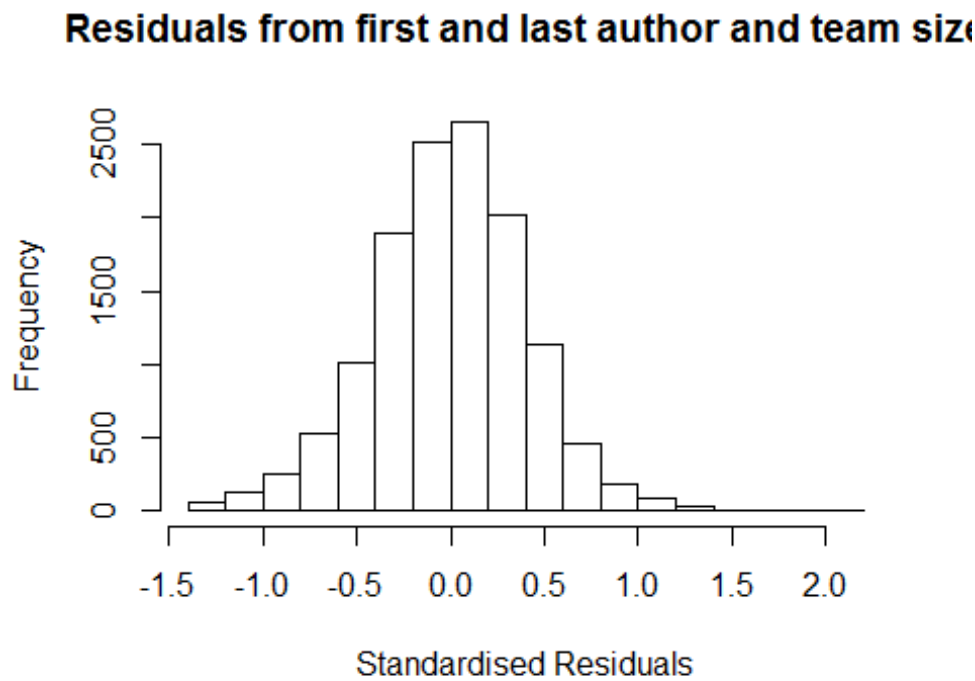
```
## 2011 2012
##  912 1009
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  965  825  862  842  711  464  735  783  667  652  748  700  743  797  828
## 2011 2012
##  760  856
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
##  Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 83, df = 16, p-value = 5e-11
```




```
##
##  Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 12, df = 1, p-value = 6e-04
```

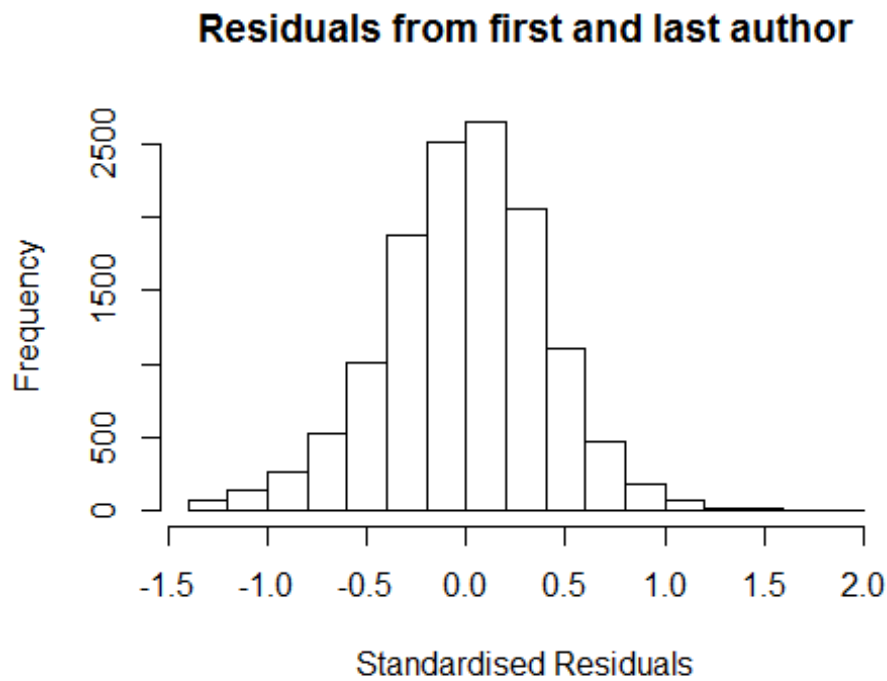



```
## [1] "Female first author team size 2018 geometric mean: 4.6196485790188"
## [1] "Male first author team size 2018 geometric mean: 4.0979418435252"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 68000, p-value = 0.004
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 4.64330947945002"
## [1] "Male last author team size 2018 geometric mean: 4.13408268329736"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 56000, p-value = 0.01
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##                     GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.020  1           1.010
## LastAuthorFemale  1.016  1           1.008
## UniqueAuthors     1.082  4           1.010
## Year              1.105 16           1.003
```



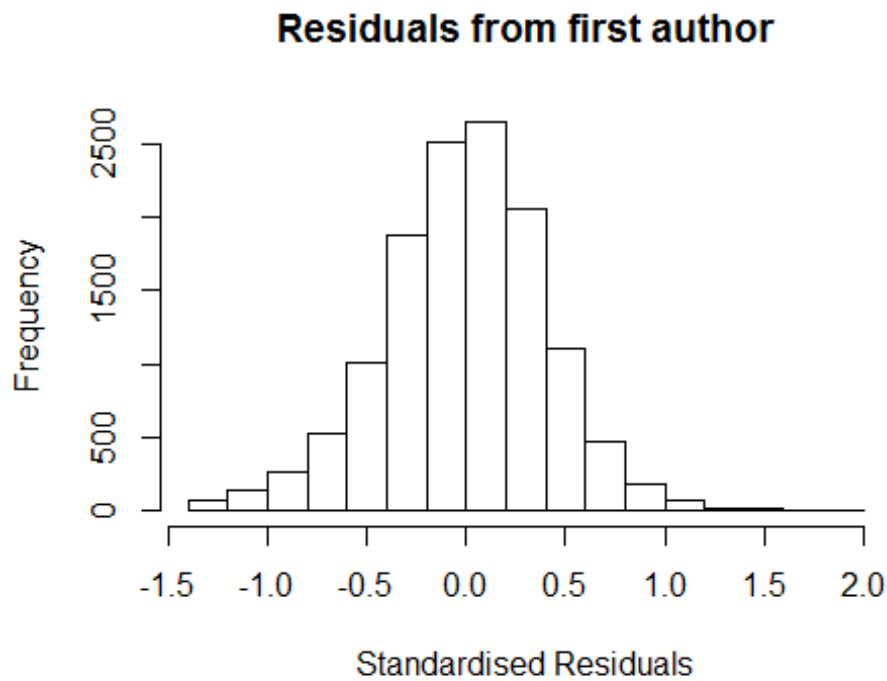

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
UniqueAuthors +
##     Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##     k.max = 1000))
##  \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.33795 -0.25371  0.00524  0.25752  2.17132
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.08870    0.02926   37.21  < 2e-16 ***
## FirstAuthorFemale1  0.02549    0.00821    3.11  0.00190 **
## LastAuthorFemale1  -0.00176    0.01067   -0.17  0.86866
## UniqueAuthors2      0.16145    0.02667    6.05  1.4e-09 ***
## UniqueAuthors3      0.16304    0.02641    6.17  6.9e-10 ***
## UniqueAuthors4      0.19174    0.02666    7.19  6.8e-13 ***
## UniqueAuthors5      0.22376    0.02640    8.48  < 2e-16 ***
## Year1997           -0.02331    0.02103   -1.11  0.26783
## Year1998           -0.00312    0.01983   -0.16  0.87499
## Year1999           -0.07007    0.02098   -3.34  0.00084 ***
```

```
## Year2000           -0.05011    0.02020   -2.48  0.01312 *
## Year2001           -0.07468    0.02350   -3.18  0.00149 **
## Year2002           -0.08623    0.02037   -4.23  2.3e-05 ***
## Year2003           -0.08035    0.01988   -4.04  5.3e-05 ***
## Year2004           -0.08957    0.02024   -4.43  9.7e-06 ***
## Year2005           -0.09961    0.01995   -4.99  6.0e-07 ***
## Year2006           -0.12672    0.01958   -6.47  1.0e-10 ***
## Year2007           -0.11740    0.01996   -5.88  4.1e-09 ***
## Year2008           -0.10305    0.01993   -5.17  2.4e-07 ***
## Year2009           -0.10909    0.01987   -5.49  4.1e-08 ***
## Year2010           -0.10350    0.01957   -5.29  1.3e-07 ***
## Year2011           -0.13782    0.02025   -6.81  1.0e-11 ***
## Year2012           -0.12402    0.01969   -6.30  3.1e-10 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.381
## Multiple R-squared:  0.0243, Adjusted R-squared:  0.0226
## Convergence in 15 IRWLS iterations
##
## Robustness weights:
##  2 observations c(8310,12613) are outliers with |weight| = 0 ( < 7.7e-06);
##  1074 weights are ~= 1. The remaining 11862 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0205  0.8680  0.9500  0.8980  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          7.73e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.020  1           1.010
## LastAuthorFemale  1.015  1           1.008
## Year              1.027 16           1.001
```



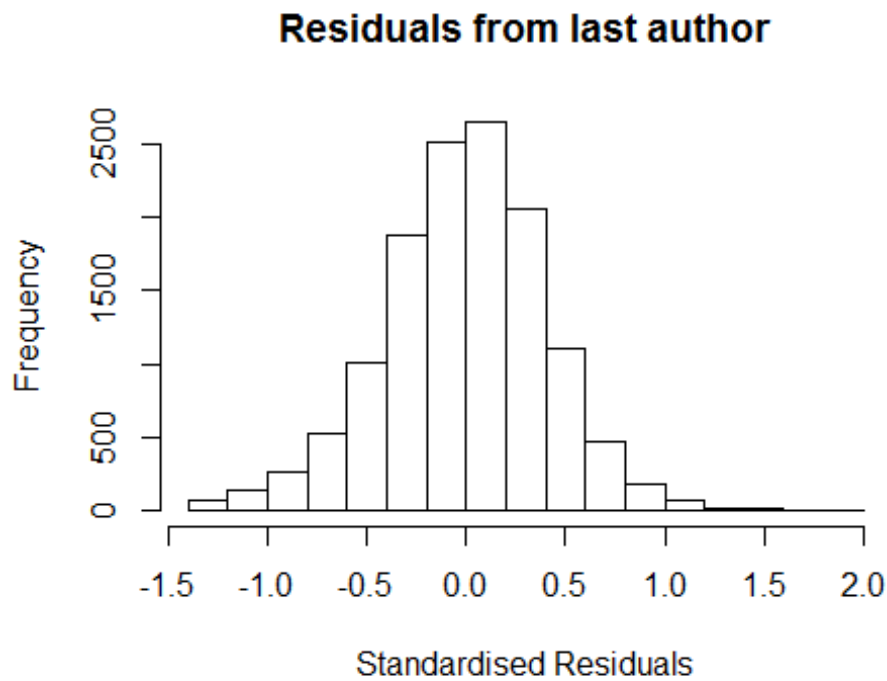

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.29545 -0.25766  0.00675  0.25858  1.98718
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)          1.26436    0.01424   88.80  < 2e-16 ***
## FirstAuthorFemale1  0.03108    0.00822    3.78  0.00016 ***
## LastAuthorFemale1  -0.00109    0.01065   -0.10  0.91845
## Year1997           -0.02698    0.02113   -1.28  0.20177
## Year1998           -0.00639    0.01989   -0.32  0.74783
## Year1999           -0.08887    0.02084   -4.26  2.0e-05 ***
## Year2000           -0.05603    0.02024   -2.77  0.00564 **
## Year2001           -0.08102    0.02380   -3.40  0.00067 ***
## Year2002           -0.08362    0.02041   -4.10  4.2e-05 ***
## Year2003           -0.07450    0.02001   -3.72  0.00020 ***
## Year2004           -0.08567    0.02030   -4.22  2.5e-05 ***
## Year2005           -0.09157    0.01998   -4.58  4.6e-06 ***
```

```
## Year2006           -0.11956    0.01971   -6.06  1.4e-09 ***
## Year2007           -0.10887    0.02004   -5.43  5.6e-08 ***
## Year2008           -0.09771    0.01998   -4.89  1.0e-06 ***
## Year2009           -0.10313    0.02002   -5.15  2.6e-07 ***
## Year2010           -0.09616    0.01960   -4.91  9.4e-07 ***
## Year2011           -0.12933    0.02037   -6.35  2.2e-10 ***
## Year2012           -0.11554    0.01984   -5.82  5.9e-09 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.382
## Multiple R-squared:  0.0103, Adjusted R-squared:  0.00897
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  observation 12613 is an outlier with |weight| = 0 ( < 7.7e-06);
##  1081 weights are ~= 1. The remaining 11856 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.025   0.868   0.951   0.897   0.985   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          7.73e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.016  1           1.008
## Year              1.016 16           1.001
```




```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.29528 -0.25757  0.00669  0.25875  1.98733
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)          1.26426    0.01421   88.96  < 2e-16 ***
## FirstAuthorFemale1   0.03101    0.00821    3.78  0.00016 ***
## Year1997            -0.02695    0.02113   -1.28  0.20209
## Year1998            -0.00640    0.01989   -0.32  0.74773
## Year1999            -0.08887    0.02084   -4.26  2.0e-05 ***
## Year2000            -0.05603    0.02024   -2.77  0.00565 **
## Year2001            -0.08102    0.02380   -3.40  0.00067 ***
## Year2002            -0.08359    0.02041   -4.10  4.2e-05 ***
## Year2003            -0.07452    0.02001   -3.72  0.00020 ***
## Year2004            -0.08568    0.02030   -4.22  2.4e-05 ***
## Year2005            -0.09160    0.01998   -4.59  4.6e-06 ***
## Year2006            -0.11954    0.01971   -6.06  1.4e-09 ***
```

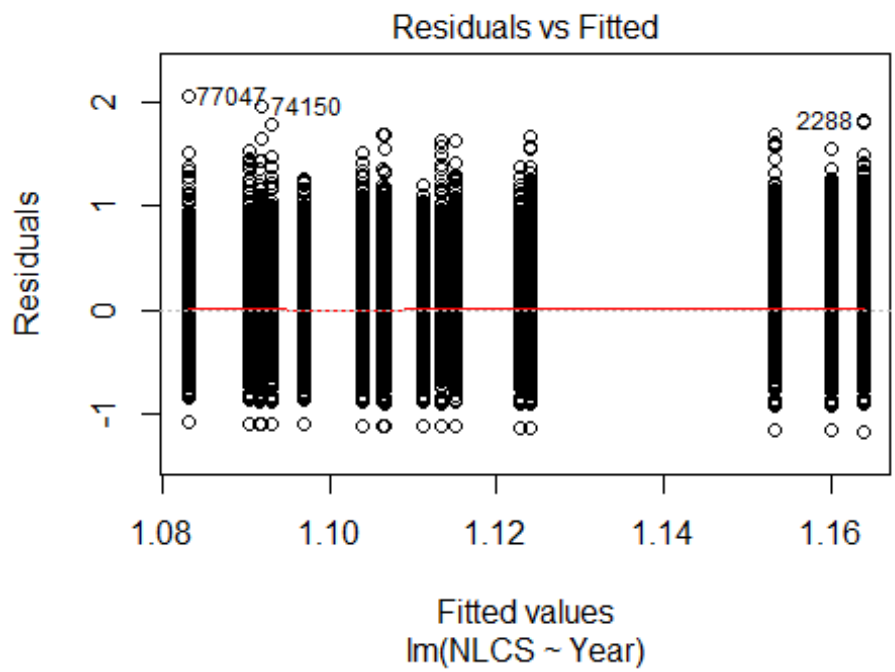
```
## Year2007           -0.10890    0.02003   -5.44  5.5e-08 ***
## Year2008           -0.09778    0.01996   -4.90  9.7e-07 ***
## Year2009           -0.10314    0.02002   -5.15  2.6e-07 ***
## Year2010           -0.09621    0.01958   -4.91  9.1e-07 ***
## Year2011           -0.12938    0.02036   -6.35  2.2e-10 ***
## Year2012           -0.11559    0.01984   -5.83  5.8e-09 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.382
## Multiple R-squared:  0.0103, Adjusted R-squared:  0.00905
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  observation 12613 is an outlier with |weight| = 0 ( < 7.7e-06);
##  1075 weights are ~= 1. The remaining 11862 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0249  0.8680  0.9510  0.8970  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          7.73e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##                   GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.012  1           1.006
## Year             1.012 16           1.000
```



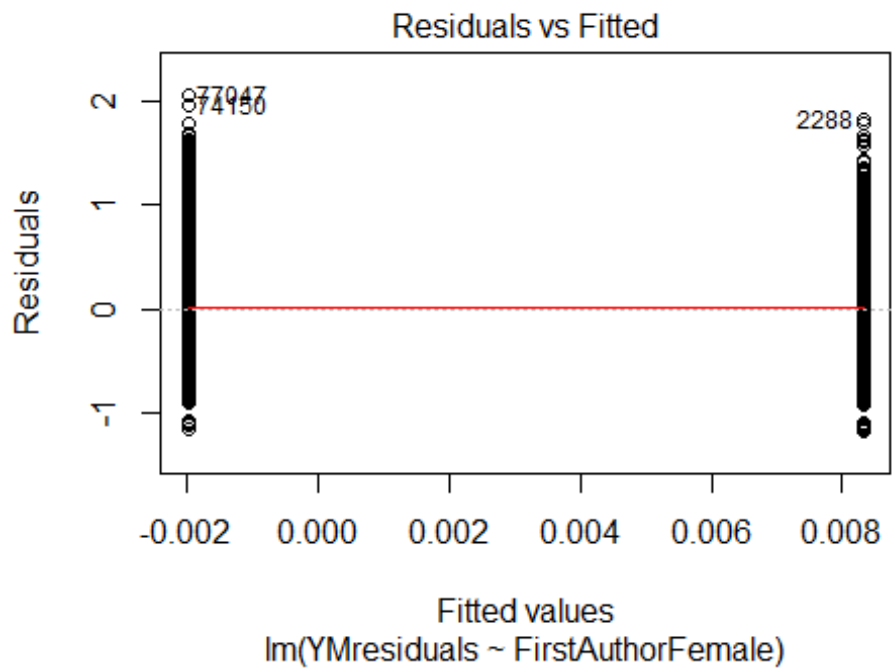

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.27282 -0.25806  0.00403  0.25831  1.97921
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.27073    0.01409   90.17  < 2e-16 ***
## LastAuthorFemale1  0.00209    0.01066    0.20  0.84471
## Year1997           -0.02716    0.02112   -1.29  0.19853
## Year1998           -0.00749    0.01986   -0.38  0.70621
## Year1999           -0.08996    0.02081   -4.32  1.6e-05 ***
## Year2000           -0.05718    0.02023   -2.83  0.00471 **
## Year2001           -0.08075    0.02381   -3.39  0.00070 ***
## Year2002           -0.08393    0.02038   -4.12  3.8e-05 ***
## Year2003           -0.07368    0.01999   -3.69  0.00023 ***
## Year2004           -0.08477    0.02031   -4.17  3.0e-05 ***
## Year2005           -0.09072    0.01997   -4.54  5.6e-06 ***
## Year2006           -0.11851    0.01972   -6.01  1.9e-09 ***
```

```
## Year2007          -0.10798    0.02003   -5.39  7.2e-08 ***
## Year2008          -0.09708    0.01998   -4.86  1.2e-06 ***
## Year2009          -0.10245    0.02002   -5.12  3.1e-07 ***
## Year2010          -0.09334    0.01955   -4.77  1.8e-06 ***
## Year2011          -0.12797    0.02037   -6.28  3.5e-10 ***
## Year2012          -0.11394    0.01983   -5.75  9.3e-09 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.382
## Multiple R-squared:  0.00925,    Adjusted R-squared:  0.00795
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  observation 12613 is an outlier with |weight| = 0 ( < 7.7e-06);
##  1103 weights are ~= 1. The remaining 11834 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0278  0.8680  0.9500  0.8970  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          7.73e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Sample size for the above analysis:  12938"
## [1] ""
## [1] ""
## [1] "####################################"
## [1] "Analysis of AJSC 1605"
## [1] "####################################"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 4522 4214 4153 4273 4089 4019 4189 4103 4145 4145 4339 4083 4270 4372 4247
## 2011 2012
## 4159 4136
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 2900 2658 2567 2658 2155 1856 2618 2601 2538 2629 2644 2572 2669 2761 2644
```

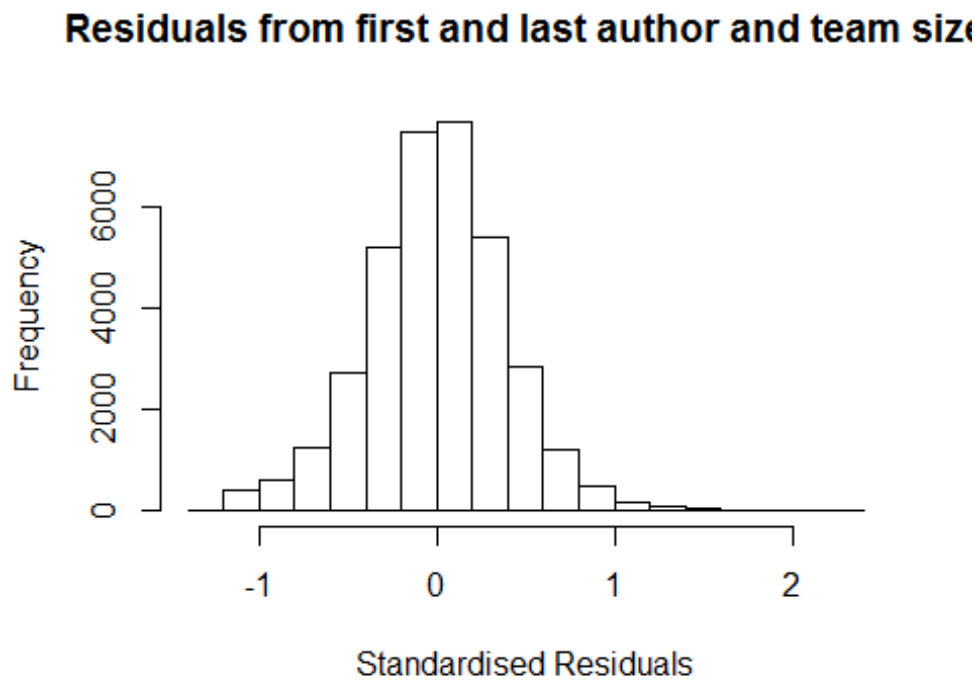
```
## 2011 2012
## 2610 2542
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 2335 2145 2099 2132 1736 1524 2153 2120 2063 2095 2126 2078 2174 2214 2193
## 2011 2012
## 2097 2118
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
##  Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 350, df = 16, p-value <2e-16
```




```
##
##  Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 0.21, df = 1, p-value = 0.7
```

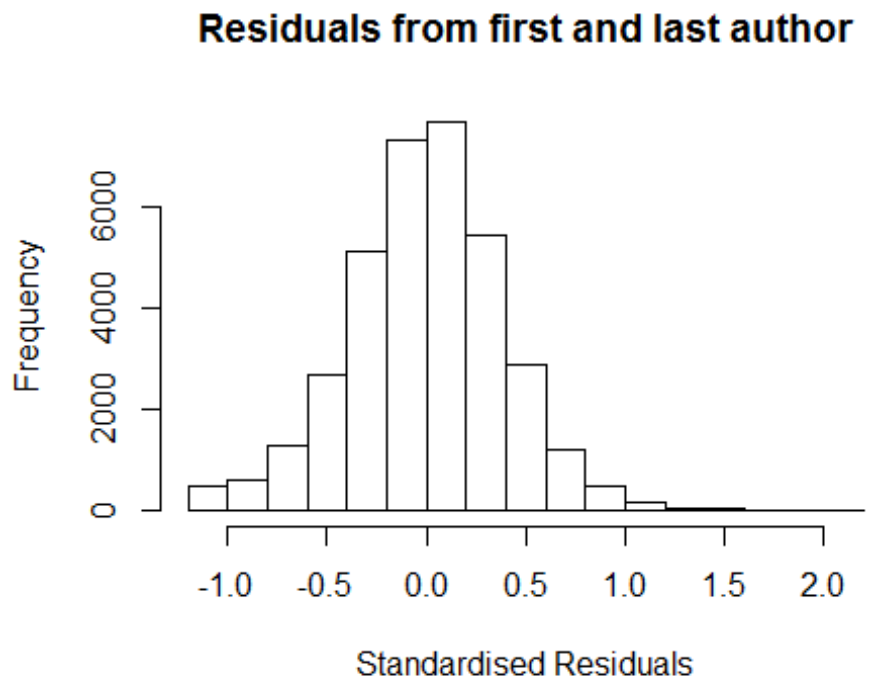



```
## [1] "Female first author team size 2018 geometric mean: 4.40539552204048"
## [1] "Male first author team size 2018 geometric mean: 4.22489375926577"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 190000, p-value = 0.3
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 4.13041890668187"
## [1] "Male last author team size 2018 geometric mean: 4.30558613203412"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 120000, p-value = 0.3
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.013  1           1.006
## LastAuthorFemale  1.007  1           1.003
## UniqueAuthors     1.053  4           1.007
## Year              1.053 16           1.002
```



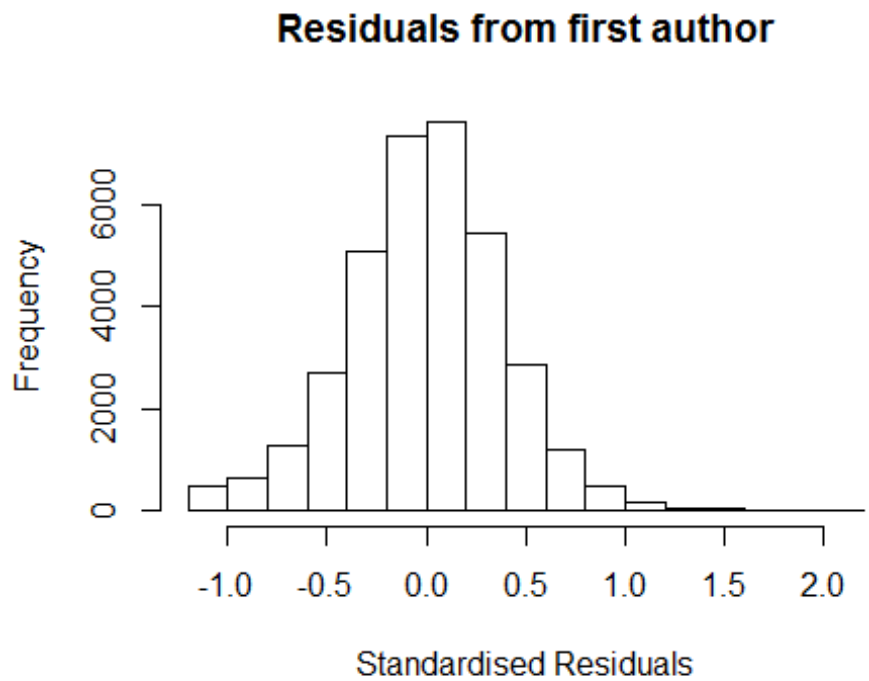

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
UniqueAuthors +
##     Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##     k.max = 1000))
##  \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.21712 -0.24043  0.00227  0.23973  2.21241
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.02109    0.01498   68.15  < 2e-16 ***
## FirstAuthorFemale1  0.00748    0.00508    1.47  0.14138
## LastAuthorFemale1  -0.01944    0.00601   -3.24  0.00121 **
## UniqueAuthors2      0.15884    0.01306   12.16  < 2e-16 ***
## UniqueAuthors3      0.15868    0.01310   12.11  < 2e-16 ***
## UniqueAuthors4      0.17643    0.01348   13.09  < 2e-16 ***
## UniqueAuthors5      0.18855    0.01293   14.58  < 2e-16 ***
## Year1997           -0.01361    0.01317   -1.03  0.30134
## Year1998           -0.01161    0.01262   -0.92  0.35792
## Year1999           -0.04430    0.01260   -3.52  0.00044 ***
```

```
## Year2000           -0.04343    0.01298   -3.35  0.00082 ***
## Year2001           -0.06319    0.01301   -4.86  1.2e-06 ***
## Year2002           -0.07089    0.01183   -5.99  2.1e-09 ***
## Year2003           -0.06928    0.01190   -5.82  6.0e-09 ***
## Year2004           -0.08092    0.01196   -6.76  1.4e-11 ***
## Year2005           -0.09526    0.01188   -8.02  1.1e-15 ***
## Year2006           -0.09226    0.01184   -7.79  6.7e-15 ***
## Year2007           -0.07856    0.01195   -6.57  5.0e-11 ***
## Year2008           -0.06338    0.01199   -5.29  1.3e-07 ***
## Year2009           -0.08720    0.01188   -7.34  2.1e-13 ***
## Year2010           -0.09737    0.01224   -7.96  1.8e-15 ***
## Year2011           -0.08866    0.01229   -7.22  5.5e-13 ***
## Year2012           -0.09750    0.01262   -7.72  1.2e-14 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.36
## Multiple R-squared:  0.0158, Adjusted R-squared:  0.0152
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  6 observations c(6270,8151,19495,22675,33160,34469)
##   are outliers with |weight| = 0 ( < 2.8e-06);
##  2997 weights are ~= 1. The remaining 32399 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.000   0.865   0.951   0.895   0.986   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          2.82e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.009  1           1.004
## LastAuthorFemale  1.002  1           1.001
## Year              1.010 16           1.000
```



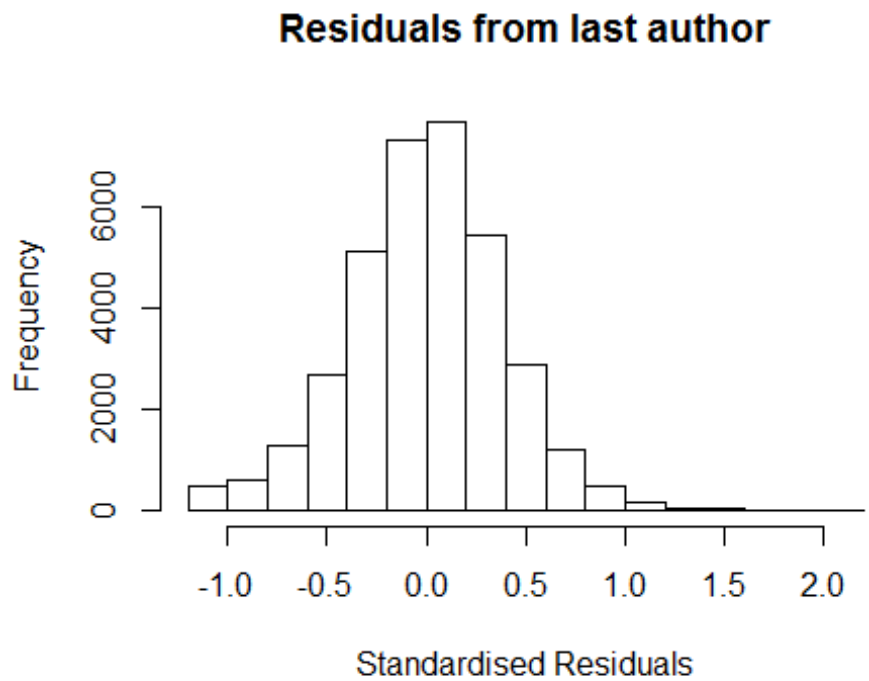

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.18899 -0.24104  0.00557  0.24286  2.04506
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)          1.17629    0.00907  129.67  < 2e-16 ***
## FirstAuthorFemale1  0.01270    0.00509    2.50  0.01257 *
## LastAuthorFemale1  -0.01553    0.00602   -2.58  0.00988 **
## Year1997           -0.01374    0.01324   -1.04  0.29952
## Year1998           -0.01009    0.01265   -0.80  0.42503
## Year1999           -0.04374    0.01265   -3.46  0.00054 ***
## Year2000           -0.04193    0.01302   -3.22  0.00128 **
## Year2001           -0.05797    0.01306   -4.44  9.0e-06 ***
## Year2002           -0.06286    0.01185   -5.31  1.1e-07 ***
## Year2003           -0.05989    0.01192   -5.02  5.1e-07 ***
## Year2004           -0.07083    0.01201   -5.90  3.7e-09 ***
## Year2005           -0.08470    0.01191   -7.11  1.2e-12 ***
```

```
## Year2006           -0.08140    0.01185   -6.87  6.5e-12 ***
## Year2007           -0.06687    0.01198   -5.58  2.4e-08 ***
## Year2008           -0.05253    0.01202   -4.37  1.2e-05 ***
## Year2009           -0.07616    0.01189   -6.41  1.5e-10 ***
## Year2010           -0.08553    0.01224   -6.99  2.9e-12 ***
## Year2011           -0.07653    0.01230   -6.22  4.9e-10 ***
## Year2012           -0.08535    0.01265   -6.75  1.5e-11 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.361
## Multiple R-squared:  0.00542,    Adjusted R-squared:  0.00492
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  3 observations c(19495,33160,34469)
##   are outliers with |weight| = 0 ( < 2.8e-06);
##  3007 weights are ~= 1. The remaining 32392 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.000   0.865   0.951   0.895   0.985   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          2.82e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.008  1           1.004
## Year              1.008 16           1.000
```




```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.18637 -0.24146  0.00602  0.24263  2.04703
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.17426    0.00904  129.91  < 2e-16 ***
## FirstAuthorFemale1  0.01211    0.00509    2.38  0.01727 *
## Year1997           -0.01328    0.01325   -1.00  0.31617
## Year1998           -0.00989    0.01266   -0.78  0.43462
## Year1999           -0.04347    0.01265   -3.44  0.00059 ***
## Year2000           -0.04180    0.01303   -3.21  0.00134 **
## Year2001           -0.05787    0.01307   -4.43  9.5e-06 ***
## Year2002           -0.06262    0.01185   -5.29  1.3e-07 ***
## Year2003           -0.05982    0.01193   -5.02  5.3e-07 ***
## Year2004           -0.07078    0.01201   -5.89  3.8e-09 ***
## Year2005           -0.08450    0.01191   -7.09  1.3e-12 ***
## Year2006           -0.08128    0.01185   -6.86  7.1e-12 ***
```

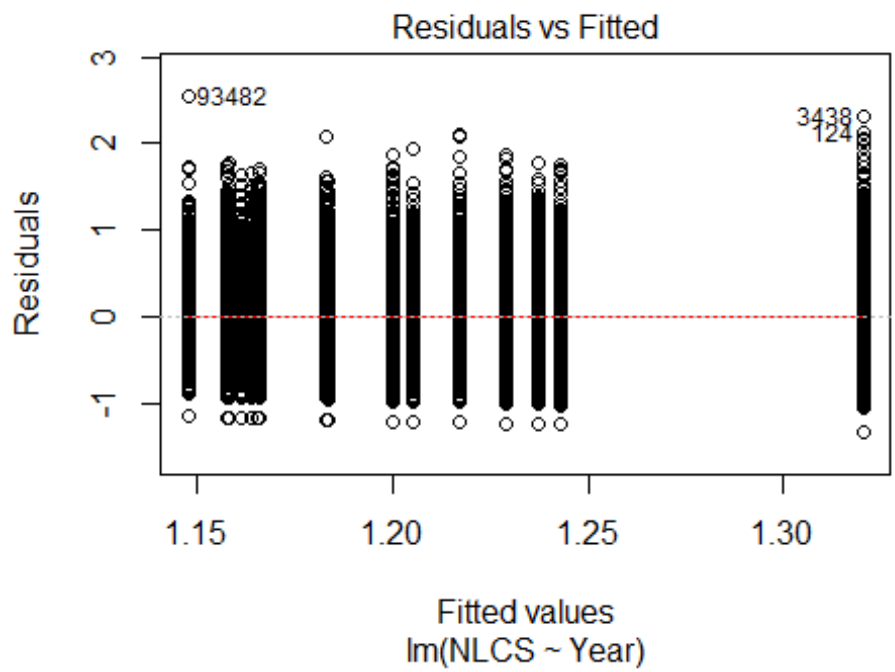
```
## Year2007           -0.06660    0.01199   -5.56  2.8e-08 ***
## Year2008           -0.05244    0.01202   -4.36  1.3e-05 ***
## Year2009           -0.07597    0.01189   -6.39  1.7e-10 ***
## Year2010           -0.08560    0.01225   -6.99  2.8e-12 ***
## Year2011           -0.07675    0.01230   -6.24  4.5e-10 ***
## Year2012           -0.08529    0.01265   -6.74  1.6e-11 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.361
## Multiple R-squared:  0.00522,    Adjusted R-squared:  0.00475
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  3 observations c(19495,33160,34469)
##   are outliers with |weight| = 0 ( < 2.8e-06);
##  3038 weights are ~= 1. The remaining 32361 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.000   0.864   0.951   0.895   0.985   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          2.82e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##                   GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.002  1           1.001
## Year             1.002 16           1.000
```



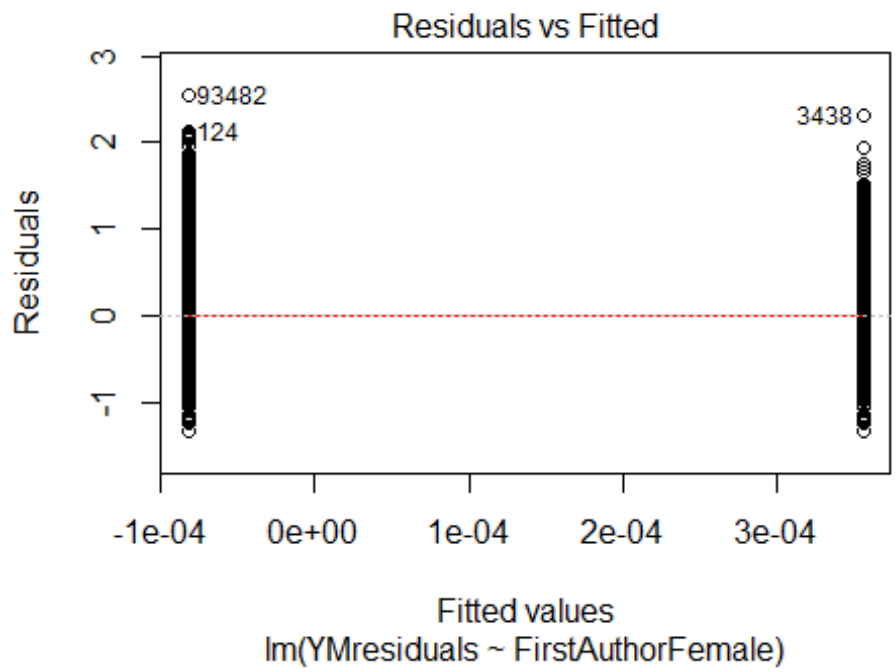

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.1781 -0.2418  0.0051  0.2424  2.0422
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.17808    0.00904  130.31  < 2e-16 ***
## LastAuthorFemale1 -0.01486    0.00602   -2.47   0.0136 *
## Year1997          -0.01353    0.01324   -1.02   0.3069
## Year1998          -0.00982    0.01265   -0.78   0.4375
## Year1999          -0.04337    0.01264   -3.43   0.0006 ***
## Year2000          -0.04183    0.01302   -3.21   0.0013 **
## Year2001          -0.05785    0.01306   -4.43  9.4e-06 ***
## Year2002          -0.06247    0.01185   -5.27  1.3e-07 ***
## Year2003          -0.05923    0.01192   -4.97  6.7e-07 ***
## Year2004          -0.07021    0.01201   -5.85  5.1e-09 ***
## Year2005          -0.08419    0.01191   -7.07  1.6e-12 ***
## Year2006          -0.08058    0.01184   -6.80  1.0e-11 ***
```

```
## Year2007          -0.06601    0.01198   -5.51  3.6e-08 ***
## Year2008          -0.05143    0.01200   -4.28  1.8e-05 ***
## Year2009          -0.07521    0.01189   -6.33  2.5e-10 ***
## Year2010          -0.08420    0.01223   -6.89  5.8e-12 ***
## Year2011          -0.07531    0.01229   -6.13  8.9e-10 ***
## Year2012          -0.08429    0.01265   -6.67  2.7e-11 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.361
## Multiple R-squared:  0.00524,    Adjusted R-squared:  0.00476
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  3 observations c(19495,33160,34469)
##   are outliers with |weight| = 0 ( < 2.8e-06);
##  2986 weights are ~= 1. The remaining 32413 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0001  0.8650  0.9510  0.8950  0.9860  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          2.82e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                  psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Sample size for the above analysis:  35402"
## [1] ""
## [1] ""
## [1] "##################################"
## [1] "Analysis of AJSC 1606"
## [1] "##################################"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 5135 4676 4740 4987 5149 4749 4832 4739 4949 5079 5436 5185 5103 5215 4974
## 2011 2012
## 5281 5019
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
```

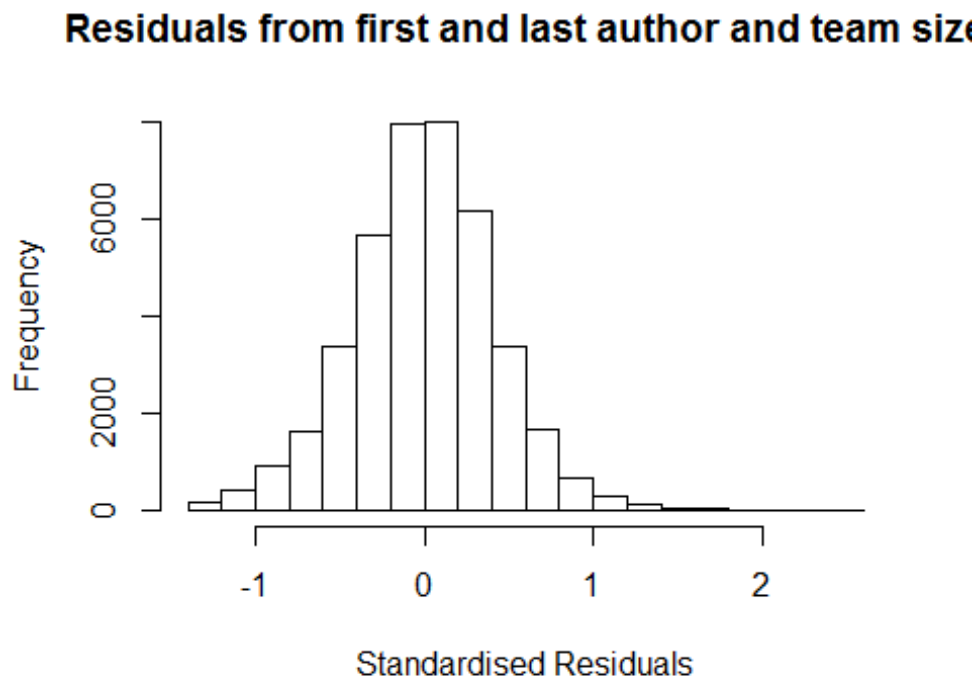
```
## 2867 2622 2738 2889 2790 2275 2767 2720 2877 2971 3144 2988 2998 3122 2970
## 2011 2012
## 3177 3077
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 2409 2175 2303 2424 2297 1889 2318 2258 2353 2420 2560 2437 2446 2554 2451
## 2011 2012
## 2557 2540
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
##  Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 330, df = 16, p-value <2e-16
```




```
##
##  Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 76, df = 1, p-value <2e-16
```

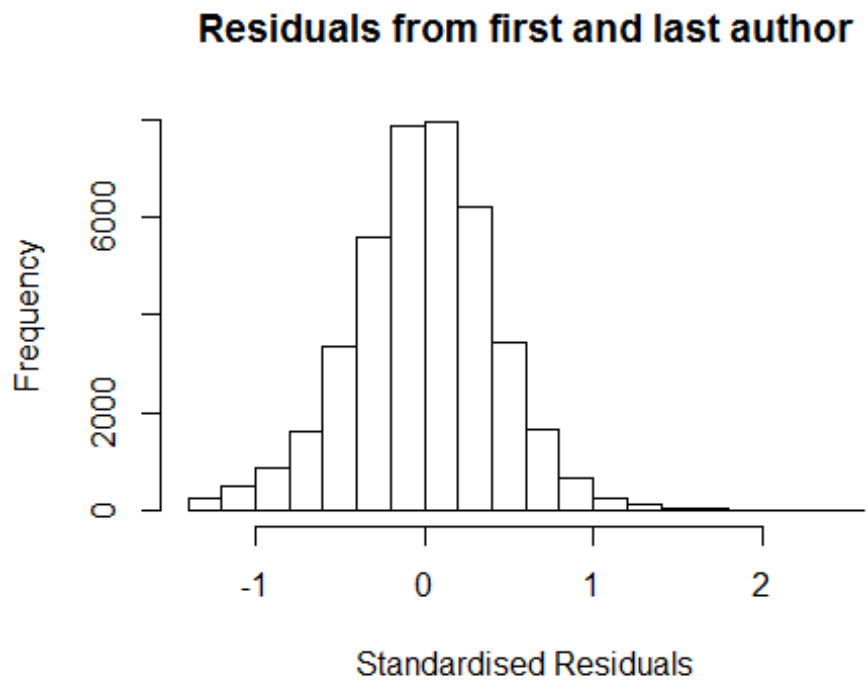



```
## [1] "Female first author team size 2018 geometric mean: 3.82347129770221"
## [1] "Male first author team size 2018 geometric mean: 3.39385941942971"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 550000, p-value = 1e-04
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 3.86868641074394"
## [1] "Male last author team size 2018 geometric mean: 3.42664317017058"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 410000, p-value = 5e-04
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.018  1           1.009
## LastAuthorFemale  1.009  1           1.004
## UniqueAuthors     1.047  4           1.006
## Year              1.045 16           1.001
```



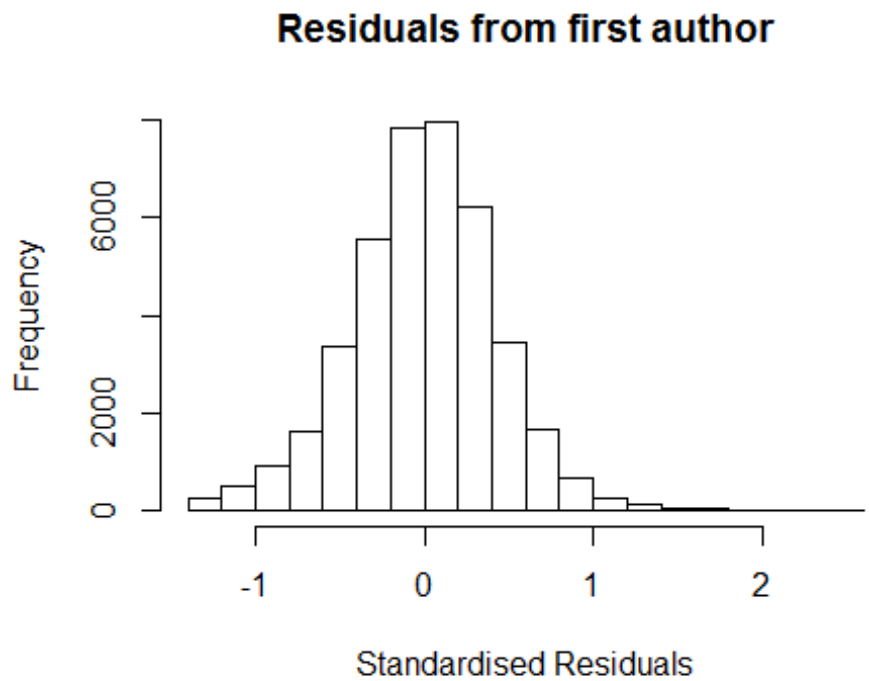

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
UniqueAuthors +
##     Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##     k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.3852 -0.2617  0.0033  0.2646  2.4966
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.15431    0.01315   87.80  < 2e-16 ***
## FirstAuthorFemale1 -0.01387    0.00506   -2.74   0.0061 **
## LastAuthorFemale1  -0.00114    0.00642   -0.18   0.8595
## UniqueAuthors2      0.16809    0.00960   17.51  < 2e-16 ***
## UniqueAuthors3      0.19091    0.00965   19.78  < 2e-16 ***
## UniqueAuthors4      0.20986    0.01008   20.82  < 2e-16 ***
## UniqueAuthors5      0.23090    0.00993   23.26  < 2e-16 ***
## Year1997           -0.06488    0.01385   -4.68  2.8e-06 ***
## Year1998           -0.08209    0.01387   -5.92  3.2e-09 ***
## Year1999           -0.10890    0.01338   -8.14  4.1e-16 ***
```

```
## Year2000           -0.11385    0.01334   -8.54  < 2e-16 ***
## Year2001           -0.08428    0.01400   -6.02  1.8e-09 ***
## Year2002           -0.12370    0.01332   -9.29  < 2e-16 ***
## Year2003           -0.13765    0.01323  -10.41  < 2e-16 ***
## Year2004           -0.16861    0.01348  -12.51  < 2e-16 ***
## Year2005           -0.16370    0.01312  -12.48  < 2e-16 ***
## Year2006           -0.17892    0.01296  -13.81  < 2e-16 ***
## Year2007           -0.17659    0.01335  -13.23  < 2e-16 ***
## Year2008           -0.17885    0.01311  -13.64  < 2e-16 ***
## Year2009           -0.15644    0.01301  -12.03  < 2e-16 ***
## Year2010           -0.17702    0.01299  -13.63  < 2e-16 ***
## Year2011           -0.17346    0.01306  -13.28  < 2e-16 ***
## Year2012           -0.18877    0.01333  -14.16  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.392
## Multiple R-squared:  0.0338, Adjusted R-squared:  0.0332
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
##  16 observations
c(63,619,848,852,1540,1731,1944,7713,9389,9394,11478,14687,18978,27961,31848,
38879)
##   are outliers with |weight| = 0 ( < 2.5e-06);
##  3406 weights are ~= 1. The remaining 36969 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0001  0.8660  0.9500  0.8950  0.9860  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          2.48e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.010  1           1.005
## LastAuthorFemale  1.008  1           1.004
## Year              1.011 16           1.000
```



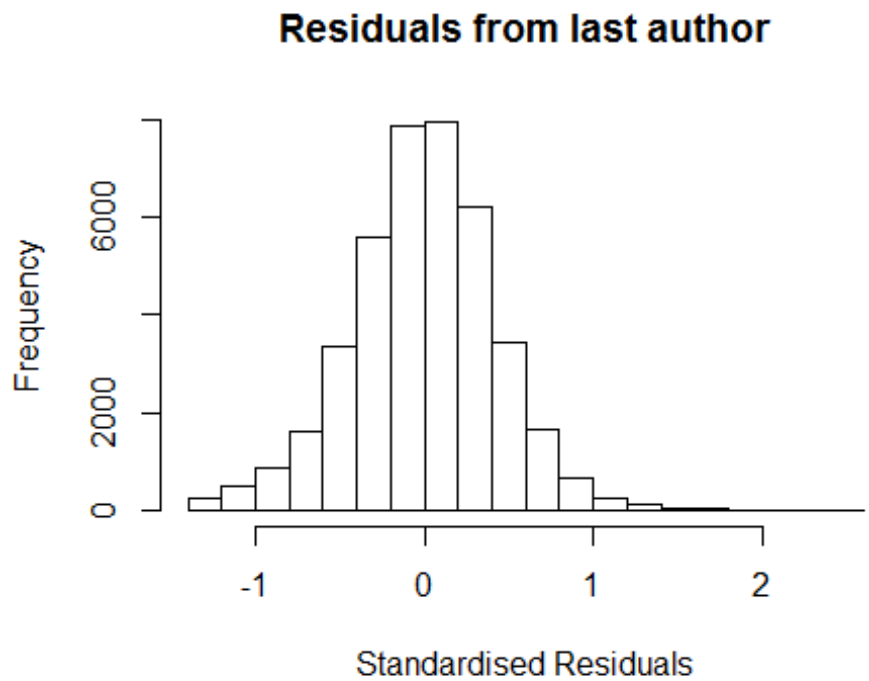

```
## [1] "List of  1 outliers with residuals above  2.5"
##          ScopusId  NLCS Year OneField Fields residuals
## 93482 84867288643 3.693 2012     1606      4     2.543
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.32334 -0.26366  0.00456  0.26422  2.54318
##
## Coefficients:
##                      Estimate Std. Error t value Pr(>|t|)
## (Intercept)          1.318780   0.010416  126.61  < 2e-16 ***
## FirstAuthorFemale1   0.000325   0.005049    0.06     0.95
## LastAuthorFemale1    0.004239   0.006416    0.66     0.51
## Year1997            -0.070172   0.013949   -5.03  4.9e-07 ***
## Year1998            -0.082810   0.013922   -5.95  2.7e-09 ***
## Year1999            -0.115818   0.013451   -8.61  < 2e-16 ***
## Year2000            -0.106786   0.013381   -7.98  1.5e-15 ***
## Year2001            -0.073027   0.014074   -5.19  2.1e-07 ***
## Year2002            -0.114338   0.013401   -8.53  < 2e-16 ***
## Year2003            -0.125751   0.013296   -9.46  < 2e-16 ***
## Year2004            -0.156918   0.013569  -11.56  < 2e-16 ***
## Year2005            -0.150311   0.013231  -11.36  < 2e-16 ***
```

```
## Year2006           -0.165109   0.013094  -12.61  < 2e-16 ***
## Year2007           -0.158441   0.013421  -11.81  < 2e-16 ***
## Year2008           -0.160581   0.013181  -12.18  < 2e-16 ***
## Year2009           -0.139459   0.013066  -10.67  < 2e-16 ***
## Year2010           -0.157634   0.013038  -12.09  < 2e-16 ***
## Year2011           -0.153208   0.013129  -11.67  < 2e-16 ***
## Year2012           -0.168959   0.013423  -12.59  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.394
## Multiple R-squared:  0.0114, Adjusted R-squared:  0.0109
## Convergence in 10 IRWLS iterations
##
## Robustness weights:
##  14 observations
c(63,619,848,852,1540,1731,4909,7713,9389,9394,11478,14687,31848,38879)
##   are outliers with |weight| = 0 ( < 2.5e-06);
##  3337 weights are ~= 1. The remaining 37040 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0007  0.8670  0.9510  0.8940  0.9860  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          2.48e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                    psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.006  1           1.003
## Year              1.006 16           1.000
```




```
## [1] "List of  1 outliers with residuals above  2.5"
##          ScopusId  NLCS Year OneField Fields residuals
## 93482 84867288643 3.693 2012     1606      4     2.543
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.31972 -0.26330  0.00491  0.26437  2.54264
##
## Coefficients:
##                      Estimate Std. Error t value Pr(>|t|)
## (Intercept)          1.319095   0.010400  126.83  < 2e-16 ***
## FirstAuthorFemale1   0.000624   0.005042    0.12      0.9
## Year1997            -0.070187   0.013949   -5.03  4.9e-07 ***
## Year1998            -0.082782   0.013922   -5.95  2.8e-09 ***
## Year1999            -0.115695   0.013450   -8.60  < 2e-16 ***
## Year2000            -0.106675   0.013379   -7.97  1.6e-15 ***
## Year2001            -0.072918   0.014073   -5.18  2.2e-07 ***
## Year2002            -0.114325   0.013401   -8.53  < 2e-16 ***
## Year2003            -0.125637   0.013294   -9.45  < 2e-16 ***
## Year2004            -0.156807   0.013569  -11.56  < 2e-16 ***
## Year2005            -0.150174   0.013229  -11.35  < 2e-16 ***
## Year2006            -0.165037   0.013093  -12.60  < 2e-16 ***
```

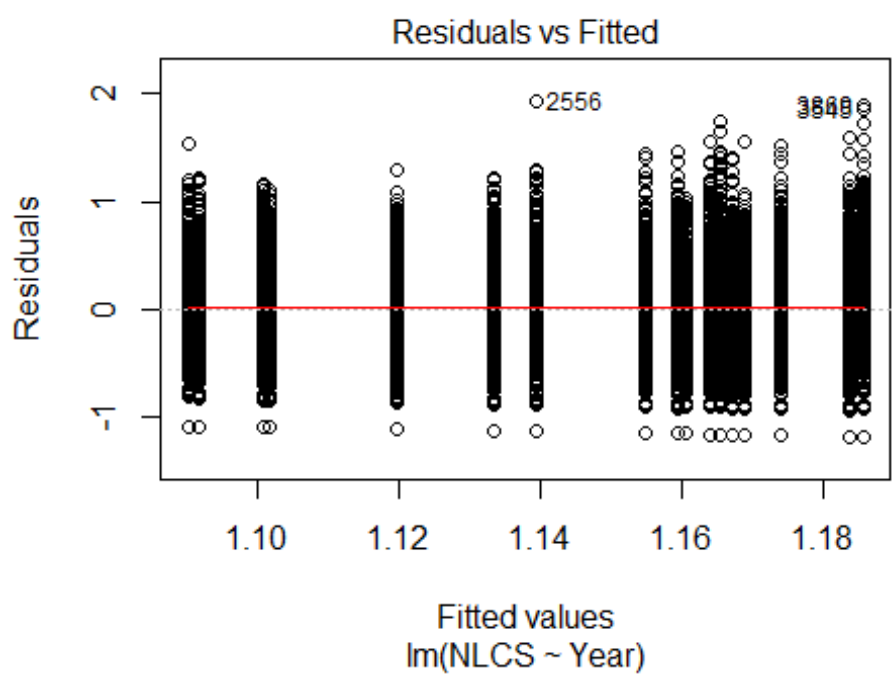
```
## Year2007           -0.158283   0.013418  -11.80  < 2e-16 ***
## Year2008           -0.160376   0.013175  -12.17  < 2e-16 ***
## Year2009           -0.139295   0.013065  -10.66  < 2e-16 ***
## Year2010           -0.157507   0.013039  -12.08  < 2e-16 ***
## Year2011           -0.153010   0.013127  -11.66  < 2e-16 ***
## Year2012           -0.168731   0.013418  -12.57  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.394
## Multiple R-squared:  0.0114, Adjusted R-squared:  0.011
## Convergence in 10 IRWLS iterations
##
## Robustness weights:
##  14 observations
c(63,619,848,852,1540,1731,4909,7713,9389,9394,11478,14687,31848,38879)
##   are outliers with |weight| = 0 ( < 2.5e-06);
##  3346 weights are ~= 1. The remaining 37031 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0005  0.8670  0.9510  0.8940  0.9860  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          2.48e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##                   GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.005  1           1.002
## Year             1.005 16           1.000
```



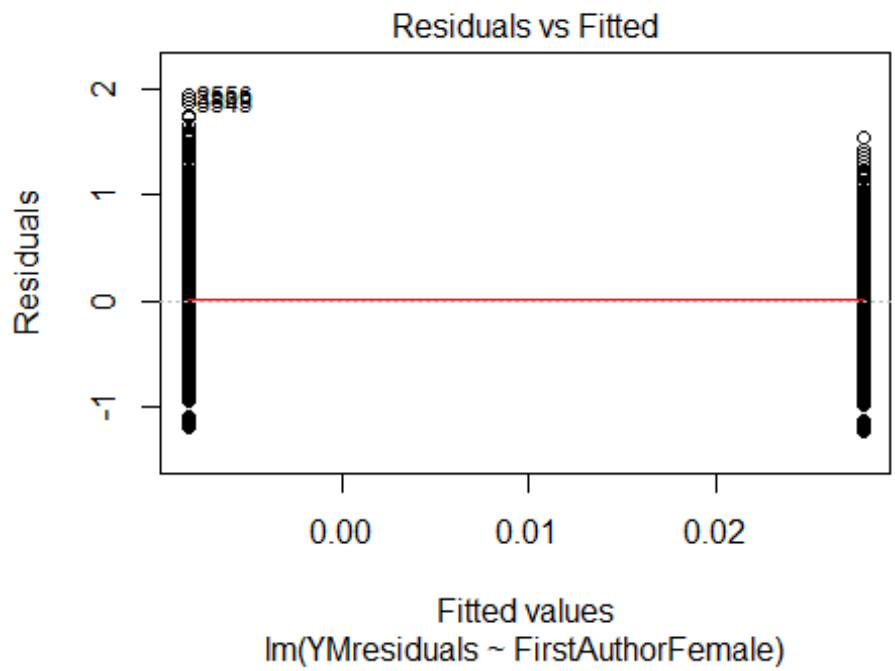

```
## [1] "List of  1 outliers with residuals above  2.5"
##          ScopusId  NLCS Year OneField Fields residuals
## 93482 84867288643 3.693 2012     1606      4     2.543
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.3231 -0.2636  0.0046  0.2643  2.5431
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.31883    0.01036  127.26  < 2e-16 ***
## LastAuthorFemale1  0.00427    0.00641    0.67      0.5
## Year1997          -0.07017    0.01395   -5.03  4.9e-07 ***
## Year1998          -0.08281    0.01392   -5.95  2.7e-09 ***
## Year1999          -0.11582    0.01345   -8.61  < 2e-16 ***
## Year2000          -0.10678    0.01338   -7.98  1.5e-15 ***
## Year2001          -0.07302    0.01407   -5.19  2.1e-07 ***
## Year2002          -0.11434    0.01340   -8.53  < 2e-16 ***
## Year2003          -0.12575    0.01330   -9.46  < 2e-16 ***
## Year2004          -0.15691    0.01357  -11.56  < 2e-16 ***
## Year2005          -0.15031    0.01323  -11.36  < 2e-16 ***
## Year2006          -0.16510    0.01309  -12.61  < 2e-16 ***
```

```
## Year2007          -0.15843    0.01342  -11.81  < 2e-16 ***
## Year2008          -0.16057    0.01318  -12.18  < 2e-16 ***
## Year2009          -0.13945    0.01307  -10.67  < 2e-16 ***
## Year2010          -0.15761    0.01304  -12.09  < 2e-16 ***
## Year2011          -0.15320    0.01313  -11.67  < 2e-16 ***
## Year2012          -0.16895    0.01342  -12.59  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.395
## Multiple R-squared:  0.0114, Adjusted R-squared:  0.011
## Convergence in 10 IRWLS iterations
##
## Robustness weights:
##  14 observations
c(63,619,848,852,1540,1731,4909,7713,9389,9394,11478,14687,31848,38879)
##   are outliers with |weight| = 0 ( < 2.5e-06);
##  3336 weights are ~= 1. The remaining 37041 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0007  0.8670  0.9510  0.8940  0.9860  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          2.48e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Sample size for the above analysis:  40391"
## [1] ""
## [1] ""
## [1] "####################################"
## [1] "Analysis of AJSC 1607"
## [1] "####################################"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 1229 1286 1184 1360 1326 1258 1250 1149 1283 1338 1352 1404 1400 1492 1497
## 2011 2012
## 1521 1491
##
```

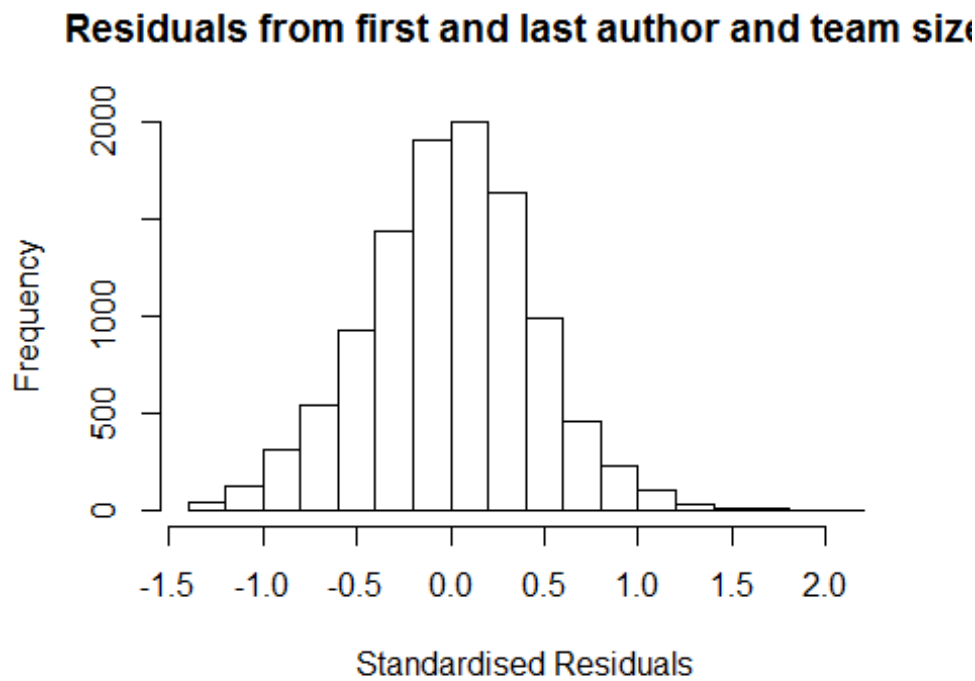
```
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  675  728  666  734  693  619  742  682  755  742  774  828  815  859  898
## 2011 2012
##  906  876
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  576  616  551  617  577  512  610  547  640  630  649  695  654  699  712
## 2011 2012
##  730  700
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
##  Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 180, df = 16, p-value <2e-16
```




```
##
##  Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 38, df = 1, p-value = 7e-10
```

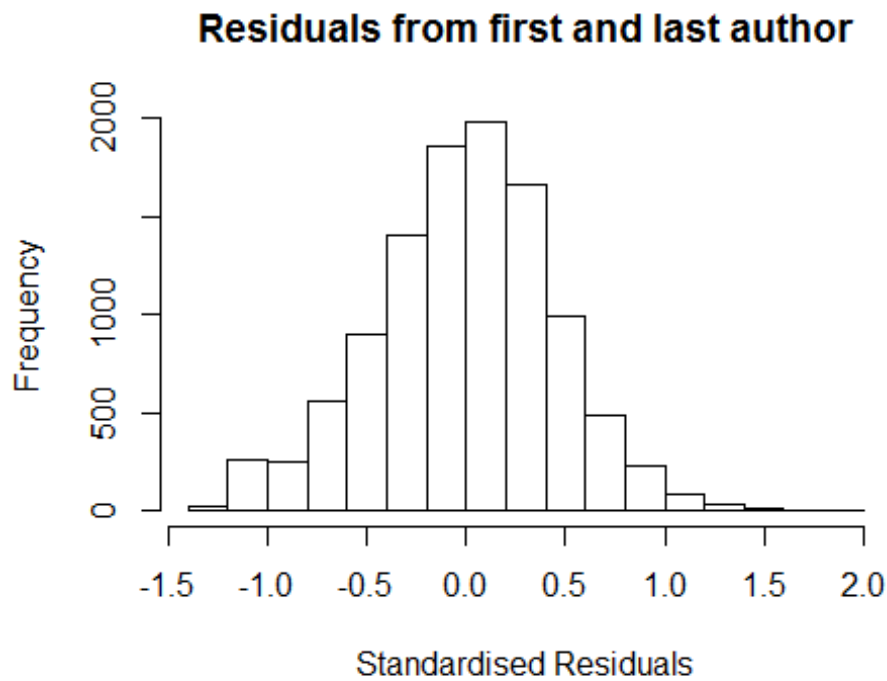



```
## [1] "Female first author team size 2018 geometric mean: 4.28183258044529"
## [1] "Male first author team size 2018 geometric mean: 3.62385086861484"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 74000, p-value = 7e-04
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 4.1257673003727"
## [1] "Male last author team size 2018 geometric mean: 3.73562223167455"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 58000, p-value = 0.06
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.027  1           1.013
## LastAuthorFemale  1.014  1           1.007
## UniqueAuthors     1.065  4           1.008
## Year              1.075 16           1.002
```



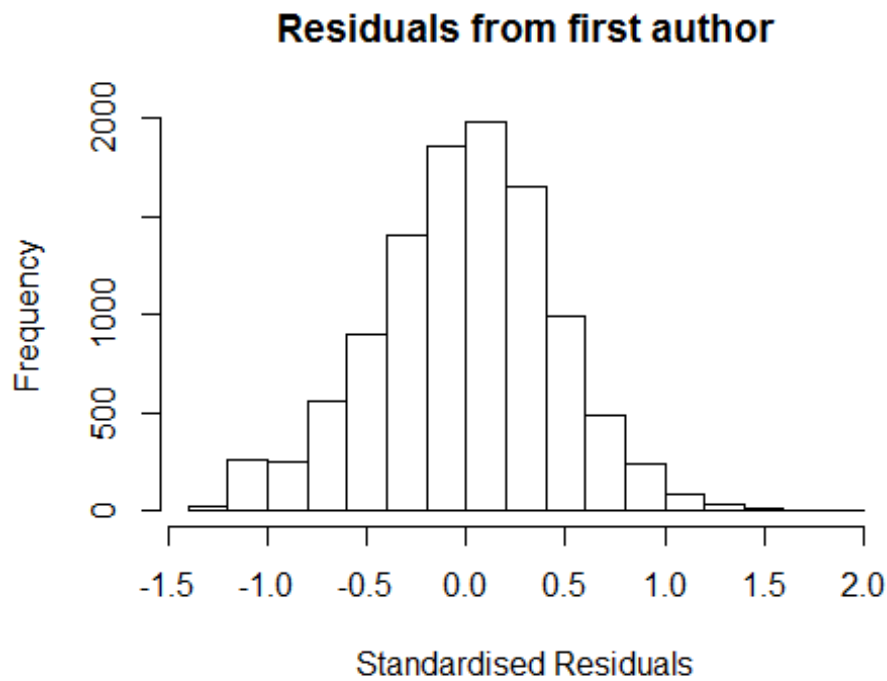

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
UniqueAuthors +
##     Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##     k.max = 1000))
##  \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.26348 -0.28499  0.00953  0.28781  2.17001
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         0.92377    0.03080   29.99  < 2e-16 ***
## FirstAuthorFemale1  0.02667    0.00987    2.70  0.00689 **
## LastAuthorFemale1   0.00786    0.01208    0.65  0.51510
## UniqueAuthors2      0.22807    0.02328    9.80  < 2e-16 ***
## UniqueAuthors3      0.28573    0.02329   12.27  < 2e-16 ***
## UniqueAuthors4      0.27992    0.02364   11.84  < 2e-16 ***
## UniqueAuthors5      0.32857    0.02331   14.10  < 2e-16 ***
## Year1997           -0.02277    0.03161   -0.72  0.47124
## Year1998            0.00555    0.03201    0.17  0.86222
## Year1999            0.00691    0.03020    0.23  0.81915
```

```
## Year2000            0.02526    0.03021    0.84  0.40310
## Year2001           -0.00447    0.03039   -0.15  0.88308
## Year2002           -0.03936    0.02943   -1.34  0.18103
## Year2003           -0.01901    0.02956   -0.64  0.52004
## Year2004           -0.03678    0.02897   -1.27  0.20421
## Year2005           -0.01791    0.02781   -0.64  0.51960
## Year2006           -0.02255    0.02764   -0.82  0.41459
## Year2007           -0.05124    0.02830   -1.81  0.07018 .
## Year2008           -0.09120    0.02859   -3.19  0.00143 **
## Year2009           -0.09635    0.02810   -3.43  0.00061 ***
## Year2010           -0.09505    0.02891   -3.29  0.00101 **
## Year2011           -0.12618    0.02771   -4.55  5.3e-06 ***
## Year2012           -0.11170    0.02942   -3.80  0.00015 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.426
## Multiple R-squared:  0.0442, Adjusted R-squared:  0.0422
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
##  observation 1110 is an outlier with |weight| = 0 ( < 9.3e-06);
##  912 weights are ~= 1. The remaining 9802 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0159  0.8680  0.9500  0.9000  0.9860  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          9.33e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.019  1           1.009
## LastAuthorFemale  1.010  1           1.005
## Year              1.022 16           1.001
```



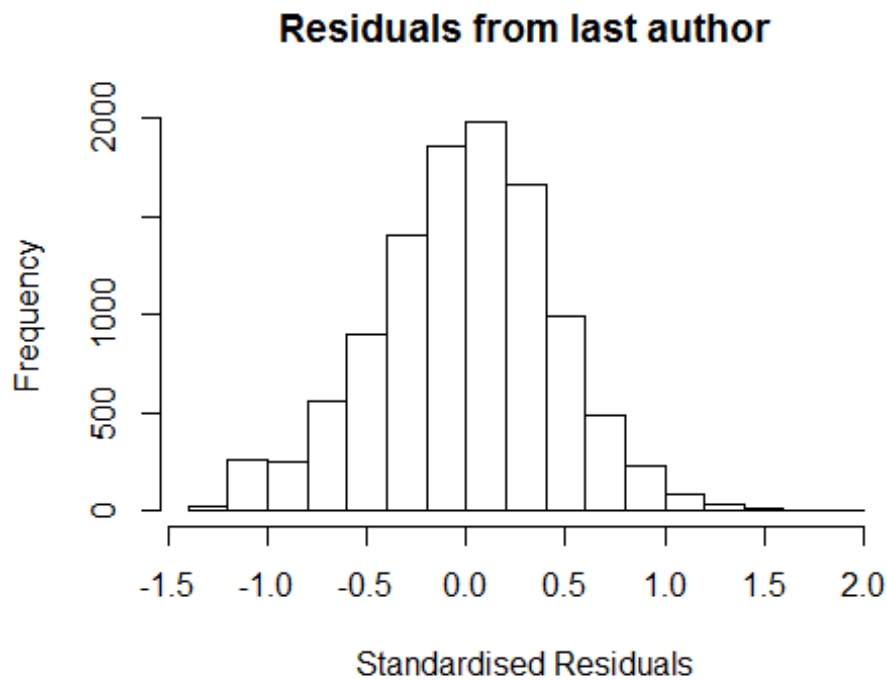

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.2536 -0.2885  0.0111  0.2905  1.9272
##
## Coefficients:
##                      Estimate Std. Error t value Pr(>|t|)
## (Intercept)          1.159051   0.023413   49.50  < 2e-16 ***
## FirstAuthorFemale1   0.044025   0.009915    4.44  9.1e-06 ***
## LastAuthorFemale1    0.013207   0.012104    1.09  0.27525
## Year1997            -0.015223   0.032022   -0.48  0.63451
## Year1998             0.013428   0.032290    0.42  0.67752
## Year1999             0.004228   0.030528    0.14  0.88986
## Year2000             0.037345   0.030323    1.23  0.21815
## Year2001             0.000697   0.030655    0.02  0.98186
## Year2002            -0.021365   0.029721   -0.72  0.47225
## Year2003             0.000808   0.029902    0.03  0.97844
## Year2004            -0.013163   0.029371   -0.45  0.65404
## Year2005             0.007869   0.028041    0.28  0.77900
```

```
## Year2006           -0.000563   0.028007   -0.02  0.98397
## Year2007           -0.026342   0.028667   -0.92  0.35817
## Year2008           -0.062468   0.028932   -2.16  0.03086 *
## Year2009           -0.067663   0.028220   -2.40  0.01651 *
## Year2010           -0.063839   0.029210   -2.19  0.02887 *
## Year2011           -0.093342   0.027962   -3.34  0.00085 ***
## Year2012           -0.079344   0.029706   -2.67  0.00757 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.43
## Multiple R-squared:  0.00853,    Adjusted R-squared:  0.00686
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
##  912 weights are ~= 1. The remaining 9803 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0074  0.8670  0.9500  0.8990  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          9.33e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.015  1           1.008
## Year              1.015 16           1.000
```




```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.2429 -0.2878  0.0111  0.2909  1.9257
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.160600   0.023359   49.68  < 2e-16 ***
## FirstAuthorFemale1  0.045001   0.009905    4.54  5.6e-06 ***
## Year1997           -0.015303   0.032025   -0.48   0.6328
## Year1998            0.013301   0.032285    0.41   0.6804
## Year1999            0.004132   0.030534    0.14   0.8924
## Year2000            0.037335   0.030319    1.23   0.2182
## Year2001            0.000451   0.030650    0.01   0.9883
## Year2002           -0.021466   0.029717   -0.72   0.4701
## Year2003            0.000896   0.029894    0.03   0.9761
## Year2004           -0.013332   0.029365   -0.45   0.6498
## Year2005            0.008240   0.028029    0.29   0.7688
## Year2006           -0.000576   0.028009   -0.02   0.9836
```

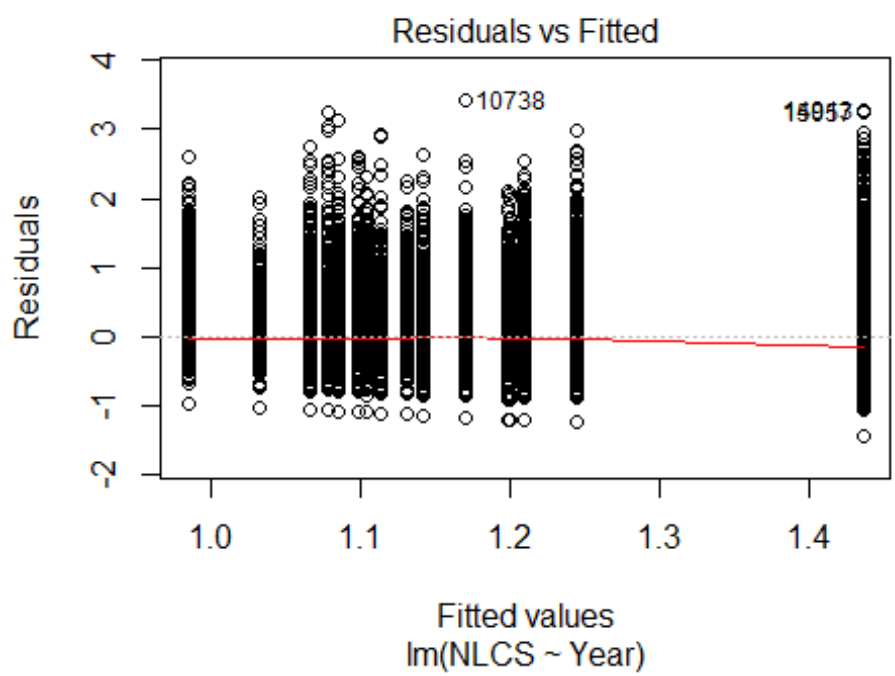
```
## Year2007           -0.026074   0.028668   -0.91   0.3631
## Year2008           -0.062217   0.028928   -2.15   0.0315 *
## Year2009           -0.067317   0.028214   -2.39   0.0171 *
## Year2010           -0.063471   0.029198   -2.17   0.0297 *
## Year2011           -0.092746   0.027939   -3.32   0.0009 ***
## Year2012           -0.078664   0.029690   -2.65   0.0081 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.43
## Multiple R-squared:  0.00841,    Adjusted R-squared:  0.00683
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
##  916 weights are ~= 1. The remaining 9799 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0078  0.8680  0.9500  0.8990  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          9.33e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##                   GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.007  1           1.003
## Year             1.007 16           1.000
```



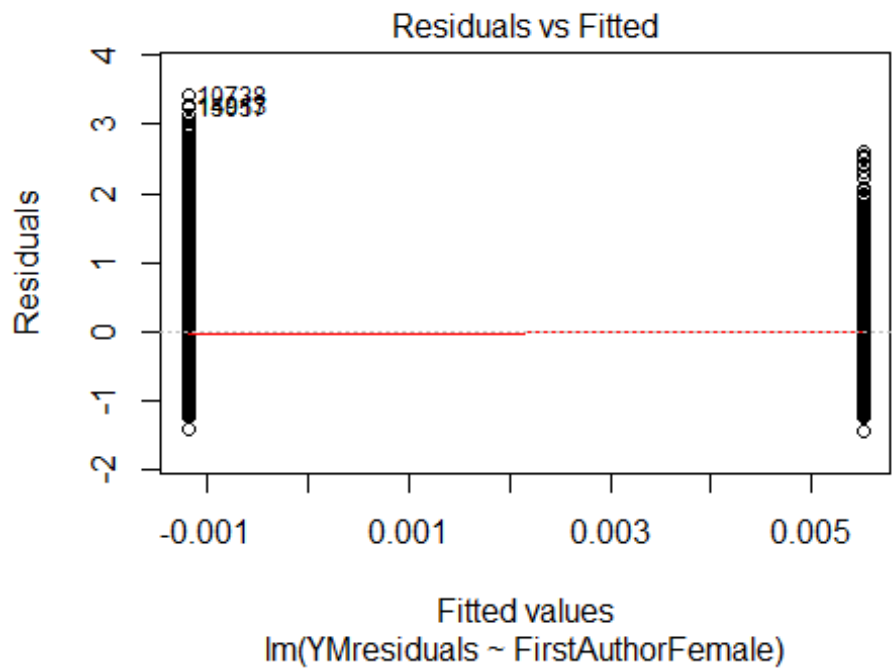

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.2238 -0.2887  0.0111  0.2908  1.9204
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.16486    0.02347   49.64   <2e-16 ***
## LastAuthorFemale1   0.01792    0.01210    1.48   0.1388
## Year1997           -0.01426    0.03213   -0.44   0.6573
## Year1998            0.01547    0.03236    0.48   0.6326
## Year1999            0.00728    0.03058    0.24   0.8119
## Year2000            0.04101    0.03034    1.35   0.1764
## Year2001            0.00396    0.03066    0.13   0.8971
## Year2002           -0.01609    0.02973   -0.54   0.5883
## Year2003            0.00678    0.02996    0.23   0.8208
## Year2004           -0.00966    0.02940   -0.33   0.7425
## Year2005            0.01171    0.02808    0.42   0.6768
## Year2006            0.00405    0.02800    0.14   0.8851
```

```
## Year2007          -0.02058    0.02870   -0.72   0.4732
## Year2008          -0.05658    0.02896   -1.95   0.0507 .
## Year2009          -0.06404    0.02824   -2.27   0.0234 *
## Year2010          -0.05870    0.02923   -2.01   0.0446 *
## Year2011          -0.08864    0.02796   -3.17   0.0015 **
## Year2012          -0.07398    0.02970   -2.49   0.0128 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.43
## Multiple R-squared:  0.00679,    Adjusted R-squared:  0.00521
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
##  907 weights are ~= 1. The remaining 9808 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0086  0.8670  0.9500  0.8990  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          9.33e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Sample size for the above analysis:  10715"
## [1] ""
## [1] ""
## [1] "####################################"
## [1] "Analysis of AJSC 1700"
## [1] "####################################"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  999  760  778  709  810  808  752  490  451  536  657  758  913 1117  889
## 2011 2012
##  864  828
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  653  467  505  491  575  511  520  321  317  363  452  531  642  817  599
## 2011 2012
```

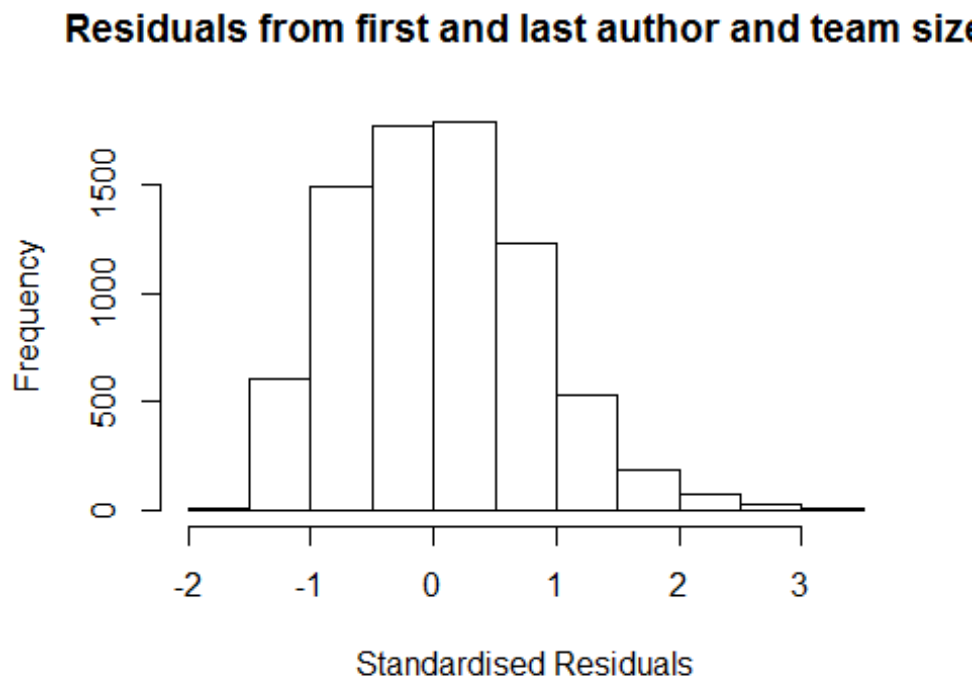
```
##  555  550
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  587  424  445  441  506  457  446  266  287  317  397  448  545  689  508
## 2011 2012
##  471  466
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
##  Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 230, df = 16, p-value <2e-16
```




```
##
##  Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 2.8, df = 1, p-value = 0.09
```



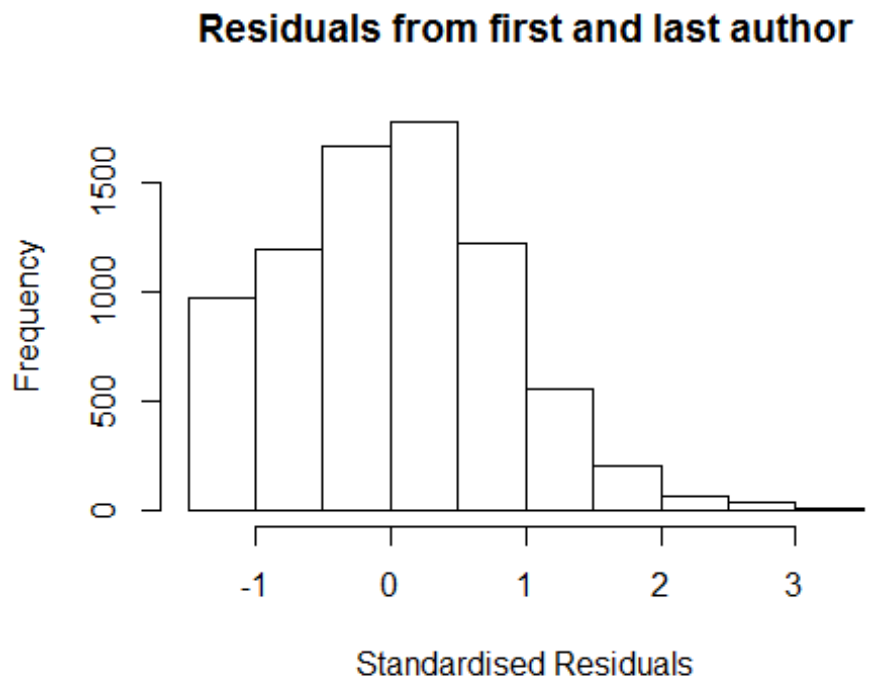

```
## [1] "Female first author team size 2018 geometric mean: 2.2231950008727"
## [1] "Male first author team size 2018 geometric mean: 2.11151208224127"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 24000, p-value = 0.6
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 2.19560498202231"
## [1] "Male last author team size 2018 geometric mean: 2.11786400322357"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 25000, p-value = 0.5
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.438  1           1.199
## LastAuthorFemale  1.438  1           1.199
## UniqueAuthors     1.094  4           1.011
## Year              1.093 16           1.003
```




```
## [1] "List of  33 outliers with residuals above  2.5"
##          ScopusId  NLCS Year OneField Fields residuals
## 1572   0031103679 4.075 1997     1700      1     2.984
## 1574   0031104254 4.116 1997     1700      1     3.025
## 1726   0031190973 3.602 1997     1700      2     2.744
## 1774   0142051871 4.315 1997     1700      1     3.457
## 1985   0032269108 3.702 1998     1700      1     2.578
## 2434   0032058184 3.621 1998     1700      1     2.698
## 2595   0032108018 3.662 1998     1700      2     2.771
## 5185   0035424017 3.551 2001     1700      1     2.503
## 6760   0037623983 3.818 2003     1700      2     2.749
## 8647  29144523061 3.760 2006     1700      7     2.618
## 9488  34247481878 4.038 2007     1700      2     2.964
## 9489  34247500374 4.020 2007     1700      2     3.178
## 10738 37549003336 4.591 2008     1700      1     3.459
## 12116 85008044987 4.215 2009     1700      1     3.027
## 12881 77950364314 3.498 2010     1700      1     2.572
## 13124 73649114265 3.748 2010     1700      1     2.589
## 13652 80053974840 3.721 2011     1700      2     2.548
## 13915 79955370378 3.941 2011     1700      1     2.768
## 14045 79951942070 3.898 2011     1700      1     2.725
## 14533 84872070696 4.014 2012     1700      1     2.618
## 14597 84865517174 4.039 2012     1700      1     2.759
## 14620 84857713208 3.859 2012     1700      2     2.569
## 14671 84864696820 4.212 2012     1700      1     3.176
## 14734 84863758078 4.014 2012     1700      1     2.618
## 14801 84865483551 4.316 2012     1700      1     2.920
```

```
## 14851 84863765049 4.149 2012     1700      1     2.880
## 15047 84861902326 4.058 2012     1700      1     2.789
## 15057 84857641127 4.677 2012     1700      1     3.408
## 15074 84857657877 3.882 2012     1700      1     2.613
## 15139 84986596759 3.778 2012     1700      1     2.509
## 15177 84856033479 4.051 2012     1700      1     3.015
## 15179 84856041489 4.001 2012     1700      1     2.711
## 15216 80052530633 4.385 2012     1700      2     3.349
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
UniqueAuthors +
##     Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##     k.max = 1000))
##  \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.57607 -0.54187 -0.00444  0.53965  3.45855
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)          0.7881     0.0351   22.42  < 2e-16 ***
## FirstAuthorFemale1   0.0107     0.0287    0.37  0.70822
## LastAuthorFemale1    0.0206     0.0282    0.73  0.46474
## UniqueAuthors2       0.2329     0.0219   10.64  < 2e-16 ***
## UniqueAuthors3       0.3594     0.0283   12.68  < 2e-16 ***
## UniqueAuthors4       0.3962     0.0376   10.52  < 2e-16 ***
## UniqueAuthors5       0.5190     0.0368   14.09  < 2e-16 ***
## Year1997             0.0703     0.0513    1.37  0.17095
## Year1998             0.1033     0.0517    2.00  0.04568 *
## Year1999             0.0964     0.0510    1.89  0.05868 .
## Year2000             0.1232     0.0485    2.54  0.01107 *
## Year2001             0.0273     0.0491    0.56  0.57750
## Year2002             0.1927     0.0500    3.86  0.00012 ***
## Year2003             0.0484     0.0559    0.86  0.38725
## Year2004             0.0547     0.0525    1.04  0.29761
## Year2005             0.1882     0.0541    3.48  0.00051 ***
## Year2006             0.1211     0.0504    2.41  0.01619 *
## Year2007             0.0534     0.0476    1.12  0.26223
## Year2008             0.1114     0.0464    2.40  0.01640 *
## Year2009             0.0403     0.0446    0.90  0.36594
## Year2010             0.1375     0.0516    2.67  0.00771 **
## Year2011             0.1519     0.0546    2.78  0.00541 **
## Year2012             0.2483     0.0626    3.97  7.3e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.793
## Multiple R-squared:  0.0579, Adjusted R-squared:  0.0552
```

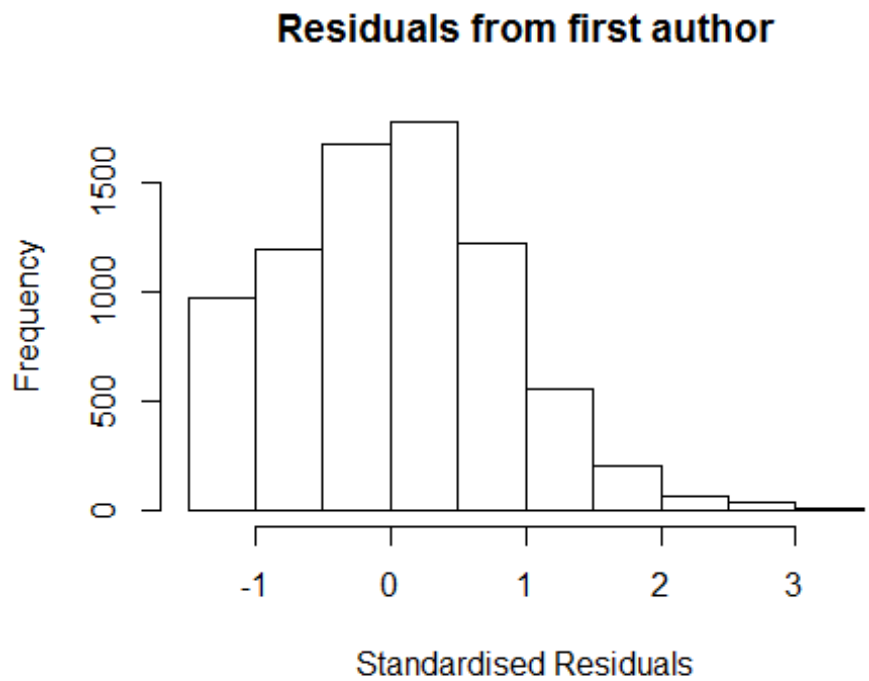
```
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
##  604 weights are ~= 1. The remaining 7096 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0175  0.8830  0.9500  0.9120  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.30e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.413  1           1.189
## LastAuthorFemale  1.411  1           1.188
## Year              1.015 16           1.000
```




```
## [1] "List of  45 outliers with residuals above  2.5"
##          ScopusId  NLCS Year OneField Fields residuals
## 627    0000289556 3.577 1996     1700      2     2.647
## 1572   0031103679 4.075 1997     1700      1     3.082
## 1574   0031104254 4.116 1997     1700      1     3.123
## 1612   0031272525 3.611 1997     1700      1     2.618
## 1726   0031190973 3.602 1997     1700      2     2.609
## 1774   0142051871 4.315 1997     1700      1     3.322
## 1985   0032269108 3.702 1998     1700      1     2.676
## 2434   0032058184 3.621 1998     1700      1     2.554
## 2576   0032057865 3.542 1998     1700      2     2.516
## 2595   0032108018 3.662 1998     1700      2     2.636
## 5185   0035424017 3.551 2001     1700      1     2.557
## 6760   0037623983 3.818 2003     1700      2     2.808
## 8647  29144523061 3.760 2006     1700      7     2.658
## 9488  34247481878 4.038 2007     1700      2     2.976
## 9489  34247500374 4.020 2007     1700      2     2.958
## 10007 52149117439 3.641 2008     1700      2     2.519
## 10408 47749140025 3.719 2008     1700      1     2.572
## 10738 37549003336 4.591 2008     1700      1     3.469
## 11104 70350769233 3.653 2009     1700      1     2.597
## 12116 85008044987 4.215 2009     1700      1     3.175
## 13124 73649114265 3.748 2010     1700      1     2.592
## 13527 80053551322 3.818 2011     1700      1     2.648
## 13652 80053974840 3.721 2011     1700      2     2.551
## 13915 79955370378 3.941 2011     1700      1     2.771
## 14045 79951942070 3.898 2011     1700      1     2.728
```

```
## 14348 84872196709 3.890 2012     1700      1     2.651
## 14478 84870549803 3.912 2012     1700      1     2.673
## 14533 84872070696 4.014 2012     1700      1     2.775
## 14597 84865517174 4.039 2012     1700      1     2.775
## 14620 84857713208 3.859 2012     1700      2     2.603
## 14671 84864696820 4.212 2012     1700      1     2.973
## 14734 84863758078 4.014 2012     1700      1     2.775
## 14801 84865483551 4.316 2012     1700      1     3.077
## 14818 84865522374 3.812 2012     1700      1     2.573
## 14851 84863765049 4.149 2012     1700      1     2.910
## 15047 84861902326 4.058 2012     1700      1     2.819
## 15057 84857641127 4.677 2012     1700      1     3.438
## 15074 84857657877 3.882 2012     1700      1     2.643
## 15135 84856912379 3.812 2012     1700      1     2.573
## 15139 84986596759 3.778 2012     1700      1     2.539
## 15177 84856033479 4.051 2012     1700      1     2.812
## 15179 84856041489 4.001 2012     1700      1     2.745
## 15181 84857712777 3.852 2012     1700      1     2.588
## 15182 84859369652 3.752 2012     1700      1     2.513
## 15216 80052530633 4.385 2012     1700      2     3.146
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.28076 -0.57095  0.00152  0.54916  3.46898
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)          0.9137     0.0355   25.77  < 2e-16 ***
## FirstAuthorFemale1   0.0252     0.0289    0.87  0.38324
## LastAuthorFemale1    0.0167     0.0283    0.59  0.55634
## Year1997             0.0797     0.0535    1.49  0.13621
## Year1998             0.1118     0.0527    2.12  0.03386 *
## Year1999             0.1122     0.0525    2.14  0.03269 *
## Year2000             0.1812     0.0490    3.70  0.00022 ***
## Year2001             0.0800     0.0500    1.60  0.10987
## Year2002             0.2613     0.0509    5.14  2.9e-07 ***
## Year2003             0.0965     0.0578    1.67  0.09497 .
## Year2004             0.0958     0.0546    1.76  0.07923 .
## Year2005             0.2558     0.0556    4.60  4.3e-06 ***
## Year2006             0.1883     0.0515    3.66  0.00026 ***
## Year2007             0.1486     0.0489    3.04  0.00238 **
## Year2008             0.2083     0.0473    4.40  1.1e-05 ***
## Year2009             0.1258     0.0457    2.75  0.00593 **
## Year2010             0.2419     0.0525    4.61  4.1e-06 ***
## Year2011             0.2558     0.0561    4.56  5.2e-06 ***
```

```
## Year2012             0.3252     0.0653    4.98  6.5e-07 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.829
## Multiple R-squared:  0.0114, Adjusted R-squared:  0.00909
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
##  617 weights are ~= 1. The remaining 7083 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.041   0.869   0.949   0.915   0.985   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.30e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                    psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.009  1           1.005
## Year              1.009 16           1.000
```




```
## [1] "List of  45 outliers with residuals above  2.5"
##          ScopusId  NLCS Year OneField Fields residuals
## 627    0000289556 3.577 1996     1700      2     2.647
## 1572   0031103679 4.075 1997     1700      1     3.082
## 1574   0031104254 4.116 1997     1700      1     3.123
## 1612   0031272525 3.611 1997     1700      1     2.618
## 1726   0031190973 3.602 1997     1700      2     2.609
## 1774   0142051871 4.315 1997     1700      1     3.322
## 1985   0032269108 3.702 1998     1700      1     2.676
## 2434   0032058184 3.621 1998     1700      1     2.554
## 2576   0032057865 3.542 1998     1700      2     2.516
## 2595   0032108018 3.662 1998     1700      2     2.636
## 5185   0035424017 3.551 2001     1700      1     2.557
## 6760   0037623983 3.818 2003     1700      2     2.808
## 8647  29144523061 3.760 2006     1700      7     2.658
## 9488  34247481878 4.038 2007     1700      2     2.976
## 9489  34247500374 4.020 2007     1700      2     2.958
## 10007 52149117439 3.641 2008     1700      2     2.519
## 10408 47749140025 3.719 2008     1700      1     2.572
## 10738 37549003336 4.591 2008     1700      1     3.469
## 11104 70350769233 3.653 2009     1700      1     2.597
## 12116 85008044987 4.215 2009     1700      1     3.175
## 13124 73649114265 3.748 2010     1700      1     2.592
## 13527 80053551322 3.818 2011     1700      1     2.648
## 13652 80053974840 3.721 2011     1700      2     2.551
## 13915 79955370378 3.941 2011     1700      1     2.771
## 14045 79951942070 3.898 2011     1700      1     2.728
```

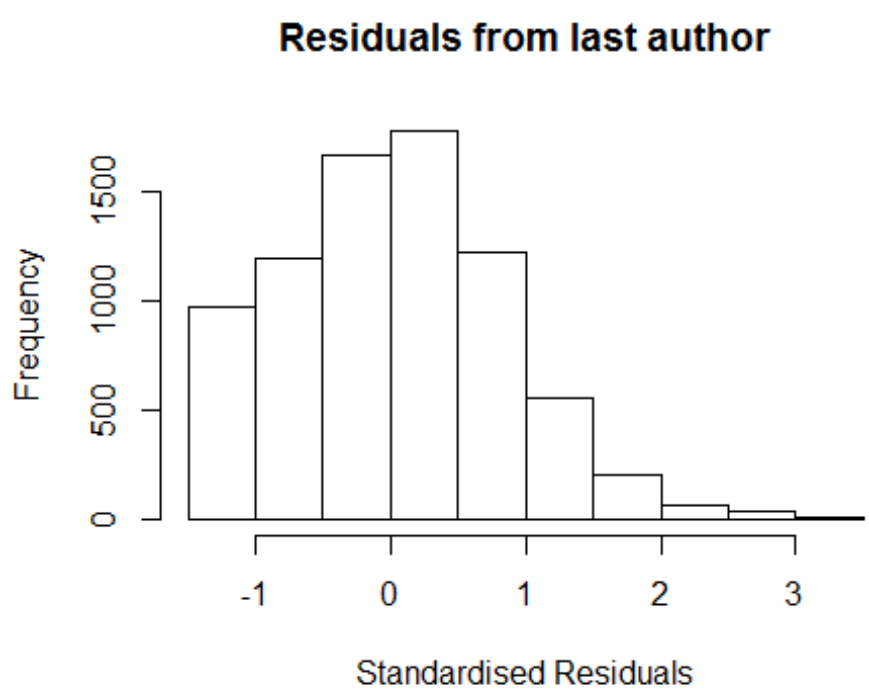
```
## 14348 84872196709 3.890 2012     1700      1     2.651
## 14478 84870549803 3.912 2012     1700      1     2.673
## 14533 84872070696 4.014 2012     1700      1     2.775
## 14597 84865517174 4.039 2012     1700      1     2.775
## 14620 84857713208 3.859 2012     1700      2     2.603
## 14671 84864696820 4.212 2012     1700      1     2.973
## 14734 84863758078 4.014 2012     1700      1     2.775
## 14801 84865483551 4.316 2012     1700      1     3.077
## 14818 84865522374 3.812 2012     1700      1     2.573
## 14851 84863765049 4.149 2012     1700      1     2.910
## 15047 84861902326 4.058 2012     1700      1     2.819
## 15057 84857641127 4.677 2012     1700      1     3.438
## 15074 84857657877 3.882 2012     1700      1     2.643
## 15135 84856912379 3.812 2012     1700      1     2.573
## 15139 84986596759 3.778 2012     1700      1     2.539
## 15177 84856033479 4.051 2012     1700      1     2.812
## 15179 84856041489 4.001 2012     1700      1     2.745
## 15181 84857712777 3.852 2012     1700      1     2.588
## 15182 84859369652 3.752 2012     1700      1     2.513
## 15216 80052530633 4.385 2012     1700      2     3.146
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.27452 -0.56942  0.00205  0.54898  3.46771
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)          0.9150     0.0354   25.83  < 2e-16 ***
## FirstAuthorFemale1   0.0339     0.0244    1.39  0.16424
## Year1997             0.0800     0.0535    1.50  0.13489
## Year1998             0.1116     0.0527    2.12  0.03432 *
## Year1999             0.1126     0.0525    2.14  0.03218 *
## Year2000             0.1810     0.0490    3.69  0.00022 ***
## Year2001             0.0802     0.0500    1.60  0.10901
## Year2002             0.2617     0.0509    5.14  2.8e-07 ***
## Year2003             0.0966     0.0578    1.67  0.09441 .
## Year2004             0.0954     0.0546    1.75  0.08076 .
## Year2005             0.2560     0.0556    4.60  4.3e-06 ***
## Year2006             0.1884     0.0515    3.66  0.00026 ***
## Year2007             0.1486     0.0489    3.04  0.00238 **
## Year2008             0.2083     0.0473    4.40  1.1e-05 ***
## Year2009             0.1258     0.0457    2.75  0.00592 **
## Year2010             0.2422     0.0525    4.62  4.0e-06 ***
## Year2011             0.2559     0.0561    4.56  5.2e-06 ***
## Year2012             0.3256     0.0653    4.99  6.2e-07 ***
```

```
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.829
## Multiple R-squared:  0.0114, Adjusted R-squared:  0.00918
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
##  613 weights are ~= 1. The remaining 7087 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0411  0.8690  0.9490  0.9150  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.30e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##                   GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.008  1           1.004
## Year             1.008 16           1.000

## [1] "List of  45 outliers with residuals above  2.5"
##          ScopusId  NLCS Year OneField Fields residuals
## 627    0000289556 3.577 1996     1700      2     2.647
## 1572   0031103679 4.075 1997     1700      1     3.082
## 1574   0031104254 4.116 1997     1700      1     3.123
## 1612   0031272525 3.611 1997     1700      1     2.618
## 1726   0031190973 3.602 1997     1700      2     2.609
## 1774   0142051871 4.315 1997     1700      1     3.322
## 1985   0032269108 3.702 1998     1700      1     2.676
## 2434   0032058184 3.621 1998     1700      1     2.554
## 2576   0032057865 3.542 1998     1700      2     2.516
## 2595   0032108018 3.662 1998     1700      2     2.636
## 5185   0035424017 3.551 2001     1700      1     2.557
## 6760   0037623983 3.818 2003     1700      2     2.808
## 8647  29144523061 3.760 2006     1700      7     2.658
## 9488  34247481878 4.038 2007     1700      2     2.976
## 9489  34247500374 4.020 2007     1700      2     2.958
## 10007 52149117439 3.641 2008     1700      2     2.519
```

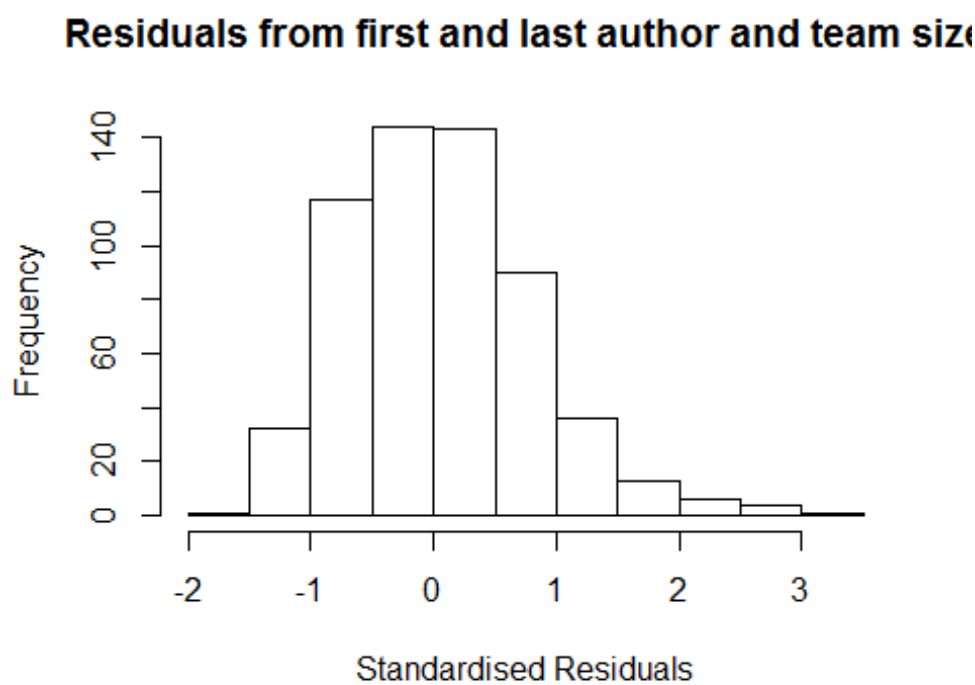
```
## 10408 47749140025 3.719 2008     1700      1     2.572
## 10738 37549003336 4.591 2008     1700      1     3.469
## 11104 70350769233 3.653 2009     1700      1     2.597
## 12116 85008044987 4.215 2009     1700      1     3.175
## 13124 73649114265 3.748 2010     1700      1     2.592
## 13527 80053551322 3.818 2011     1700      1     2.648
## 13652 80053974840 3.721 2011     1700      2     2.551
## 13915 79955370378 3.941 2011     1700      1     2.771
## 14045 79951942070 3.898 2011     1700      1     2.728
## 14348 84872196709 3.890 2012     1700      1     2.651
## 14478 84870549803 3.912 2012     1700      1     2.673
## 14533 84872070696 4.014 2012     1700      1     2.775
## 14597 84865517174 4.039 2012     1700      1     2.775
## 14620 84857713208 3.859 2012     1700      2     2.603
## 14671 84864696820 4.212 2012     1700      1     2.973
## 14734 84863758078 4.014 2012     1700      1     2.775
## 14801 84865483551 4.316 2012     1700      1     3.077
## 14818 84865522374 3.812 2012     1700      1     2.573
## 14851 84863765049 4.149 2012     1700      1     2.910
## 15047 84861902326 4.058 2012     1700      1     2.819
## 15057 84857641127 4.677 2012     1700      1     3.438
## 15074 84857657877 3.882 2012     1700      1     2.643
## 15135 84856912379 3.812 2012     1700      1     2.573
## 15139 84986596759 3.778 2012     1700      1     2.539
## 15177 84856033479 4.051 2012     1700      1     2.812
## 15179 84856041489 4.001 2012     1700      1     2.745
## 15181 84857712777 3.852 2012     1700      1     2.588
## 15182 84859369652 3.752 2012     1700      1     2.513
## 15216 80052530633 4.385 2012     1700      2     3.146
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.27250 -0.56913 -0.00022  0.54900  3.46697
##
## Coefficients:
##                   Estimate Std. Error t value Pr(>|t|)
## (Intercept)         0.9158     0.0353   25.91  < 2e-16 ***
## LastAuthorFemale1   0.0300     0.0239    1.26  0.20917
## Year1997            0.0802     0.0534    1.50  0.13366
## Year1998            0.1124     0.0527    2.13  0.03287 *
## Year1999            0.1122     0.0525    2.14  0.03262 *
## Year2000            0.1814     0.0490    3.70  0.00022 ***
## Year2001            0.0810     0.0499    1.62  0.10496
## Year2002            0.2611     0.0509    5.13  2.9e-07 ***
## Year2003            0.0959     0.0577    1.66  0.09666 .
```

```
## Year2004            0.0962     0.0546    1.76  0.07784 .
## Year2005            0.2550     0.0556    4.59  4.6e-06 ***
## Year2006            0.1883     0.0515    3.66  0.00026 ***
## Year2007            0.1482     0.0489    3.03  0.00246 **
## Year2008            0.2082     0.0473    4.40  1.1e-05 ***
## Year2009            0.1260     0.0457    2.76  0.00588 **
## Year2010            0.2423     0.0525    4.62  3.9e-06 ***
## Year2011            0.2564     0.0561    4.57  4.9e-06 ***
## Year2012            0.3267     0.0653    5.00  5.8e-07 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.831
## Multiple R-squared:  0.0113, Adjusted R-squared:  0.0091
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
##  610 weights are ~= 1. The remaining 7090 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0426  0.8690  0.9490  0.9160  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.30e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Sample size for the above analysis:  7700"
## [1] ""
## [1] ""
## [1] "####################################"
## [1] "Analysis of AJSC 1701"
## [1] "####################################"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    5    1    1    5    7   15   26   32   35   93  165  160  138  111   92
## 2011 2012
##  130  103
##
```

```
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    5    1    1    5    5    6   18   16   19   45  100  104   89   73   60
## 2011 2012
##   82   67
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    5    1    1    5    4    5   16   14   17   34   80   92   74   62   51
## 2011 2012
##   68   58
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
## [1] "Female first author team size 2018 geometric mean: 2.68040357823952"
## [1] "Male first author team size 2018 geometric mean: 2.54278189708881"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 350, p-value = 0.7
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 2.69765144187547"
## [1] "Male last author team size 2018 geometric mean: 2.53212302115941"

## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties
```



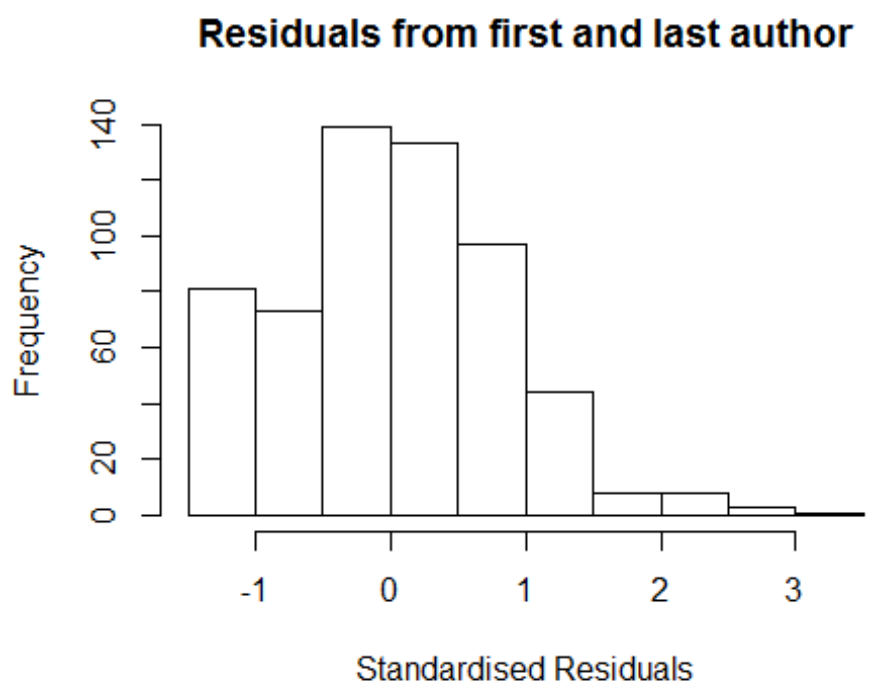
```
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 400, p-value = 0.5
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##                        GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.937e+12  1       1.392e+06
## LastAuthorFemale  1.965e+00  1       1.402e+00
## UniqueAuthors     1.067e+13  4       4.251e+01
## Year              1.675e+13 16       2.590e+00
```




```
## [1] "List of  5 outliers with residuals above  2.5"
##         ScopusId  NLCS Year OneField Fields residuals
## 79    3042559983 3.635 2003     1701      1     2.710
## 660  48449095896 4.549 2008     1701      2     3.187
## 835  68849094099 3.357 2009     1701      2     2.517
## 1049 82055180716 3.935 2011     1701      1     2.554
## 1243 84859170855 3.881 2012     1701      1     2.820
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
UniqueAuthors +
##     Year, data = AllScopusDataOlderFirstLastGendered, control =
```

```
lmrob.control(fast.s.large.n = Inf,
##     k.max = 1000))
##  \--> method = "MM"
## Residuals:
##       Min        1Q    Median        3Q       Max
## -1.61e+00 -5.05e-01 -2.22e-15  5.28e-01  3.19e+00
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)           0.4849     0.2554    1.90  0.05815 .
## FirstAuthorFemale1    0.0693     0.0968    0.72  0.47402
## LastAuthorFemale1    -0.0171     0.0908   -0.19  0.85035
## UniqueAuthors2        0.4751     0.0836    5.68  2.1e-08 ***
## UniqueAuthors3        0.3670     0.0983    3.73  0.00021 ***
## UniqueAuthors4        0.2417     0.1209    2.00  0.04604 *
## UniqueAuthors5        0.3237     0.1408    2.30  0.02187 *
## Year1997              0.6410     0.2842    2.26  0.02447 *
## Year1998              2.3561     0.2554    9.22  < 2e-16 ***
## Year1999              0.2292     0.3683    0.62  0.53405
## Year2000              0.1068     0.4521    0.24  0.81331
## Year2001              0.4265     0.5599    0.76  0.44658
## Year2002              0.3707     0.3805    0.97  0.33033
## Year2003              0.1159     0.3551    0.33  0.74438
## Year2004              0.6840     0.3141    2.18  0.02983 *
## Year2005              0.5019     0.2944    1.70  0.08878 .
## Year2006              0.4450     0.2742    1.62  0.10510
## Year2007              0.3984     0.2620    1.52  0.12894
## Year2008              0.4191     0.2642    1.59  0.11328
## Year2009              0.3029     0.2778    1.09  0.27593
## Year2010              0.3606     0.2736    1.32  0.18813
## Year2011              0.4386     0.2667    1.64  0.10069
## Year2012              0.2823     0.2822    1.00  0.31754
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.751
## Multiple R-squared:  0.0905, Adjusted R-squared:  0.055
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
##  58 weights are ~= 1. The remaining 529 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0322  0.8720  0.9480  0.9030  0.9840  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.70e-04          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
```

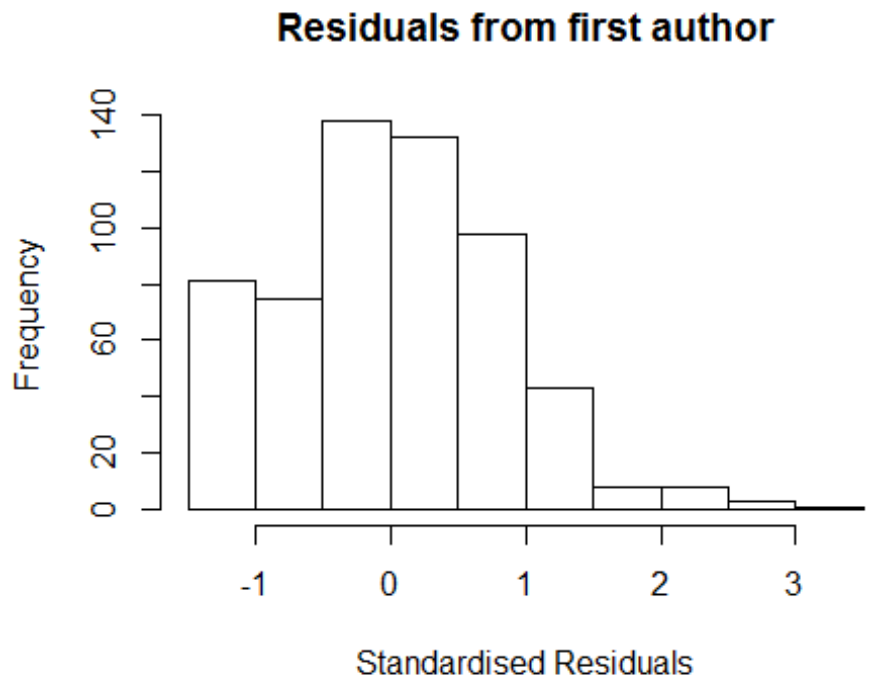
```
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##                        GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 7.011e+12  1       2.648e+06
## LastAuthorFemale  1.719e+00  1       1.311e+00
## Year              9.381e+12 16       2.543e+00
```




```
## [1] "List of  4 outliers with residuals above  2.5"
##         ScopusId  NLCS Year OneField Fields residuals
## 79    3042559983 3.635 2003     1701      1     2.704
## 660  48449095896 4.549 2008     1701      2     3.397
## 1049 82055180716 3.935 2011     1701      1     2.769
## 1243 84859170855 3.881 2012     1701      1     2.801
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
```

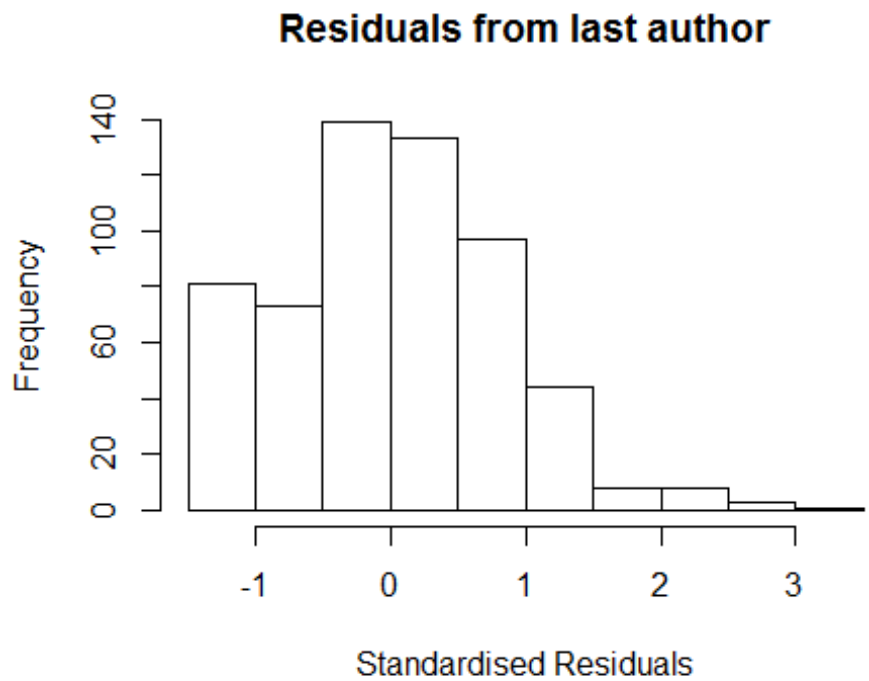
```
##  \--> method = "MM"
## Residuals:
##       Min        1Q    Median        3Q       Max
## -1.405518 -0.548793  0.000876  0.550538  3.396875
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)          0.5815     0.2721    2.14   0.0330 *
## FirstAuthorFemale1   0.0566     0.0991    0.57   0.5680
## LastAuthorFemale1    0.0164     0.0918    0.18   0.8585
## Year1997             0.7989     0.2825    2.83   0.0049 **
## Year1998             2.2595     0.2721    8.30  7.4e-16 ***
## Year1999             0.2009     0.4333    0.46   0.6430
## Year2000             0.2838     0.3932    0.72   0.4707
## Year2001             0.3832     0.5057    0.76   0.4489
## Year2002             0.5774     0.3758    1.54   0.1250
## Year2003             0.3490     0.3917    0.89   0.3732
## Year2004             0.7674     0.3228    2.38   0.0178 *
## Year2005             0.5572     0.3161    1.76   0.0785 .
## Year2006             0.5966     0.2853    2.09   0.0370 *
## Year2007             0.5497     0.2799    1.96   0.0500 .
## Year2008             0.5542     0.2802    1.98   0.0484 *
## Year2009             0.5089     0.2946    1.73   0.0846 .
## Year2010             0.5417     0.2952    1.84   0.0670 .
## Year2011             0.5682     0.2868    1.98   0.0481 *
## Year2012             0.4252     0.2963    1.44   0.1518
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.813
## Multiple R-squared:  0.0236, Adjusted R-squared:  -0.00739
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
##  62 weights are ~= 1. The remaining 525 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.042   0.865   0.948   0.911   0.986   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.70e-04          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
```

```
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##                        GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 2.961e+12  1       1.721e+06
## Year              2.961e+12 16       2.453e+00
```




```
## [1] "List of  4 outliers with residuals above  2.5"
##         ScopusId  NLCS Year OneField Fields residuals
## 79    3042559983 3.635 2003     1701      1     2.704
## 660  48449095896 4.549 2008     1701      2     3.397
## 1049 82055180716 3.935 2011     1701      1     2.769
## 1243 84859170855 3.881 2012     1701      1     2.801
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##       Min        1Q    Median        3Q       Max
## -1.41e+00 -5.48e-01  1.11e-15  5.49e-01  3.41e+00
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)          0.5864     0.2708    2.17   0.0308 *
```

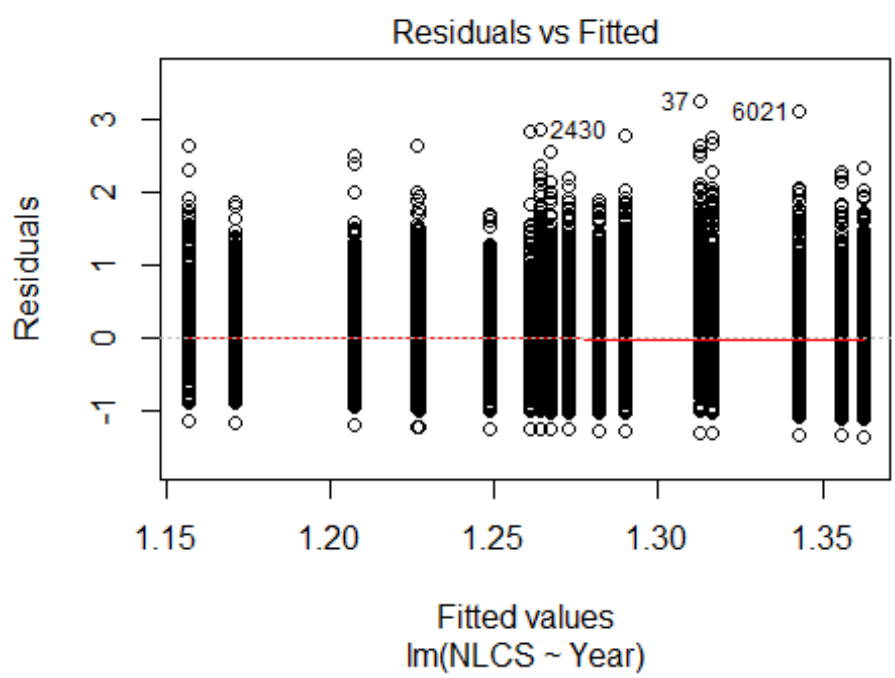
```
## FirstAuthorFemale1   0.0642     0.0873    0.74   0.4624
## Year1997             0.7864     0.2746    2.86   0.0043 **
## Year1998             2.2546     0.2708    8.32  6.3e-16 ***
## Year1999             0.1959     0.4325    0.45   0.6508
## Year2000             0.2832     0.3942    0.72   0.4728
## Year2001             0.3799     0.5063    0.75   0.4534
## Year2002             0.5744     0.3752    1.53   0.1263
## Year2003             0.3426     0.3908    0.88   0.3811
## Year2004             0.7622     0.3226    2.36   0.0185 *
## Year2005             0.5522     0.3148    1.75   0.0800 .
## Year2006             0.5938     0.2847    2.09   0.0374 *
## Year2007             0.5466     0.2794    1.96   0.0509 .
## Year2008             0.5512     0.2793    1.97   0.0489 *
## Year2009             0.5062     0.2937    1.72   0.0853 .
## Year2010             0.5399     0.2945    1.83   0.0673 .
## Year2011             0.5655     0.2856    1.98   0.0482 *
## Year2012             0.4223     0.2959    1.43   0.1542
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.811
## Multiple R-squared:  0.0236, Adjusted R-squared:  -0.00561
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
##  64 weights are ~= 1. The remaining 523 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0378  0.8640  0.9480  0.9100  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.70e-04          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##                   GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.314  1           1.146
## Year             1.314 16           1.009
```



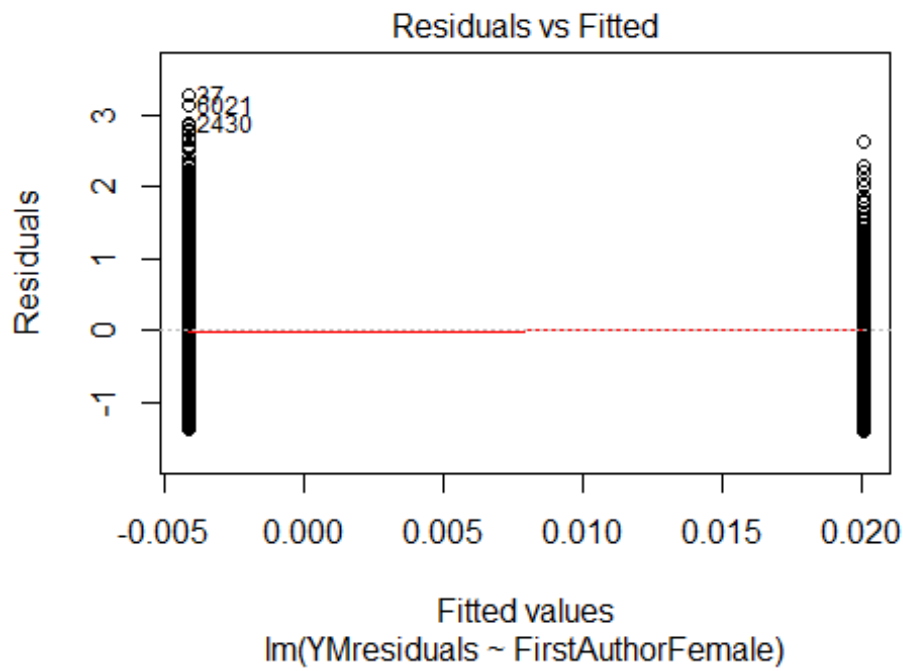

```
## [1] "List of  4 outliers with residuals above  2.5"
##         ScopusId  NLCS Year OneField Fields residuals
## 79    3042559983 3.635 2003     1701      1     2.704
## 660  48449095896 4.549 2008     1701      2     3.397
## 1049 82055180716 3.935 2011     1701      1     2.769
## 1243 84859170855 3.881 2012     1701      1     2.801
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.35736 -0.55299  0.00262  0.55873  3.36502
##
## Coefficients:
##                   Estimate Std. Error t value Pr(>|t|)
## (Intercept)         0.5985     0.2673    2.24   0.0255 *
## LastAuthorFemale1   0.0425     0.0811    0.52   0.6001
## Year1997            0.8385     0.2673    3.14   0.0018 **
## Year1998            2.2425     0.2673    8.39  3.9e-16 ***
## Year1999            0.1844     0.4299    0.43   0.6682
## Year2000            0.2600     0.3863    0.67   0.5012
## Year2001            0.3780     0.5078    0.74   0.4570
## Year2002            0.5574     0.3724    1.50   0.1350
## Year2003            0.3386     0.3857    0.88   0.3804
```

```
## Year2004            0.7589     0.3173    2.39   0.0171 *
## Year2005            0.5484     0.3126    1.75   0.0799 .
## Year2006            0.5873     0.2812    2.09   0.0372 *
## Year2007            0.5399     0.2758    1.96   0.0507 .
## Year2008            0.5430     0.2763    1.97   0.0499 *
## Year2009            0.5012     0.2900    1.73   0.0845 .
## Year2010            0.5317     0.2912    1.83   0.0684 .
## Year2011            0.5555     0.2823    1.97   0.0496 *
## Year2012            0.4122     0.2915    1.41   0.1578
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.817
## Multiple R-squared:  0.0228, Adjusted R-squared:  -0.00637
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  62 weights are ~= 1. The remaining 525 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0519  0.8660  0.9500  0.9120  0.9860  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.70e-04          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Sample size for the above analysis:  587"
## [1] ""
## [1] ""
## [1] "####################################"
## [1] "Analysis of AJSC 1702"
## [1] "####################################"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 1015  998  931  998 1084 1059  995  758  767  873  875  962  885 1050  991
## 2011 2012
##  903  946
##
```

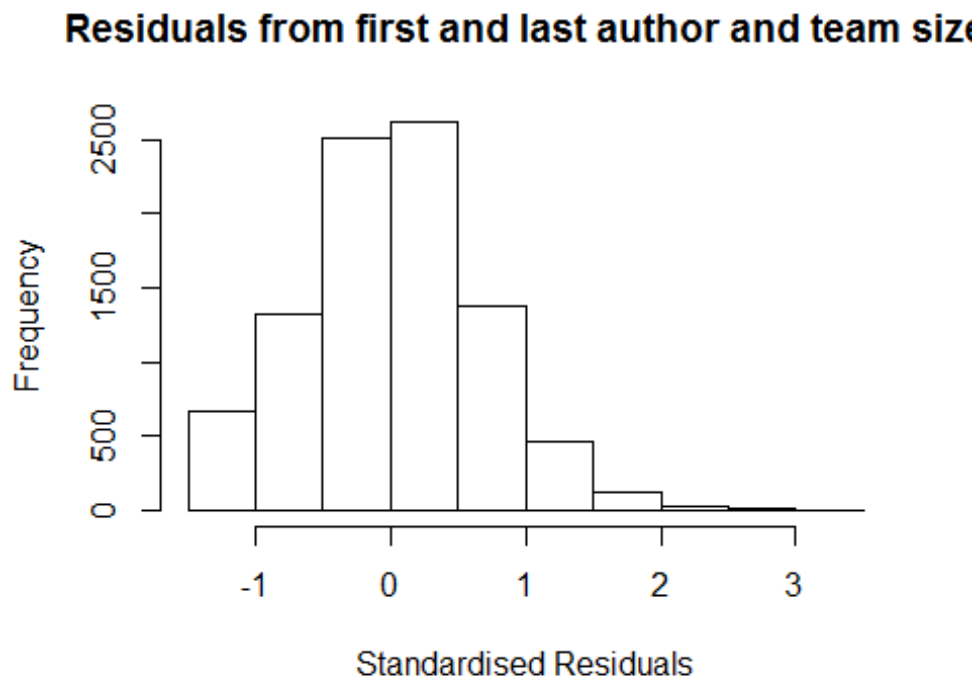
```
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  698  709  653  644  725  671  672  502  527  562  573  612  565  698  673
## 2011 2012
##  593  648
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  585  609  569  545  637  589  573  427  428  461  504  513  467  598  571
## 2011 2012
##  508  544
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
##  Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 140, df = 16, p-value <2e-16
```




```
##
##  Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 34, df = 1, p-value = 5e-09
```



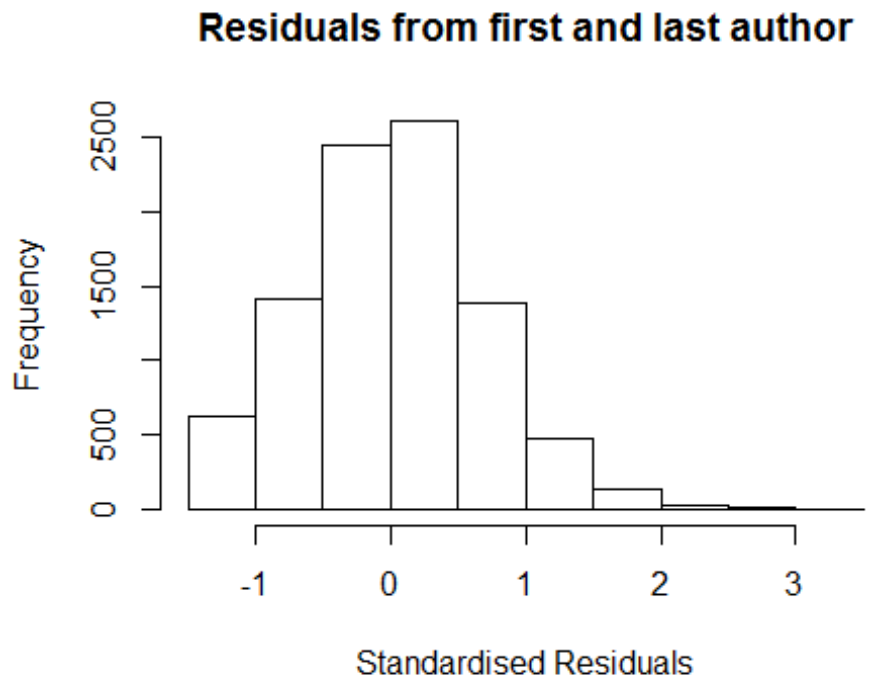

```
## [1] "Female first author team size 2018 geometric mean: 2.37047542361374"
## [1] "Male first author team size 2018 geometric mean: 2.29689322498408"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 900, p-value = 0.9
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 2.24960541678009"
## [1] "Male last author team size 2018 geometric mean: 2.3271613081618"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 750, p-value = 0.9
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##                     GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.263  1           1.124
## LastAuthorFemale  1.259  1           1.122
## UniqueAuthors     1.046  4           1.006
## Year              1.053 16           1.002
```




```
## [1] "List of  11 outliers with residuals above  2.5"
##          ScopusId  NLCS Year OneField Fields residuals
## 24     0029679044 3.953 1996     1702      2     2.534
## 37     0030211964 4.574 1996     1702      2     3.428
## 93     0030142764 3.833 1996     1702      4     2.687
## 1345   0031381525 3.993 1997     1203      3     2.654
## 1408   0031185845 4.068 1997     1702      5     2.682
## 2430   0032204063 4.140 1998     1702      4     2.800
## 4461   4243148480 4.090 1999     1702      5     2.972
## 6021   0035478854 4.469 2001     1702      2     3.316
## 8582   0141607824 4.104 2003     1702      4     2.786
## 11493 33646023117 3.828 2006     1702      4     2.728
## 17406 79953048649 3.731 2011     1702      5     2.500
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
UniqueAuthors +
##     Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##     k.max = 1000))
##  \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.46339 -0.43936  0.00988  0.44087  3.42752
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
```

```
## (Intercept)         1.14648    0.03465   33.08  < 2e-16 ***
## FirstAuthorFemale1  0.02746    0.01974    1.39  0.16422
## LastAuthorFemale1   0.04375    0.01996    2.19  0.02840 *
## UniqueAuthors2      0.19703    0.01719   11.46  < 2e-16 ***
## UniqueAuthors3      0.24462    0.02098   11.66  < 2e-16 ***
## UniqueAuthors4      0.23450    0.02750    8.53  < 2e-16 ***
## UniqueAuthors5      0.25818    0.03256    7.93  2.4e-15 ***
## Year1997           -0.00486    0.04542   -0.11  0.91483
## Year1998           -0.05100    0.04545   -1.12  0.26182
## Year1999           -0.02809    0.04840   -0.58  0.56162
## Year2000           -0.03516    0.04343   -0.81  0.41817
## Year2001            0.00636    0.04407    0.14  0.88523
## Year2002            0.03720    0.04388    0.85  0.39657
## Year2003           -0.07306    0.04616   -1.58  0.11355
## Year2004            0.01120    0.04746    0.24  0.81341
## Year2005           -0.05881    0.04681   -1.26  0.20904
## Year2006           -0.04692    0.04359   -1.08  0.28177
## Year2007           -0.09735    0.04362   -2.23  0.02563 *
## Year2008           -0.06934    0.04269   -1.62  0.10430
## Year2009           -0.11472    0.04210   -2.72  0.00644 **
## Year2010           -0.17036    0.04172   -4.08  4.5e-05 ***
## Year2011           -0.15001    0.04335   -3.46  0.00054 ***
## Year2012           -0.17123    0.04306   -3.98  7.0e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.653
## Multiple R-squared:  0.0338, Adjusted R-squared:  0.0315
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  2 observations c(24,3069) are outliers with |weight| = 0 ( < 1.1e-05);
##  754 weights are ~= 1. The remaining 8372 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0031  0.8690  0.9510  0.9050  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.10e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
```

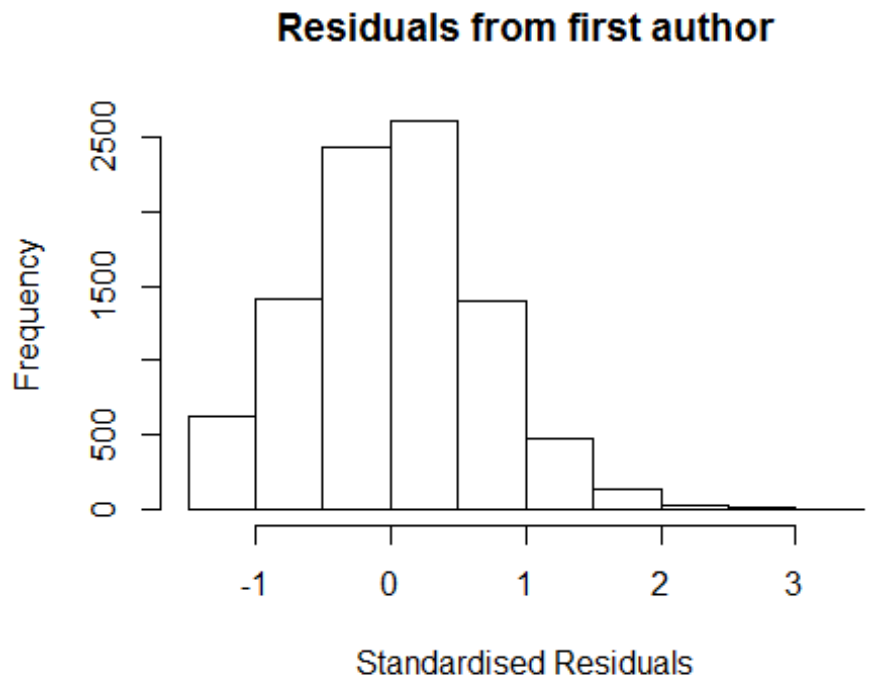
```
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##                     GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.252  1           1.119
## LastAuthorFemale  1.249  1           1.118
## Year              1.012 16           1.000
```




```
## [1] "List of  11 outliers with residuals above  2.5"
##          ScopusId  NLCS Year OneField Fields residuals
## 24     0029679044 3.953 1996     1702      2     2.646
## 37     0030211964 4.574 1996     1702      2     3.300
## 93     0030142764 3.833 1996     1702      4     2.559
## 1345   0031381525 3.993 1997     1203      3     2.713
## 1408   0031185845 4.068 1997     1702      5     2.788
## 2430   0032204063 4.140 1998     1702      4     2.914
## 4461   4243148480 4.090 1999     1702      5     2.843
## 6021   0035478854 4.469 2001     1702      2     3.173
## 8582   0141607824 4.104 2003     1702      4     2.883
## 11493 33646023117 3.828 2006     1702      4     2.574
## 17406 79953048649 3.731 2011     1702      5     2.584
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
```

```
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.4096 -0.4377  0.0147  0.4429  3.2996
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.27436    0.03323   38.35  < 2e-16 ***
## FirstAuthorFemale1  0.03279    0.01973    1.66  0.09654 .
## LastAuthorFemale1   0.03959    0.01996    1.98  0.04736 *
## Year1997            0.00577    0.04516    0.13  0.89827
## Year1998           -0.04812    0.04532   -1.06  0.28843
## Year1999           -0.02727    0.04830   -0.56  0.57239
## Year2000           -0.02244    0.04325   -0.52  0.60393
## Year2001            0.02138    0.04407    0.49  0.62760
## Year2002            0.06285    0.04351    1.44  0.14867
## Year2003           -0.05344    0.04658   -1.15  0.25128
## Year2004            0.03303    0.04775    0.69  0.48909
## Year2005           -0.03921    0.04713   -0.83  0.40549
## Year2006           -0.02075    0.04350   -0.48  0.63344
## Year2007           -0.07543    0.04340   -1.74  0.08227 .
## Year2008           -0.03961    0.04266   -0.93  0.35319
## Year2009           -0.08538    0.04199   -2.03  0.04207 *
## Year2010           -0.16086    0.04194   -3.84  0.00013 ***
## Year2011           -0.12763    0.04336   -2.94  0.00326 **
## Year2012           -0.13789    0.04293   -3.21  0.00132 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.658
## Multiple R-squared:  0.00902,    Adjusted R-squared:  0.00706
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  2 observations c(24,3069) are outliers with |weight| = 0 ( < 1.1e-05);
##  808 weights are ~= 1. The remaining 8318 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0113  0.8660  0.9500  0.9040  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.10e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
```

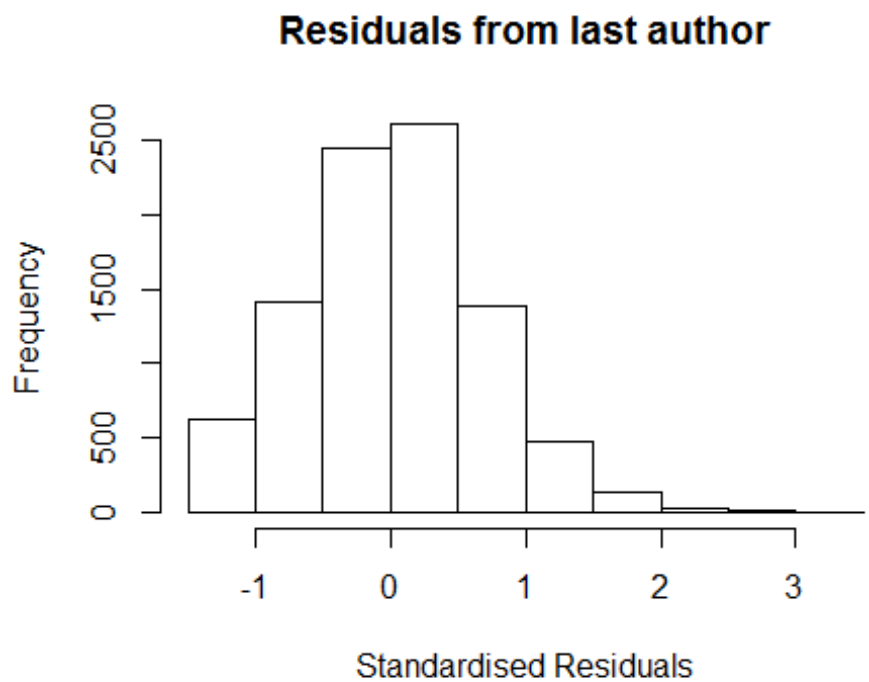
```
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.008  1           1.004
## Year              1.008 16           1.000
```




```
## [1] "List of  11 outliers with residuals above  2.5"
##          ScopusId  NLCS Year OneField Fields residuals
## 24     0029679044 3.953 1996     1702      2     2.646
## 37     0030211964 4.574 1996     1702      2     3.300
## 93     0030142764 3.833 1996     1702      4     2.559
## 1345   0031381525 3.993 1997     1203      3     2.713
## 1408   0031185845 4.068 1997     1702      5     2.788
## 2430   0032204063 4.140 1998     1702      4     2.914
## 4461   4243148480 4.090 1999     1702      5     2.843
## 6021   0035478854 4.469 2001     1702      2     3.173
## 8582   0141607824 4.104 2003     1702      4     2.883
## 11493 33646023117 3.828 2006     1702      4     2.574
## 17406 79953048649 3.731 2011     1702      5     2.584
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
```

```
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.3905 -0.4376  0.0151  0.4426  3.2972
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.27681    0.03314   38.53  < 2e-16 ***
## FirstAuthorFemale1  0.04986    0.01771    2.82  0.00488 **
## Year1997            0.00661    0.04513    0.15  0.88355
## Year1998           -0.04689    0.04523   -1.04  0.29988
## Year1999           -0.02658    0.04826   -0.55  0.58187
## Year2000           -0.02034    0.04322   -0.47  0.63796
## Year2001            0.02276    0.04403    0.52  0.60527
## Year2002            0.06380    0.04347    1.47  0.14222
## Year2003           -0.05247    0.04658   -1.13  0.26005
## Year2004            0.03363    0.04774    0.70  0.48120
## Year2005           -0.03824    0.04709   -0.81  0.41680
## Year2006           -0.01917    0.04346   -0.44  0.65918
## Year2007           -0.07383    0.04335   -1.70  0.08853 .
## Year2008           -0.03786    0.04261   -0.89  0.37431
## Year2009           -0.08367    0.04193   -2.00  0.04602 *
## Year2010           -0.15883    0.04191   -3.79  0.00015 ***
## Year2011           -0.12578    0.04333   -2.90  0.00370 **
## Year2012           -0.13592    0.04288   -3.17  0.00153 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.659
## Multiple R-squared:  0.00863,    Adjusted R-squared:  0.00678
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  2 observations c(24,3069) are outliers with |weight| = 0 ( < 1.1e-05);
##  811 weights are ~= 1. The remaining 8315 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0123  0.8670  0.9500  0.9040  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.10e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
```

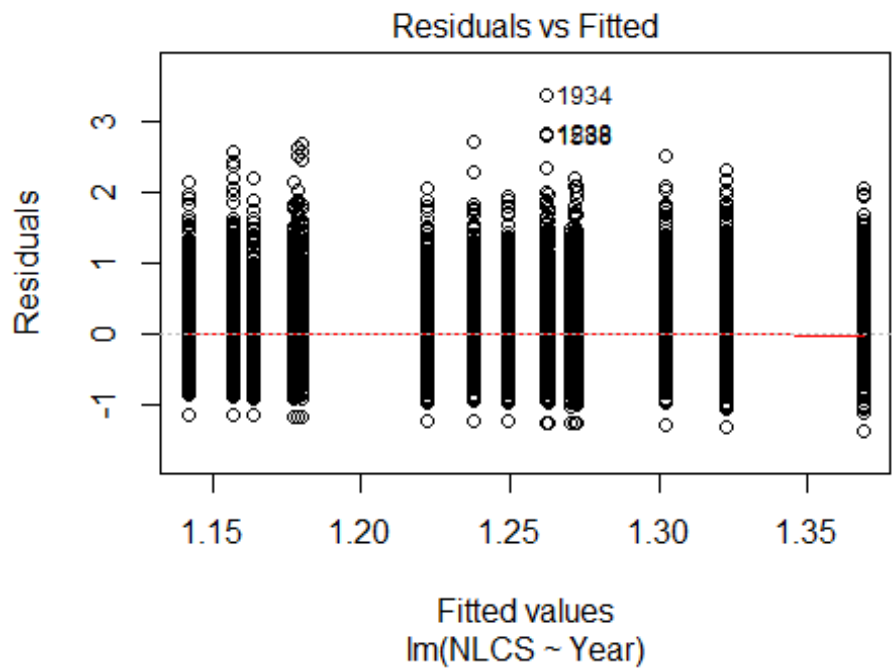
```
##                    "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.006  1           1.003
## Year             1.006 16           1.000
```



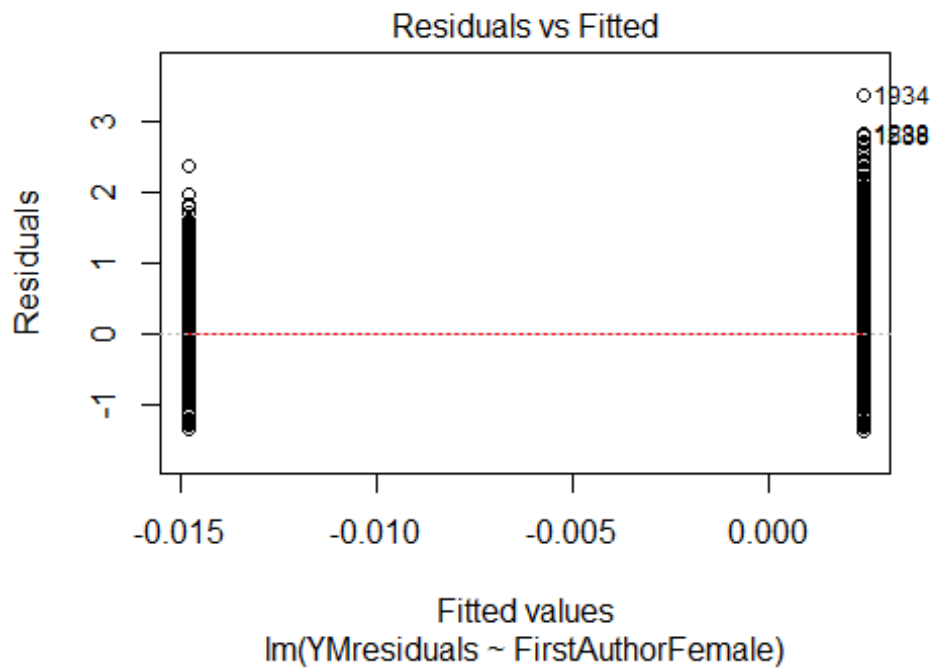

```
## [1] "List of  11 outliers with residuals above  2.5"
##          ScopusId  NLCS Year OneField Fields residuals
## 24     0029679044 3.953 1996     1702      2     2.646
## 37     0030211964 4.574 1996     1702      2     3.300
## 93     0030142764 3.833 1996     1702      4     2.559
## 1345   0031381525 3.993 1997     1203      3     2.713
## 1408   0031185845 4.068 1997     1702      5     2.788
## 2430   0032204063 4.140 1998     1702      4     2.914
## 4461   4243148480 4.090 1999     1702      5     2.843
## 6021   0035478854 4.469 2001     1702      2     3.173
## 8582   0141607824 4.104 2003     1702      4     2.883
## 11493 33646023117 3.828 2006     1702      4     2.574
## 17406 79953048649 3.731 2011     1702      5     2.584
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
```

```
##     Min      1Q  Median      3Q     Max
## -1.3959 -0.4377  0.0133  0.4452  3.2971
##
## Coefficients:
##                   Estimate Std. Error t value Pr(>|t|)
## (Intercept)        1.27687    0.03320   38.45  < 2e-16 ***
## LastAuthorFemale1  0.05427    0.01792    3.03  0.00246 **
## Year1997           0.00584    0.04520    0.13  0.89719
## Year1998          -0.04844    0.04535   -1.07  0.28547
## Year1999          -0.02692    0.04828   -0.56  0.57710
## Year2000          -0.02184    0.04326   -0.50  0.61367
## Year2001           0.02204    0.04410    0.50  0.61728
## Year2002           0.06475    0.04353    1.49  0.13693
## Year2003          -0.05283    0.04659   -1.13  0.25680
## Year2004           0.03439    0.04782    0.72  0.47204
## Year2005          -0.03709    0.04717   -0.79  0.43173
## Year2006          -0.01935    0.04355   -0.44  0.65674
## Year2007          -0.07401    0.04343   -1.70  0.08837 .
## Year2008          -0.03912    0.04268   -0.92  0.35941
## Year2009          -0.08370    0.04201   -1.99  0.04636 *
## Year2010          -0.16009    0.04195   -3.82  0.00014 ***
## Year2011          -0.12605    0.04339   -2.91  0.00368 **
## Year2012          -0.13636    0.04291   -3.18  0.00149 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.658
## Multiple R-squared:  0.00873,    Adjusted R-squared:  0.00688
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  2 observations c(24,3069) are outliers with |weight| = 0 ( < 1.1e-05);
##  796 weights are ~= 1. The remaining 8330 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0119  0.8670  0.9500  0.9040  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.10e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
```

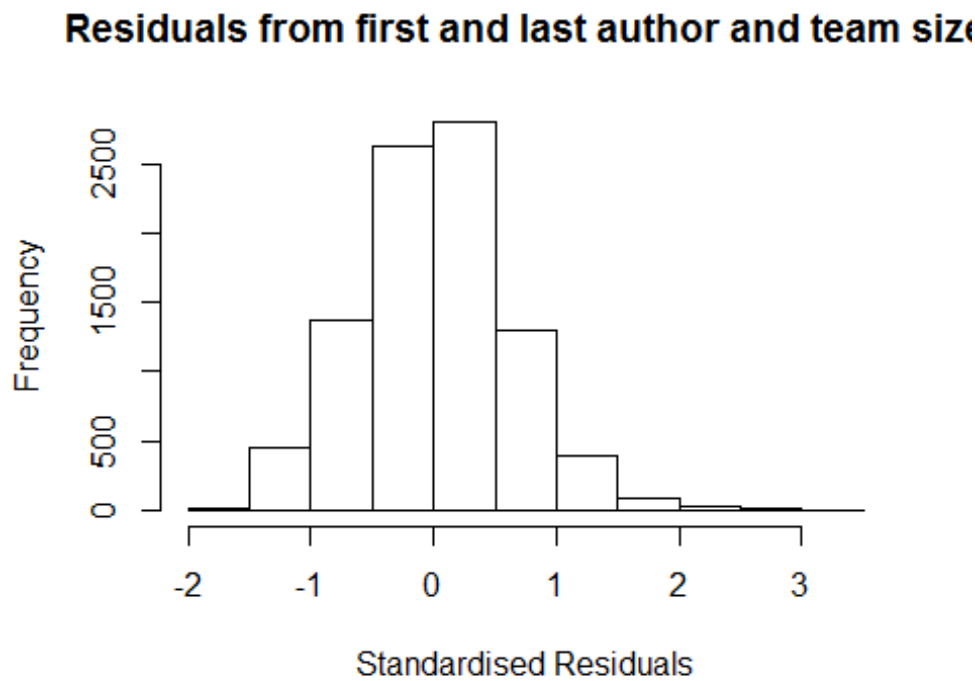
```
## seed : int(0)
## [1] "Sample size for the above analysis:  9128"
## [1] ""
## [1] ""
## [1] "####################################"
## [1] "Analysis of AJSC 1703"
## [1] "####################################"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 1043  966  910  884  890  981  994  981 1048 1126 1240 1307 1298 1255 1210
## 2011 2012
## 1205 1314
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  582  574  476  507  499  518  619  565  613  632  725  784  784  761  736
## 2011 2012
##  761  816
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  484  483  407  416  419  433  509  460  506  522  607  631  660  643  598
## 2011 2012
##  622  682
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
##  Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 140, df = 16, p-value <2e-16
```




```
##
##  Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 22, df = 1, p-value = 3e-06
```



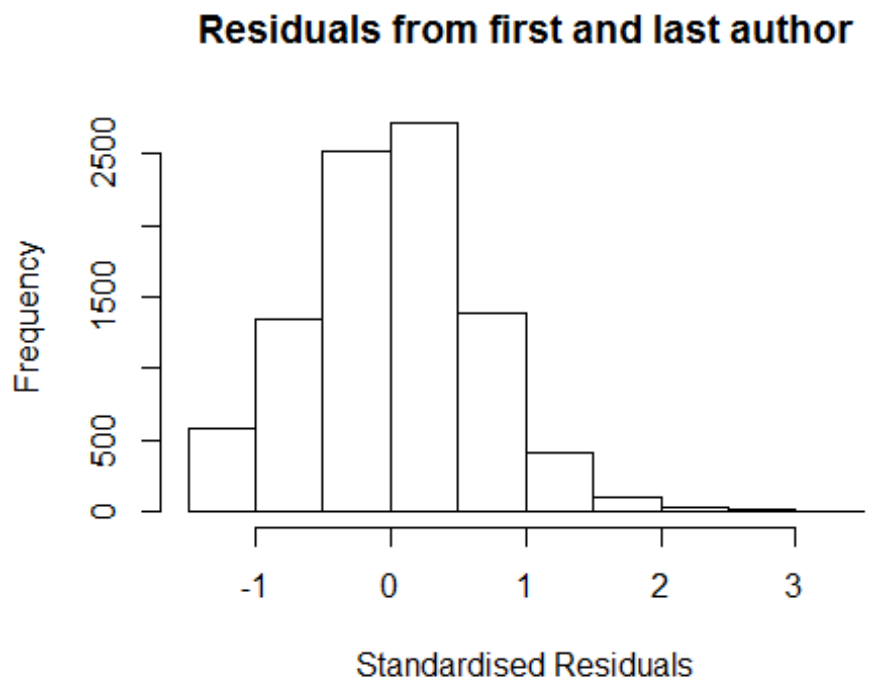

```
## [1] "Female first author team size 2018 geometric mean: 2.96802164332637"
## [1] "Male first author team size 2018 geometric mean: 2.50493551152547"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 19000, p-value = 0.03
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 3.0457980597204"
## [1] "Male last author team size 2018 geometric mean: 2.4865200245729"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 20000, p-value = 0.007
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.102  1           1.050
## LastAuthorFemale  1.102  1           1.050
## UniqueAuthors     1.092  4           1.011
## Year              1.096 16           1.003
```




```
## [1] "List of  7 outliers with residuals above  2.5"
##          ScopusId  NLCS Year OneField Fields residuals
## 120    0030142764 3.833 1996     1702      4     2.717
## 1288   0031118203 4.089 1997     1703      3     2.814
## 1538   0031185845 4.068 1997     1702      5     2.700
## 1934   0031211090 4.640 1997     1703      4     3.365
## 3358   0031773680 3.959 1998     1303      6     2.699
## 16329 71049116435 3.639 2009     1703      4     2.506
## 17126 65449136284 3.759 2009     1303      7     2.534
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
UniqueAuthors +
##     Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##     k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.6208 -0.3963  0.0138  0.3964  3.3645
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)           1.1160     0.0369   30.24  < 2e-16 ***
## FirstAuthorFemale1   -0.0338     0.0182   -1.86  0.06359 .
## LastAuthorFemale1    -0.0216     0.0193   -1.12  0.26393
## UniqueAuthors2        0.2453     0.0175   14.00  < 2e-16 ***
```

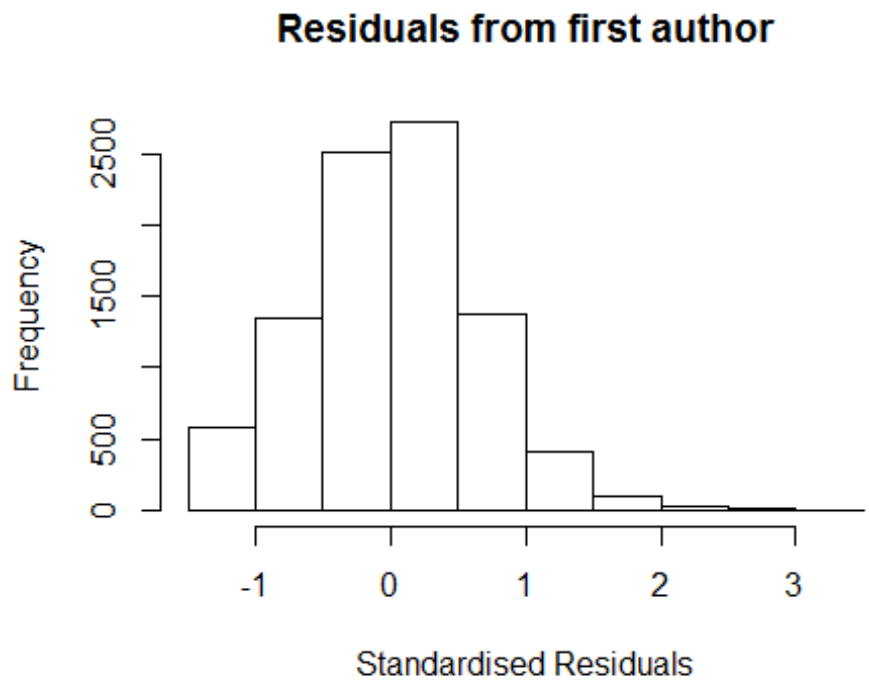
```
## UniqueAuthors3       0.3374     0.0195   17.28  < 2e-16 ***
## UniqueAuthors4       0.4067     0.0231   17.60  < 2e-16 ***
## UniqueAuthors5       0.5047     0.0225   22.43  < 2e-16 ***
## Year1997            -0.0859     0.0496   -1.73  0.08370 .
## Year1998            -0.1014     0.0506   -2.00  0.04507 *
## Year1999            -0.0710     0.0503   -1.41  0.15863
## Year2000            -0.0263     0.0467   -0.56  0.57234
## Year2001             0.0170     0.0485    0.35  0.72618
## Year2002            -0.0672     0.0450   -1.49  0.13518
## Year2003            -0.0808     0.0453   -1.78  0.07430 .
## Year2004            -0.1605     0.0456   -3.52  0.00044 ***
## Year2005            -0.1076     0.0450   -2.39  0.01677 *
## Year2006            -0.0855     0.0442   -1.93  0.05342 .
## Year2007            -0.2136     0.0422   -5.06  4.4e-07 ***
## Year2008            -0.2065     0.0432   -4.78  1.8e-06 ***
## Year2009            -0.2281     0.0420   -5.43  5.9e-08 ***
## Year2010            -0.2207     0.0430   -5.14  2.8e-07 ***
## Year2011            -0.2724     0.0437   -6.24  4.7e-10 ***
## Year2012            -0.2620     0.0423   -6.20  5.9e-10 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.596
## Multiple R-squared:  0.0805, Adjusted R-squared:  0.0783
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  2 observations c(519,817) are outliers with |weight| = 0 ( < 1.1e-05);
##  766 weights are ~= 1. The remaining 8314 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0028  0.8650  0.9510  0.9020  0.9860  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.10e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                    psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##                     GVIF Df GVIF^(1/(2*Df))
```

```
## FirstAuthorFemale 1.082  1           1.040
## LastAuthorFemale  1.081  1           1.039
## Year              1.014 16           1.000
```



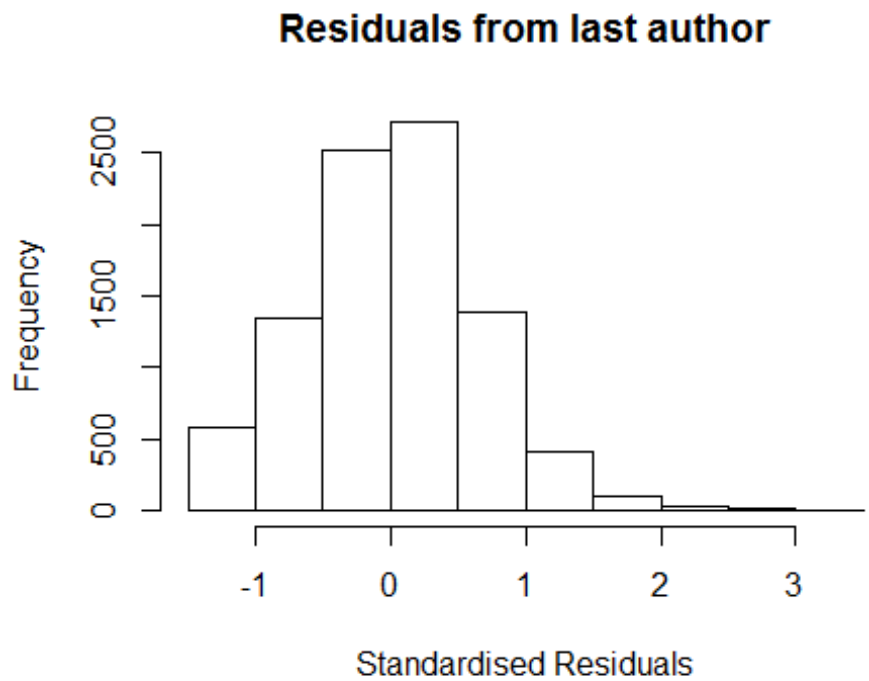

```
## [1] "List of  7 outliers with residuals above  2.5"
##          ScopusId  NLCS Year OneField Fields residuals
## 120    0030142764 3.833 1996     1702      4     2.560
## 1288   0031118203 4.089 1997     1703      3     2.894
## 1538   0031185845 4.068 1997     1702      5     2.873
## 1934   0031211090 4.640 1997     1703      4     3.445
## 3358   0031773680 3.959 1998     1303      6     2.756
## 17126 65449136284 3.759 2009     1303      7     2.610
## 20479 79953048649 3.731 2011     1702      5     2.622
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.3467 -0.4116  0.0161  0.4220  3.4447
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)          1.27335    0.03684   34.56  < 2e-16 ***
## FirstAuthorFemale1 -0.00158    0.01881   -0.08  0.93326
```

```
## LastAuthorFemale1  -0.00966    0.02011   -0.48  0.63104
## Year1997           -0.07809    0.05084   -1.54  0.12458
## Year1998           -0.07065    0.05199   -1.36  0.17421
## Year1999           -0.05989    0.05149   -1.16  0.24481
## Year2000            0.00599    0.04788    0.13  0.90040
## Year2001            0.07337    0.04968    1.48  0.13980
## Year2002           -0.00328    0.04579   -0.07  0.94295
## Year2003           -0.01132    0.04599   -0.25  0.80560
## Year2004           -0.08653    0.04689   -1.85  0.06501 .
## Year2005           -0.02409    0.04590   -0.52  0.59978
## Year2006           -0.00379    0.04527   -0.08  0.93327
## Year2007           -0.14037    0.04311   -3.26  0.00113 **
## Year2008           -0.11743    0.04474   -2.62  0.00869 **
## Year2009           -0.12428    0.04315   -2.88  0.00398 **
## Year2010           -0.12560    0.04416   -2.84  0.00447 **
## Year2011           -0.16469    0.04451   -3.70  0.00022 ***
## Year2012           -0.15688    0.04336   -3.62  0.00030 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.624
## Multiple R-squared:  0.0112, Adjusted R-squared:  0.00923
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
##  observation 817 is an outlier with |weight| = 0 ( < 1.1e-05);
##  738 weights are ~= 1. The remaining 8343 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0004  0.8700  0.9510  0.9050  0.9860  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.10e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.007  1           1.004
## Year              1.007 16           1.000
```




```
## [1] "List of  7 outliers with residuals above  2.5"
##          ScopusId  NLCS Year OneField Fields residuals
## 120    0030142764 3.833 1996     1702      4     2.560
## 1288   0031118203 4.089 1997     1703      3     2.894
## 1538   0031185845 4.068 1997     1702      5     2.873
## 1934   0031211090 4.640 1997     1703      4     3.445
## 3358   0031773680 3.959 1998     1303      6     2.756
## 17126 65449136284 3.759 2009     1303      7     2.610
## 20479 79953048649 3.731 2011     1702      5     2.622
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.3462 -0.4116  0.0163  0.4226  3.4455
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.27250    0.03675   34.63  < 2e-16 ***
## FirstAuthorFemale1 -0.00427    0.01819   -0.23  0.81439
## Year1997           -0.07795    0.05085   -1.53  0.12533
## Year1998           -0.07052    0.05199   -1.36  0.17502
## Year1999           -0.05981    0.05151   -1.16  0.24561
## Year2000            0.00604    0.04789    0.13  0.89967
```

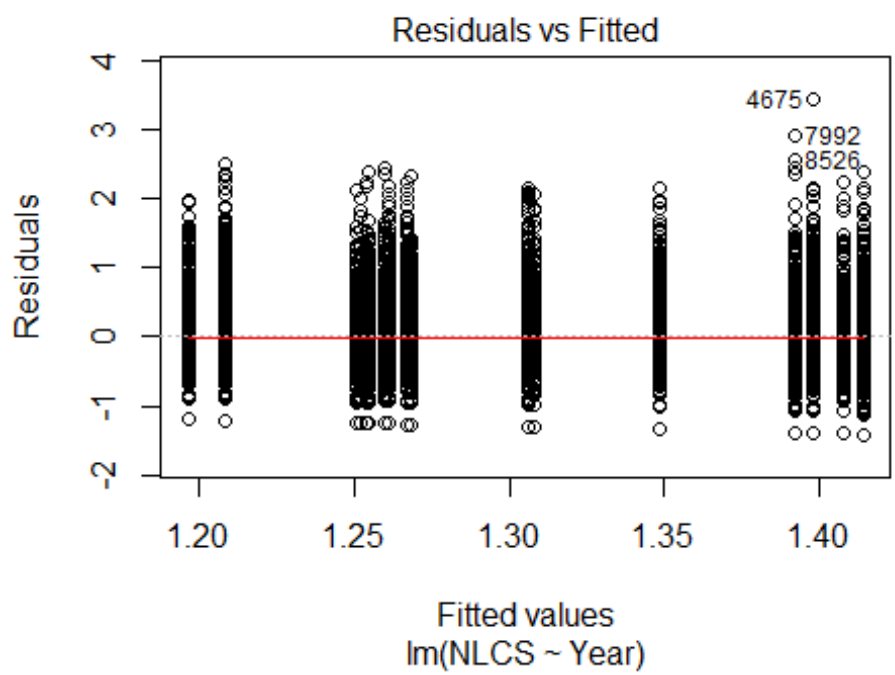
```
## Year2001            0.07365    0.04968    1.48  0.13826
## Year2002           -0.00320    0.04579   -0.07  0.94435
## Year2003           -0.01089    0.04597   -0.24  0.81282
## Year2004           -0.08641    0.04689   -1.84  0.06538 .
## Year2005           -0.02415    0.04592   -0.53  0.59893
## Year2006           -0.00380    0.04527   -0.08  0.93305
## Year2007           -0.14039    0.04312   -3.26  0.00113 **
## Year2008           -0.11743    0.04474   -2.62  0.00869 **
## Year2009           -0.12447    0.04315   -2.88  0.00393 **
## Year2010           -0.12565    0.04417   -2.84  0.00445 **
## Year2011           -0.16469    0.04451   -3.70  0.00022 ***
## Year2012           -0.15698    0.04336   -3.62  0.00030 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.624
## Multiple R-squared:  0.0112, Adjusted R-squared:  0.00932
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  observation 817 is an outlier with |weight| = 0 ( < 1.1e-05);
##  738 weights are ~= 1. The remaining 8343 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0004  0.8700  0.9510  0.9050  0.9860  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.10e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##                   GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.007  1           1.003
## Year             1.007 16           1.000
```



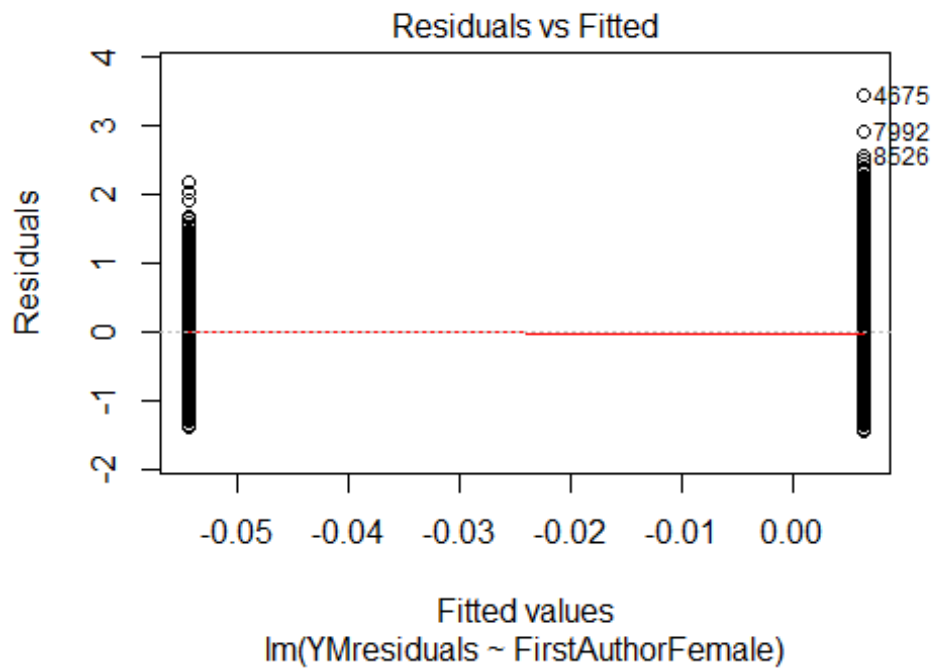

```
## [1] "List of  7 outliers with residuals above  2.5"
##          ScopusId  NLCS Year OneField Fields residuals
## 120    0030142764 3.833 1996     1702      4     2.560
## 1288   0031118203 4.089 1997     1703      3     2.894
## 1538   0031185845 4.068 1997     1702      5     2.873
## 1934   0031211090 4.640 1997     1703      4     3.445
## 3358   0031773680 3.959 1998     1303      6     2.756
## 17126 65449136284 3.759 2009     1303      7     2.610
## 20479 79953048649 3.731 2011     1702      5     2.622
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.3466 -0.4115  0.0164  0.4221  3.4449
##
## Coefficients:
##                   Estimate Std. Error t value Pr(>|t|)
## (Intercept)        1.27321    0.03680   34.60  < 2e-16 ***
## LastAuthorFemale1 -0.01015    0.01942   -0.52  0.60114
## Year1997          -0.07810    0.05085   -1.54  0.12458
## Year1998          -0.07067    0.05199   -1.36  0.17413
## Year1999          -0.05987    0.05149   -1.16  0.24498
## Year2000           0.00600    0.04788    0.13  0.90033
```

```
## Year2001           0.07340    0.04968    1.48  0.13962
## Year2002          -0.00329    0.04579   -0.07  0.94276
## Year2003          -0.01135    0.04599   -0.25  0.80503
## Year2004          -0.08656    0.04689   -1.85  0.06492 .
## Year2005          -0.02410    0.04591   -0.53  0.59954
## Year2006          -0.00384    0.04525   -0.08  0.93231
## Year2007          -0.14042    0.04310   -3.26  0.00113 **
## Year2008          -0.11747    0.04473   -2.63  0.00865 **
## Year2009          -0.12434    0.04314   -2.88  0.00396 **
## Year2010          -0.12563    0.04416   -2.84  0.00445 **
## Year2011          -0.16476    0.04449   -3.70  0.00021 ***
## Year2012          -0.15692    0.04335   -3.62  0.00030 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.624
## Multiple R-squared:  0.0112, Adjusted R-squared:  0.00934
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
##  observation 817 is an outlier with |weight| = 0 ( < 1.1e-05);
##  740 weights are ~= 1. The remaining 8341 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0004  0.8700  0.9510  0.9050  0.9860  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.10e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Sample size for the above analysis:  9082"
## [1] ""
## [1] ""
## [1] "####################################"
## [1] "Analysis of AJSC 1704"
## [1] "####################################"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
```

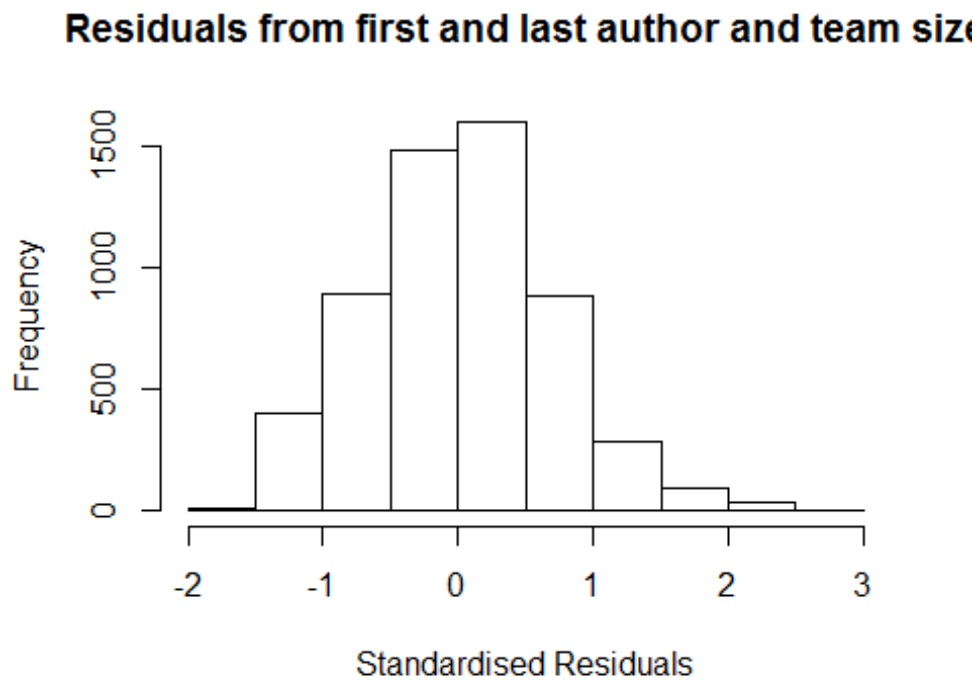
```
##  815  745  788  749  718  783  597  534  570  678  847  781  707  662  619
## 2011 2012
##  682  627
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  486  460  449  432  454  429  362  305  336  405  476  476  418  355  349
## 2011 2012
##  372  363
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  403  401  366  357  397  355  298  240  269  332  377  397  322  272  287
## 2011 2012
##  302  298
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
##  Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 130, df = 16, p-value <2e-16
```




```
##
##  Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 0.63, df = 1, p-value = 0.4
```

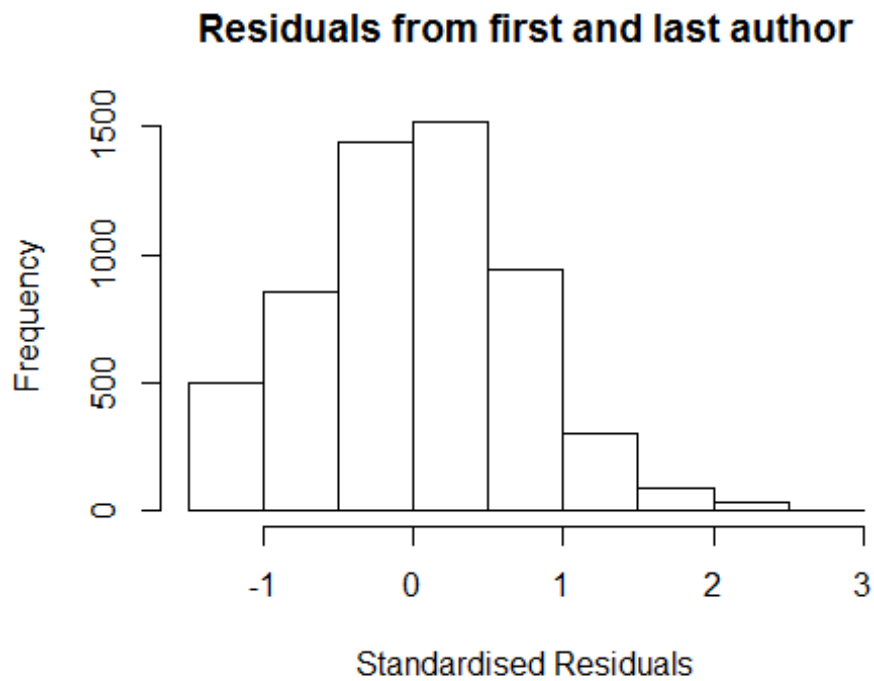



```
## [1] "Female first author team size 2018 geometric mean: 2.63969719926521"
## [1] "Male first author team size 2018 geometric mean: 2.7203646776569"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 3500, p-value = 0.8
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 2.0511508006562"
## [1] "Male last author team size 2018 geometric mean: 2.81086749971627"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 1900, p-value = 0.01
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.117  1           1.057
## LastAuthorFemale  1.123  1           1.060
## UniqueAuthors     1.099  4           1.012
## Year              1.116 16           1.003
```



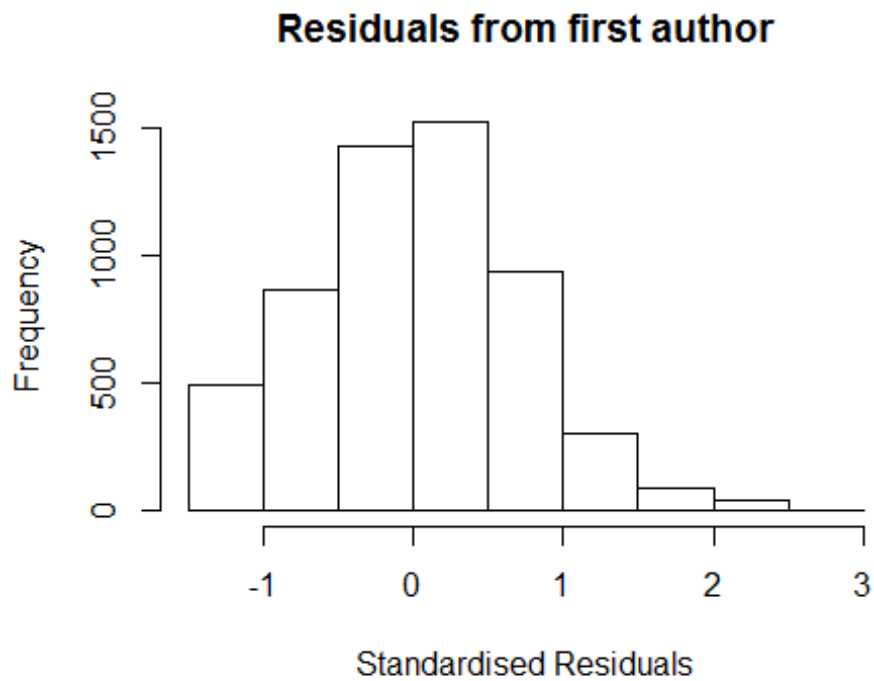

```
## [1] "List of  2 outliers with residuals above  2.5"
##         ScopusId  NLCS Year OneField Fields residuals
## 6596  3843102623 3.655 2004     1704      3     2.539
## 8526 31144478351 3.964 2006     1704      2     2.570
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
UniqueAuthors +
##     Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##     k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.5574 -0.4517  0.0156  0.4587  2.5700
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.02272    0.04203   24.33   <2e-16 ***
## FirstAuthorFemale1 -0.05157    0.03222   -1.60   0.1095
## LastAuthorFemale1  -0.05123    0.03164   -1.62   0.1054
## UniqueAuthors2      0.33573    0.02582   13.00   <2e-16 ***
## UniqueAuthors3      0.43959    0.02800   15.70   <2e-16 ***
## UniqueAuthors4      0.43113    0.03362   12.82   <2e-16 ***
## UniqueAuthors5      0.49910    0.03511   14.22   <2e-16 ***
## Year1997           -0.04827    0.05563   -0.87   0.3856
## Year1998           -0.14594    0.05835   -2.50   0.0124 *
```

```
## Year1999           -0.08697    0.05380   -1.62   0.1060
## Year2000           -0.00459    0.05231   -0.09   0.9301
## Year2001            0.04388    0.05768    0.76   0.4468
## Year2002            0.02090    0.05893    0.35   0.7229
## Year2003           -0.01489    0.05829   -0.26   0.7983
## Year2004            0.09296    0.05423    1.71   0.0865 .
## Year2005            0.06261    0.05532    1.13   0.2578
## Year2006            0.03553    0.05426    0.65   0.5126
## Year2007           -0.06777    0.04929   -1.37   0.1693
## Year2008           -0.06638    0.05473   -1.21   0.2252
## Year2009           -0.07378    0.05455   -1.35   0.1763
## Year2010           -0.11141    0.05214   -2.14   0.0327 *
## Year2011           -0.15147    0.05295   -2.86   0.0042 **
## Year2012           -0.12043    0.05382   -2.24   0.0253 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.671
## Multiple R-squared:  0.0744, Adjusted R-squared:  0.0708
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  481 weights are ~= 1. The remaining 5192 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.110   0.866   0.950   0.904   0.986   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.76e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.081  1           1.040
## LastAuthorFemale  1.086  1           1.042
## Year              1.022 16           1.001
```



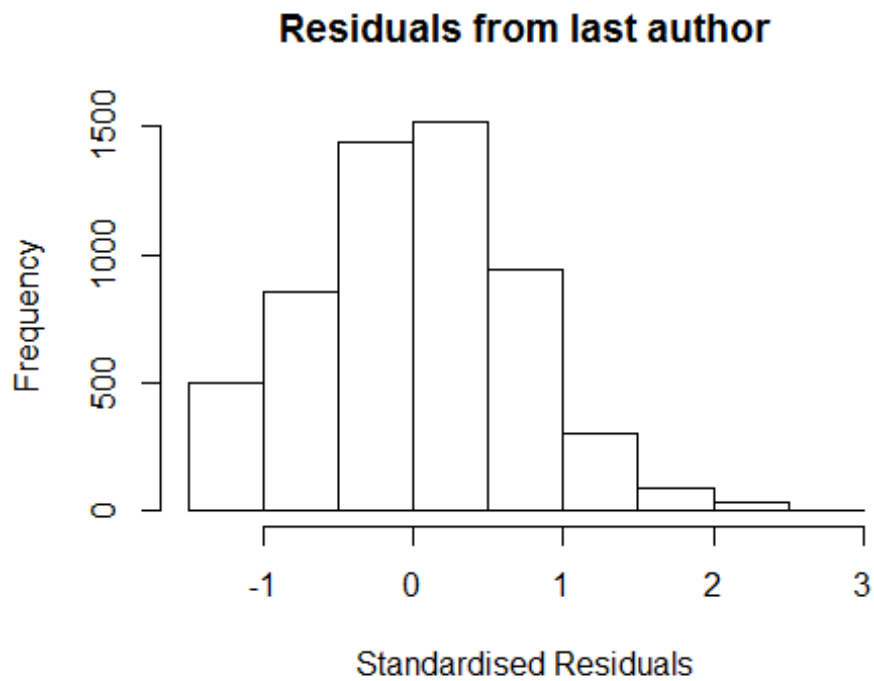

```
## [1] "List of  2 outliers with residuals above  2.5"
##         ScopusId  NLCS Year OneField Fields residuals
## 1772  0031700576 3.707 1998     1704      3     2.553
## 8526 31144478351 3.964 2006     1704      2     2.591
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.4174 -0.4644  0.0164  0.4742  2.5911
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.280966   0.041089   31.18   <2e-16 ***
## FirstAuthorFemale1 -0.049216   0.033488   -1.47    0.142
## LastAuthorFemale1  -0.041756   0.032674   -1.28    0.201
## Year1997           -0.056914   0.058536   -0.97    0.331
## Year1998           -0.126591   0.061244   -2.07    0.039 *
## Year1999           -0.098124   0.056732   -1.73    0.084 .
## Year2000           -0.000121   0.054340    0.00    0.998
## Year2001            0.090195   0.059820    1.51    0.132
## Year2002            0.034473   0.060613    0.57    0.570
## Year2003            0.006080   0.060078    0.10    0.919
## Year2004            0.136391   0.055783    2.45    0.015 *
```

```
## Year2005            0.099307   0.057494    1.73    0.084 .
## Year2006            0.091967   0.057352    1.60    0.109
## Year2007           -0.011379   0.051169   -0.22    0.824
## Year2008           -0.002237   0.056827   -0.04    0.969
## Year2009            0.025113   0.055489    0.45    0.651
## Year2010           -0.011977   0.053939   -0.22    0.824
## Year2011           -0.043976   0.055309   -0.80    0.427
## Year2012           -0.031465   0.056503   -0.56    0.578
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.697
## Multiple R-squared:  0.0102, Adjusted R-squared:  0.00704
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  470 weights are ~= 1. The remaining 5203 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.137   0.867   0.950   0.905   0.986   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.76e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                  psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.007  1           1.003
## Year              1.007 16           1.000
```




```
## [1] "List of  2 outliers with residuals above  2.5"
##         ScopusId  NLCS Year OneField Fields residuals
## 1772  0031700576 3.707 1998     1704      3     2.553
## 8526 31144478351 3.964 2006     1704      2     2.591
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.4148 -0.4655  0.0134  0.4741  2.5941
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.27742    0.04096   31.19   <2e-16 ***
## FirstAuthorFemale1 -0.06195    0.03260   -1.90    0.057 .
## Year1997           -0.05649    0.05855   -0.96    0.335
## Year1998           -0.12617    0.06134   -2.06    0.040 *
## Year1999           -0.09734    0.05676   -1.71    0.086 .
## Year2000            0.00147    0.05428    0.03    0.978
## Year2001            0.09203    0.05984    1.54    0.124
## Year2002            0.03455    0.06059    0.57    0.569
## Year2003            0.00678    0.06008    0.11    0.910
## Year2004            0.13742    0.05575    2.46    0.014 *
## Year2005            0.09909    0.05753    1.72    0.085 .
```

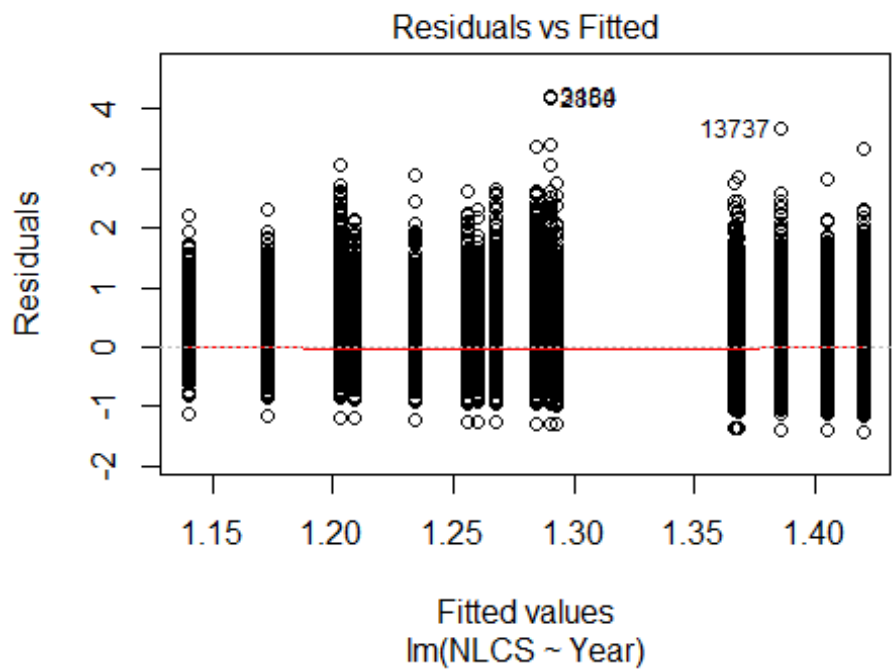
```
## Year2006            0.09243    0.05735    1.61    0.107
## Year2007           -0.01269    0.05124   -0.25    0.804
## Year2008           -0.00160    0.05685   -0.03    0.978
## Year2009            0.02678    0.05549    0.48    0.629
## Year2010           -0.01290    0.05395   -0.24    0.811
## Year2011           -0.04341    0.05531   -0.78    0.433
## Year2012           -0.03280    0.05653   -0.58    0.562
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.696
## Multiple R-squared:  0.00989,    Adjusted R-squared:  0.00691
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  471 weights are ~= 1. The remaining 5202 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.135   0.866   0.950   0.905   0.987   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.76e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##                   GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.011  1           1.006
## Year             1.011 16           1.000
```



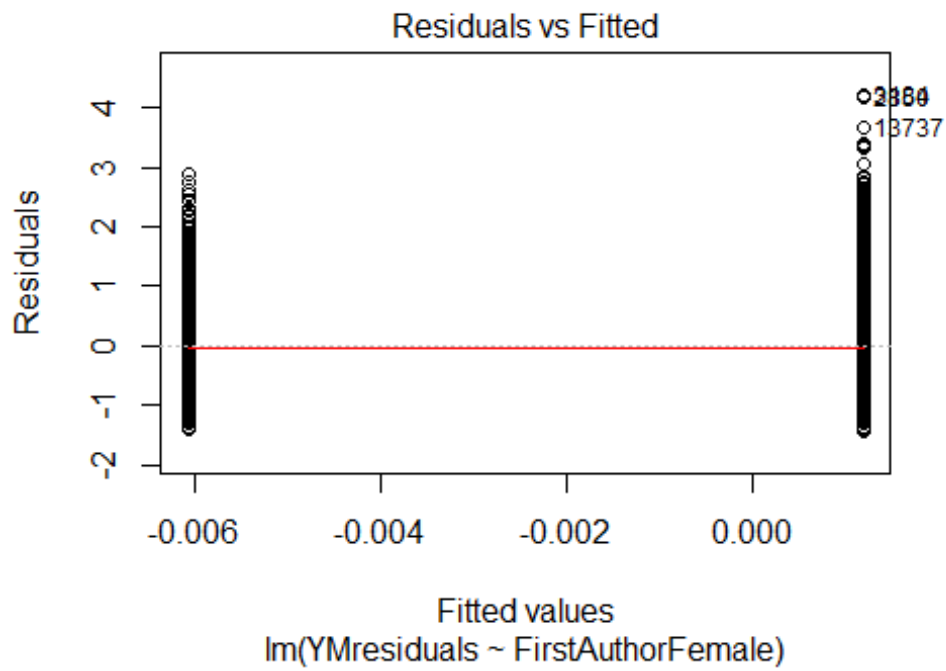

```
## [1] "List of  2 outliers with residuals above  2.5"
##         ScopusId  NLCS Year OneField Fields residuals
## 1772  0031700576 3.707 1998     1704      3     2.553
## 8526 31144478351 3.964 2006     1704      2     2.591
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.4120 -0.4643  0.0172  0.4768  2.5947
##
## Coefficients:
##                   Estimate Std. Error t value Pr(>|t|)
## (Intercept)        1.27805    0.04101   31.16   <2e-16 ***
## LastAuthorFemale1 -0.05642    0.03185   -1.77    0.077 .
## Year1997          -0.05641    0.05848   -0.96    0.335
## Year1998          -0.12684    0.06118   -2.07    0.038 *
## Year1999          -0.09841    0.05676   -1.73    0.083 .
## Year2000          -0.00129    0.05427   -0.02    0.981
## Year2001           0.08831    0.05975    1.48    0.139
## Year2002           0.03475    0.06064    0.57    0.567
## Year2003           0.00575    0.06006    0.10    0.924
## Year2004           0.13398    0.05574    2.40    0.016 *
## Year2005           0.09873    0.05736    1.72    0.085 .
```

```
## Year2006           0.09123    0.05733    1.59    0.112
## Year2007          -0.01135    0.05115   -0.22    0.824
## Year2008          -0.00348    0.05682   -0.06    0.951
## Year2009           0.02274    0.05537    0.41    0.681
## Year2010          -0.01241    0.05392   -0.23    0.818
## Year2011          -0.04554    0.05533   -0.82    0.411
## Year2012          -0.03119    0.05653   -0.55    0.581
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.696
## Multiple R-squared:  0.00975,    Adjusted R-squared:  0.00677
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  463 weights are ~= 1. The remaining 5210 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.135   0.867   0.950   0.905   0.986   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.76e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Sample size for the above analysis:  5673"
## [1] ""
## [1] ""
## [1] "####################################"
## [1] "Analysis of AJSC 1705"
## [1] "####################################"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 1006 1110 1046 1091 1218 1464 1301  952 1037 1378 1638 1625 1536 1475 1433
## 2011 2012
## 1353 1310
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  699  787  701  716  801  873  812  587  658  788  913  894  882  881  844
```

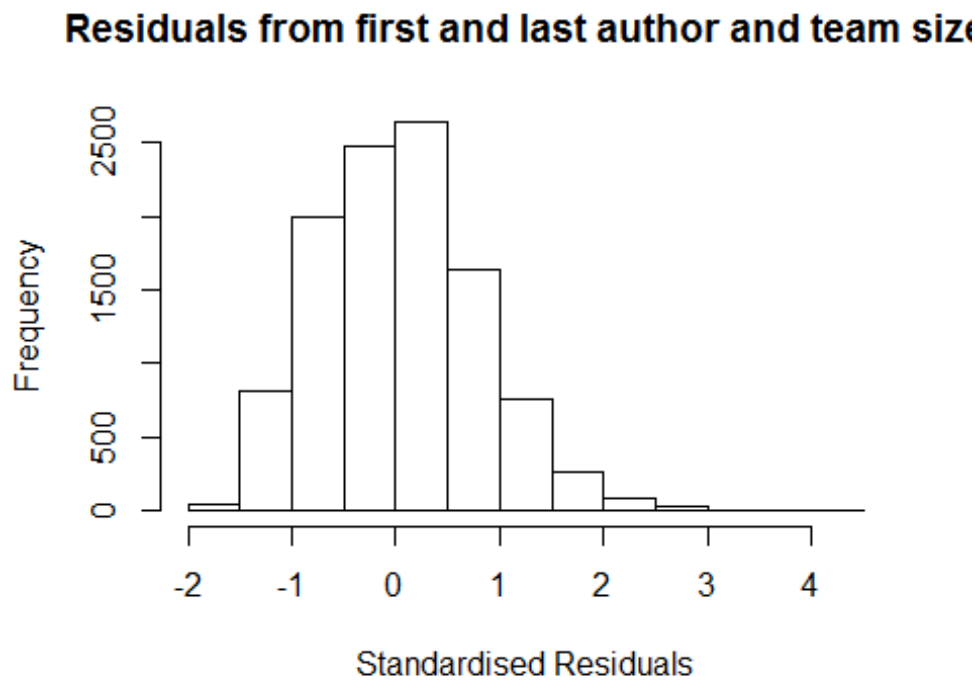
```
## 2011 2012
##  775  772
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  595  647  595  596  670  710  650  472  525  605  694  689  701  682  662
## 2011 2012
##  627  605
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
##  Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 280, df = 16, p-value <2e-16
```




```
##
##  Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 4.1, df = 1, p-value = 0.04
```



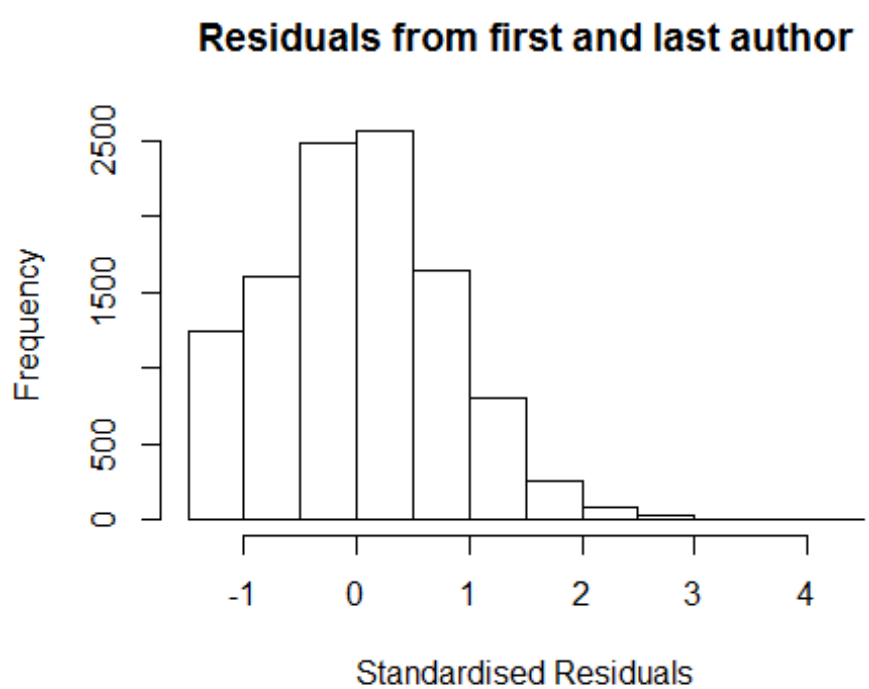

```
## [1] "Female first author team size 2018 geometric mean: 2.17269310324003"
## [1] "Male first author team size 2018 geometric mean: 2.51089443522012"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 5300, p-value = 0.1
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 2.21014624148673"
## [1] "Male last author team size 2018 geometric mean: 2.51304784105541"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 6000, p-value = 0.2
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.243  1           1.115
## LastAuthorFemale  1.244  1           1.115
## UniqueAuthors     1.064  4           1.008
## Year              1.079 16           1.002
```




```
## [1] "List of  28 outliers with residuals above  2.5"
##          ScopusId  NLCS Year OneField Fields residuals
## 17     0030262042 3.817 1996     1705      1     2.917
## 373    0030262865 4.271 1996     1705      2     3.371
## 525    0030166739 3.759 1996     1705      2     2.570
## 536    0030170353 3.905 1996     1705      2     2.716
## 537    0030171479 3.606 1996     1705      2     2.706
## 1181   0010362121 3.854 1997     1705      1     2.889
## 1445   0031383380 3.875 1997     1705      4     2.622
## 2161   0031211090 4.640 1997     1703      4     3.387
## 2296  21744462998 3.531 1997     1705      2     2.566
## 2302  27144463192 3.637 1997     1705      2     2.672
## 2311   4344578226 3.483 1997     1705      2     2.518
## 2395   0000429759 3.576 1998     1705      1     2.591
## 2850   0032183752 5.467 1998     1705      2     4.482
## 3184   0038589165 5.501 1998     1705      1     4.227
## 3221   0031999108 3.530 1998     1705      2     2.545
## 3258  85045953475 4.346 1998     1705      2     3.361
## 3497  27144489164 4.703 1998     1705      3     3.718
## 3584   0033295259 3.907 1999     1705      1     2.930
## 3781   9744277690 3.588 1999     1705      3     2.611
## 3942   0032667848 3.826 1999     1705      2     2.561
## 4096   0032632072 3.837 1999     1705      2     2.860
## 6143   0034559541 3.488 2000     1404      5     2.527
## 9130   0036688074 4.740 2002     1705      3     3.164
## 10150  0037252945 4.229 2003     1705      1     2.729
## 10738 10944266504 4.134 2004     1705      3     2.773
```

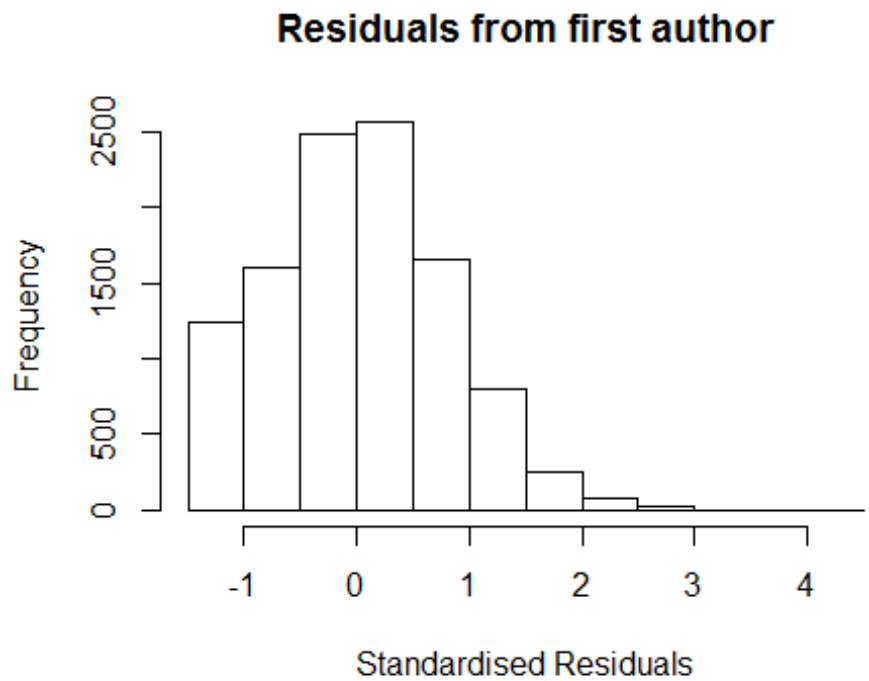
```
## 12568 13944252629 4.211 2005     1705      3     2.717
## 15605 37249053884 4.031 2007     1705      2     2.747
## 17641 46449122114 4.114 2008     1705      1     2.788
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
UniqueAuthors +
##     Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##     k.max = 1000))
##  \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.57052 -0.54240  0.00888  0.52002  4.48198
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)         0.899886   0.038399   23.44  < 2e-16 ***
## FirstAuthorFemale1  0.002566   0.022250    0.12  0.90818
## LastAuthorFemale1  -0.000419   0.022178   -0.02  0.98494
## UniqueAuthors2      0.288647   0.019178   15.05  < 2e-16 ***
## UniqueAuthors3      0.387802   0.022348   17.35  < 2e-16 ***
## UniqueAuthors4      0.461564   0.028974   15.93  < 2e-16 ***
## UniqueAuthors5      0.450282   0.031322   14.38  < 2e-16 ***
## Year1997            0.064859   0.051543    1.26  0.20829
## Year1998            0.085137   0.053205    1.60  0.10959
## Year1999            0.076759   0.051536    1.49  0.13640
## Year2000            0.061092   0.048164    1.27  0.20468
## Year2001            0.192746   0.049715    3.88  0.00011 ***
## Year2002            0.214431   0.048986    4.38  1.2e-05 ***
## Year2003            0.211975   0.051942    4.08  4.5e-05 ***
## Year2004            0.173230   0.049980    3.47  0.00053 ***
## Year2005            0.206503   0.048649    4.24  2.2e-05 ***
## Year2006            0.183667   0.047852    3.84  0.00012 ***
## Year2007            0.093125   0.045413    2.05  0.04033 *
## Year2008            0.035382   0.045284    0.78  0.43462
## Year2009            0.007989   0.045354    0.18  0.86018
## Year2010           -0.001119   0.046708   -0.02  0.98089
## Year2011           -0.081710   0.046676   -1.75  0.08005 .
## Year2012            0.068621   0.046830    1.47  0.14286
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.782
## Multiple R-squared:  0.0597, Adjusted R-squared:  0.0577
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  3 observations c(1451,1626,1789)
##   are outliers with |weight| = 0 ( < 9.3e-06);
```

```
##  947 weights are ~= 1. The remaining 9775 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0211  0.8730  0.9500  0.9110  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          9.32e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                    psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.224  1           1.106
## LastAuthorFemale  1.222  1           1.106
## Year              1.017 16           1.001
```



```
## [1] "List of  30 outliers with residuals above  2.5"
##          ScopusId  NLCS Year OneField Fields residuals
## 17     0030262042 3.817 1996     1705      1     2.703
## 89     0030413936 3.682 1996     1705      2     2.568
## 373    0030262865 4.271 1996     1705      2     3.157
## 525    0030166739 3.759 1996     1705      2     2.645
## 536    0030170353 3.905 1996     1705      2     2.791
## 964    0030370740 3.688 1996     1705      4     2.574
## 1181   0010362121 3.854 1997     1705      1     2.690
## 1445   0031383380 3.875 1997     1705      4     2.711
## 1545   0031274649 3.708 1997     1705      4     2.544
## 2161   0031211090 4.640 1997     1703      4     3.476
## 2188   0031246031 3.912 1997     1705      3     2.748
## 2293  21744433274 3.850 1997     1705      2     2.686
## 2850   0032183752 5.467 1998     1705      2     4.285
## 3059   0032074579 3.796 1998     1705      2     2.614
## 3184   0038589165 5.501 1998     1705      1     4.319
## 3258  85045953475 4.346 1998     1705      2     3.164
## 3426   0032313923 3.699 1998     1705      4     2.517
## 3497  27144489164 4.703 1998     1705      3     3.521
## 3584   0033295259 3.907 1999     1705      1     2.713
## 3942   0032667848 3.826 1999     1705      2     2.648
## 4096   0032632072 3.837 1999     1705      2     2.643
## 5552   0033721503 3.868 2000     1705      1     2.685
## 9130   0036688074 4.740 2002     1705      3     3.379
## 10150  0037252945 4.229 2003     1705      1     2.872
## 10738 10944266504 4.134 2004     1705      3     2.821
## 12568 13944252629 4.211 2005     1705      3     2.841
## 14586 33645115901 3.975 2006     1705      1     2.623
## 15605 37249053884 4.031 2007     1705      2     2.781
## 16080 34548285057 3.838 2007     1705      2     2.572
## 17641 46449122114 4.114 2008     1705      1     2.913
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.36960 -0.53746  0.00626  0.53222  4.31866
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)          1.11366    0.03803   29.28  < 2e-16 ***
## FirstAuthorFemale1 -0.00291    0.02242   -0.13  0.89673
## LastAuthorFemale1  -0.01673    0.02230   -0.75  0.45317
## Year1997            0.04996    0.05336    0.94  0.34917
## Year1998            0.06868    0.05471    1.26  0.20942
## Year1999            0.08073    0.05313    1.52  0.12868
```

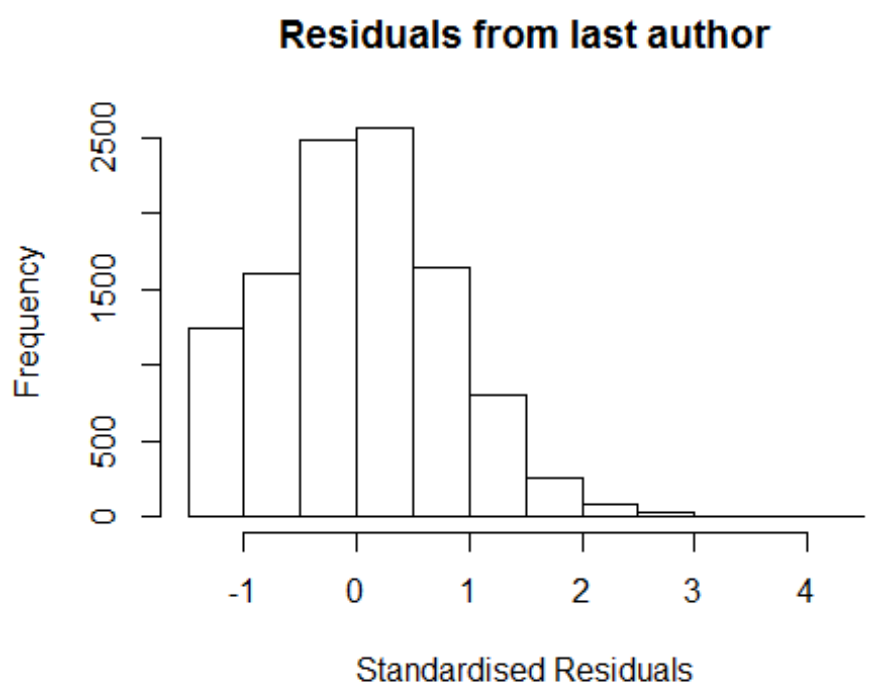
```
## Year2000            0.08595    0.04955    1.73  0.08282 .
## Year2001            0.21990    0.05048    4.36  1.3e-05 ***
## Year2002            0.24761    0.04966    4.99  6.3e-07 ***
## Year2003            0.24325    0.05349    4.55  5.5e-06 ***
## Year2004            0.21579    0.05168    4.18  3.0e-05 ***
## Year2005            0.25593    0.05001    5.12  3.1e-07 ***
## Year2006            0.23812    0.04923    4.84  1.3e-06 ***
## Year2007            0.15570    0.04622    3.37  0.00076 ***
## Year2008            0.08992    0.04658    1.93  0.05360 .
## Year2009            0.06463    0.04696    1.38  0.16872
## Year2010            0.04994    0.04792    1.04  0.29733
## Year2011           -0.01793    0.04774   -0.38  0.70726
## Year2012            0.13025    0.04803    2.71  0.00670 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.803
## Multiple R-squared:  0.0124, Adjusted R-squared:  0.0108
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  2 observations c(1451,1626) are outliers with |weight| = 0 ( < 9.3e-06);
##  874 weights are ~= 1. The remaining 9849 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0155  0.8670  0.9510  0.9120  0.9860  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          9.32e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.011  1           1.006
## Year              1.011 16           1.000
```




```
## [1] "List of  30 outliers with residuals above  2.5"
##          ScopusId  NLCS Year OneField Fields residuals
## 17     0030262042 3.817 1996     1705      1     2.703
## 89     0030413936 3.682 1996     1705      2     2.568
## 373    0030262865 4.271 1996     1705      2     3.157
## 525    0030166739 3.759 1996     1705      2     2.645
## 536    0030170353 3.905 1996     1705      2     2.791
## 964    0030370740 3.688 1996     1705      4     2.574
## 1181   0010362121 3.854 1997     1705      1     2.690
## 1445   0031383380 3.875 1997     1705      4     2.711
## 1545   0031274649 3.708 1997     1705      4     2.544
## 2161   0031211090 4.640 1997     1703      4     3.476
## 2188   0031246031 3.912 1997     1705      3     2.748
## 2293  21744433274 3.850 1997     1705      2     2.686
## 2850   0032183752 5.467 1998     1705      2     4.285
## 3059   0032074579 3.796 1998     1705      2     2.614
## 3184   0038589165 5.501 1998     1705      1     4.319
## 3258  85045953475 4.346 1998     1705      2     3.164
## 3426   0032313923 3.699 1998     1705      4     2.517
## 3497  27144489164 4.703 1998     1705      3     3.521
## 3584   0033295259 3.907 1999     1705      1     2.713
## 3942   0032667848 3.826 1999     1705      2     2.648
## 4096   0032632072 3.837 1999     1705      2     2.643
## 5552   0033721503 3.868 2000     1705      1     2.685
## 9130   0036688074 4.740 2002     1705      3     3.379
## 10150  0037252945 4.229 2003     1705      1     2.872
## 10738 10944266504 4.134 2004     1705      3     2.821
```

```
## 12568 13944252629 4.211 2005     1705      3     2.841
## 14586 33645115901 3.975 2006     1705      1     2.623
## 15605 37249053884 4.031 2007     1705      2     2.781
## 16080 34548285057 3.838 2007     1705      2     2.572
## 17641 46449122114 4.114 2008     1705      1     2.913
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.36801 -0.53836  0.00686  0.53287  4.32015
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.11245    0.03799   29.28  < 2e-16 ***
## FirstAuthorFemale1 -0.00992    0.02041   -0.49  0.62687
## Year1997            0.04991    0.05336    0.94  0.34962
## Year1998            0.06840    0.05472    1.25  0.21134
## Year1999            0.08051    0.05313    1.52  0.12970
## Year2000            0.08601    0.04955    1.74  0.08262 .
## Year2001            0.22032    0.05049    4.36  1.3e-05 ***
## Year2002            0.24752    0.04965    4.98  6.3e-07 ***
## Year2003            0.24294    0.05350    4.54  5.7e-06 ***
## Year2004            0.21553    0.05167    4.17  3.0e-05 ***
## Year2005            0.25556    0.05000    5.11  3.3e-07 ***
## Year2006            0.23728    0.04919    4.82  1.4e-06 ***
## Year2007            0.15531    0.04622    3.36  0.00078 ***
## Year2008            0.08961    0.04659    1.92  0.05445 .
## Year2009            0.06367    0.04692    1.36  0.17483
## Year2010            0.04923    0.04791    1.03  0.30416
## Year2011           -0.01850    0.04771   -0.39  0.69823
## Year2012            0.12944    0.04800    2.70  0.00702 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.803
## Multiple R-squared:  0.0124, Adjusted R-squared:  0.0108
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  2 observations c(1451,1626) are outliers with |weight| = 0 ( < 9.3e-06);
##  884 weights are ~= 1. The remaining 9839 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0153  0.8660  0.9510  0.9120  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
```

```
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          9.32e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##                  GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.01  1           1.005
## Year             1.01 16           1.000
```

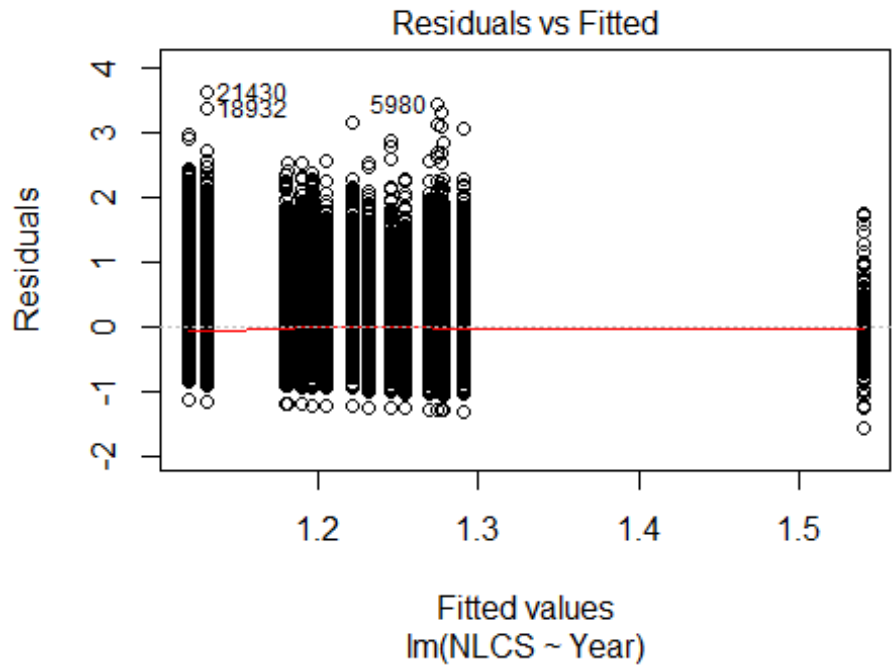



```
## [1] "List of  30 outliers with residuals above  2.5"
##          ScopusId  NLCS Year OneField Fields residuals
## 17     0030262042 3.817 1996     1705      1     2.703
## 89     0030413936 3.682 1996     1705      2     2.568
## 373    0030262865 4.271 1996     1705      2     3.157
## 525    0030166739 3.759 1996     1705      2     2.645
## 536    0030170353 3.905 1996     1705      2     2.791
## 964    0030370740 3.688 1996     1705      4     2.574
## 1181   0010362121 3.854 1997     1705      1     2.690
```

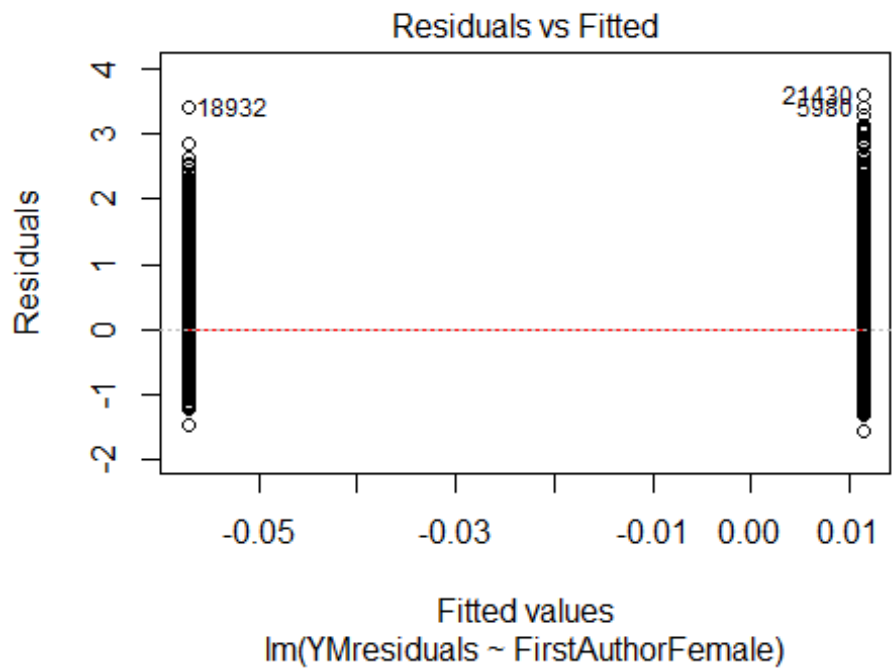
```
## 1445   0031383380 3.875 1997     1705      4     2.711
## 1545   0031274649 3.708 1997     1705      4     2.544
## 2161   0031211090 4.640 1997     1703      4     3.476
## 2188   0031246031 3.912 1997     1705      3     2.748
## 2293  21744433274 3.850 1997     1705      2     2.686
## 2850   0032183752 5.467 1998     1705      2     4.285
## 3059   0032074579 3.796 1998     1705      2     2.614
## 3184   0038589165 5.501 1998     1705      1     4.319
## 3258  85045953475 4.346 1998     1705      2     3.164
## 3426   0032313923 3.699 1998     1705      4     2.517
## 3497  27144489164 4.703 1998     1705      3     3.521
## 3584   0033295259 3.907 1999     1705      1     2.713
## 3942   0032667848 3.826 1999     1705      2     2.648
## 4096   0032632072 3.837 1999     1705      2     2.643
## 5552   0033721503 3.868 2000     1705      1     2.685
## 9130   0036688074 4.740 2002     1705      3     3.379
## 10150  0037252945 4.229 2003     1705      1     2.872
## 10738 10944266504 4.134 2004     1705      3     2.821
## 12568 13944252629 4.211 2005     1705      3     2.841
## 14586 33645115901 3.975 2006     1705      1     2.623
## 15605 37249053884 4.031 2007     1705      2     2.781
## 16080 34548285057 3.838 2007     1705      2     2.572
## 17641 46449122114 4.114 2008     1705      1     2.913
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.36932 -0.53744  0.00661  0.53250  4.31886
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)          1.1135     0.0380   29.30  < 2e-16 ***
## LastAuthorFemale1   -0.0180     0.0203   -0.89  0.37521
## Year1997             0.0499     0.0534    0.94  0.34955
## Year1998             0.0686     0.0547    1.25  0.20980
## Year1999             0.0806     0.0531    1.52  0.12916
## Year2000             0.0858     0.0495    1.73  0.08320 .
## Year2001             0.2197     0.0505    4.35  1.3e-05 ***
## Year2002             0.2475     0.0497    4.98  6.3e-07 ***
## Year2003             0.2431     0.0535    4.54  5.6e-06 ***
## Year2004             0.2157     0.0517    4.17  3.0e-05 ***
## Year2005             0.2558     0.0500    5.11  3.2e-07 ***
## Year2006             0.2380     0.0492    4.83  1.4e-06 ***
## Year2007             0.1555     0.0462    3.37  0.00077 ***
## Year2008             0.0897     0.0466    1.93  0.05404 .
## Year2009             0.0645     0.0469    1.37  0.16975
```

```
## Year2010            0.0498     0.0479    1.04  0.29901
## Year2011           -0.0181     0.0477   -0.38  0.70388
## Year2012            0.1301     0.0480    2.71  0.00675 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.803
## Multiple R-squared:  0.0124, Adjusted R-squared:  0.0109
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  2 observations c(1451,1626) are outliers with |weight| = 0 ( < 9.3e-06);
##  874 weights are ~= 1. The remaining 9849 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0155  0.8670  0.9510  0.9120  0.9860  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          9.32e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                  psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Sample size for the above analysis:  10725"
## [1] ""
## [1] ""
## [1] "####################################"
## [1] "Analysis of AJSC 1706"
## [1] "####################################"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 3016  150 2603 2538 3155 4877 4257 2718 2993 3489 3971 4271 4264 4494 4398
## 2011 2012
## 4465 4190
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 1678   91 1436 1401 1823 2598 2637 1640 1781 1999 2233 2429 2478 2690 2580
## 2011 2012
## 2659 2472
##
```

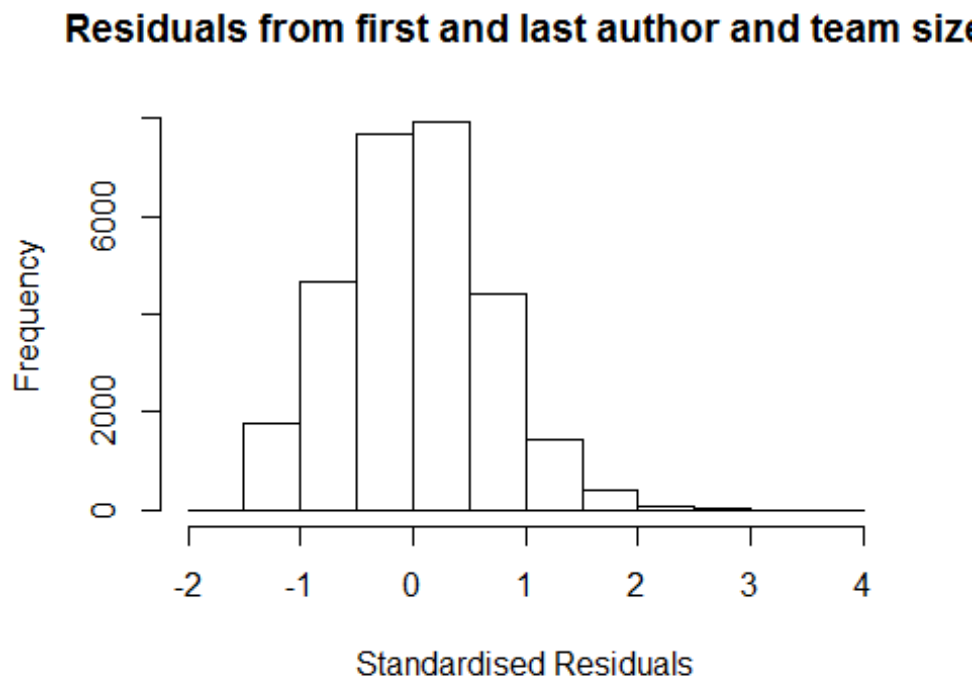
```
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 1421   76 1236 1175 1533 2142 2227 1345 1461 1618 1822 1969 2016 2179 2070
## 2011 2012
## 2123 2008
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
##  Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 620, df = 16, p-value <2e-16
```




```
##
##  Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 0.11, df = 1, p-value = 0.7
```



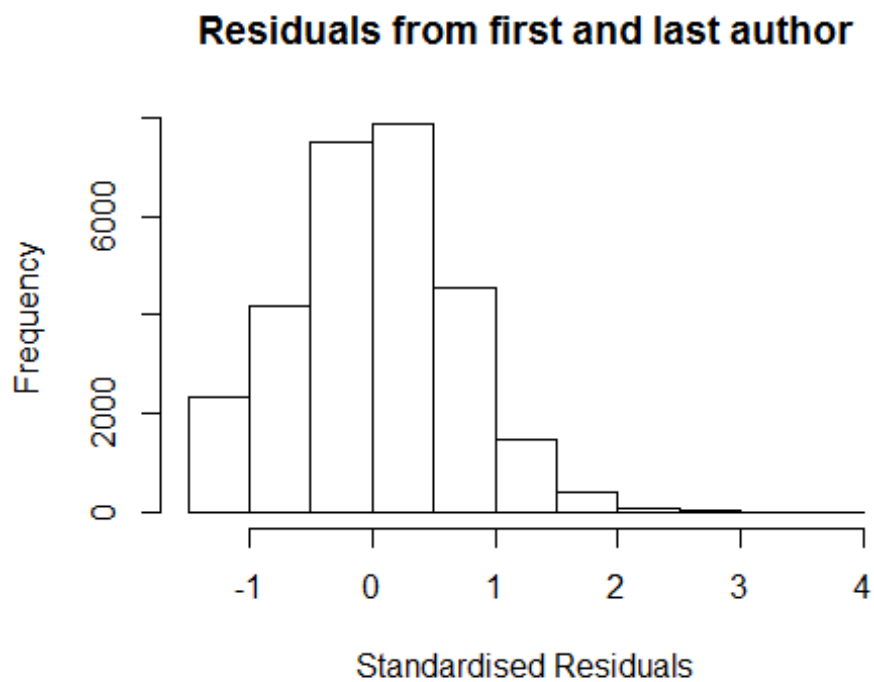

```
## [1] "Female first author team size 2018 geometric mean: 2.63208966521196"
## [1] "Male first author team size 2018 geometric mean: 2.79901922316622"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 110000, p-value = 0.3
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 2.5772767587017"
## [1] "Male last author team size 2018 geometric mean: 2.81716472952817"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 110000, p-value = 0.1
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.117  1           1.057
## LastAuthorFemale  1.114  1           1.055
## UniqueAuthors     1.056  4           1.007
## Year              1.062 16           1.002
```




```
## [1] "List of  21 outliers with residuals above  2.5"
##          ScopusId  NLCS Year OneField Fields residuals
## 188    0030107810 3.576 1996     1706      3     2.591
## 2421   0030192824 3.488 1996     1706      2     2.503
## 3756   0032047505 3.905 1998     1706      3     2.850
## 5980  27144489164 4.703 1998     1705      3     3.648
## 6323   0031773680 3.959 1998     1303      6     2.625
## 13289  9744248703 3.506 2001     1404      4     2.500
## 13475  0035517699 3.334 2001     1703      3     2.607
## 18932 33646589837 4.485 2002     1706      3     3.356
## 19654  0036670170 3.701 2002     1705      4     2.561
## 20081  0036611573 3.387 2002     1706      3     2.526
## 20084  0036611602 3.834 2002     1706      3     2.661
## 21430  0036688074 4.740 2002     1705      3     3.560
## 27598  4644244041 4.109 2004     1706      3     2.771
## 30191 29144439194 3.812 2005     1706      3     2.501
## 34404 33947416035 3.963 2006     1706      3     2.631
## 37011 33645712892 4.593 2006     1706      3     3.540
## 37366 33644997478 3.797 2006     1706      3     2.744
## 37629 31744440684 4.357 2006     1706      3     2.992
## 38895 64649083745 4.114 2007     1706      3     2.867
## 39498 37249053884 4.031 2007     1705      2     2.828
## 40665 34548285057 3.838 2007     1705      2     2.545
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
UniqueAuthors +
```

```
##     Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##     k.max = 1000))
##  \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.57852 -0.44705  0.00692  0.44925  3.64807
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         0.98477    0.02340   42.08  < 2e-16 ***
## FirstAuthorFemale1 -0.04378    0.01120   -3.91  9.3e-05 ***
## LastAuthorFemale1  -0.05674    0.01186   -4.78  1.7e-06 ***
## UniqueAuthors2      0.27915    0.01122   24.89  < 2e-16 ***
## UniqueAuthors3      0.31171    0.01249   24.97  < 2e-16 ***
## UniqueAuthors4      0.31920    0.01500   21.28  < 2e-16 ***
## UniqueAuthors5      0.28454    0.01400   20.32  < 2e-16 ***
## Year1997            0.35838    0.09130    3.93  8.7e-05 ***
## Year1998            0.07016    0.03106    2.26   0.0239 *
## Year1999            0.09008    0.03101    2.91   0.0037 **
## Year2000            0.01297    0.02974    0.44   0.6626
## Year2001           -0.15757    0.02959   -5.33  1.0e-07 ***
## Year2002           -0.12351    0.02786   -4.43  9.3e-06 ***
## Year2003            0.03179    0.02874    1.11   0.2687
## Year2004            0.07442    0.02812    2.65   0.0081 **
## Year2005            0.04678    0.02787    1.68   0.0933 .
## Year2006            0.06852    0.02750    2.49   0.0127 *
## Year2007            0.03982    0.02631    1.51   0.1301
## Year2008            0.01321    0.02608    0.51   0.6124
## Year2009           -0.00142    0.02587   -0.06   0.9561
## Year2010           -0.03085    0.02594   -1.19   0.2344
## Year2011           -0.04054    0.02603   -1.56   0.1194
## Year2012           -0.04726    0.02640   -1.79   0.0734 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.663
## Multiple R-squared:  0.0467, Adjusted R-squared:  0.046
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
##  4 observations c(2533,7736,8804,15413)
##   are outliers with |weight| = 0 ( < 3.5e-06);
##  2434 weights are ~= 1. The remaining 25983 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0053  0.8660  0.9500  0.9070  0.9860  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
```

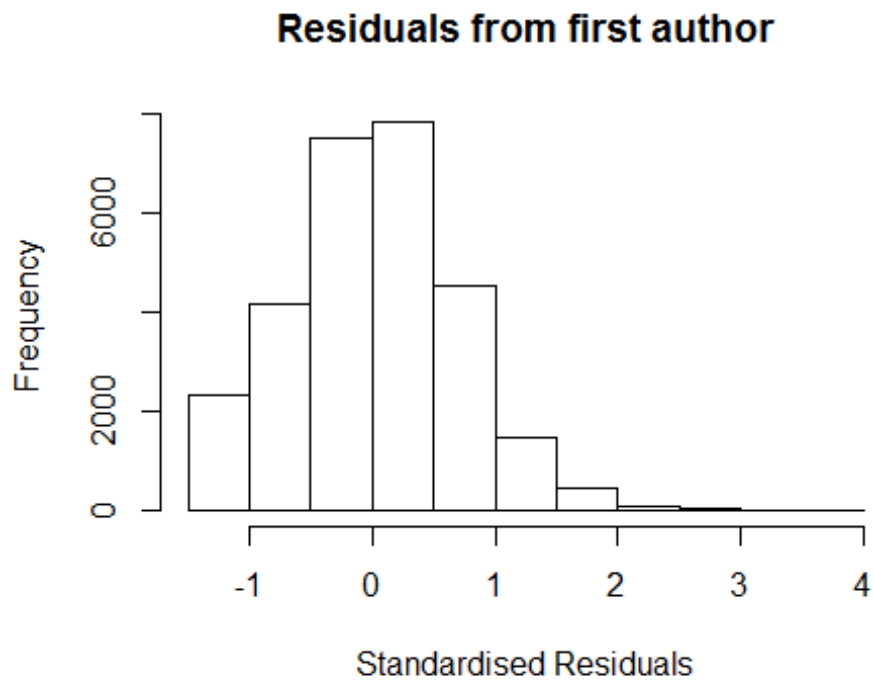
```
##          1.00e-07          1.00e-07          3.52e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.097  1           1.047
## LastAuthorFemale  1.095  1           1.046
## Year              1.007 16           1.000
```




```
## [1] "List of  23 outliers with residuals above  2.5"
##          ScopusId  NLCS Year OneField Fields residuals
## 3756   0032047505 3.905 1998     1706      3     2.660
## 5980  27144489164 4.703 1998     1705      3     3.458
## 6323   0031773680 3.959 1998     1303      6     2.714
## 13289  9744248703 3.506 2001     1404      4     2.549
## 18682 67549103767 3.604 2002     1401      2     2.550
## 18932 33646589837 4.485 2002     1706      3     3.431
```

```
## 19149  0036856989 3.589 2002     1702      5     2.535
## 19654  0036670170 3.701 2002     1705      4     2.607
## 20084  0036611602 3.834 2002     1706      3     2.740
## 21430  0036688074 4.740 2002     1705      3     3.646
## 27598  4644244041 4.109 2004     1706      3     2.834
## 30191 29144439194 3.812 2005     1706      3     2.556
## 34404 33947416035 3.963 2006     1706      3     2.678
## 37011 33645712892 4.593 2006     1706      3     3.308
## 37366 33644997478 3.797 2006     1706      3     2.512
## 37629 31744440684 4.357 2006     1706      3     3.072
## 38895 64649083745 4.114 2007     1706      3     2.929
## 39498 37249053884 4.031 2007     1705      2     2.886
## 40665 34548285057 3.838 2007     1705      2     2.622
## 45447 48849113556 3.758 2008     1706      3     2.533
## 51504 65449136284 3.759 2009     1303      7     2.546
## 66318 84867288643 3.693 2012     1606      4     2.522
## 69028 82455167807 3.570 2012     1705      3     2.510
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.4867 -0.4533  0.0113  0.4580  3.6462
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.16899    0.02281   51.24  < 2e-16 ***
## FirstAuthorFemale1 -0.03998    0.01136   -3.52  0.00043 ***
## LastAuthorFemale1  -0.07143    0.01208   -5.91  3.4e-09 ***
## Year1997            0.31767    0.08636    3.68  0.00024 ***
## Year1998            0.07614    0.03186    2.39  0.01686 *
## Year1999            0.08871    0.03141    2.82  0.00474 **
## Year2000            0.04083    0.03016    1.35  0.17579
## Year2001           -0.10017    0.02944   -3.40  0.00067 ***
## Year2002           -0.07523    0.02785   -2.70  0.00691 **
## Year2003            0.06150    0.02943    2.09  0.03665 *
## Year2004            0.10597    0.02870    3.69  0.00022 ***
## Year2005            0.08727    0.02841    3.07  0.00213 **
## Year2006            0.11577    0.02798    4.14  3.5e-05 ***
## Year2007            0.08696    0.02668    3.26  0.00112 **
## Year2008            0.05602    0.02655    2.11  0.03487 *
## Year2009            0.04381    0.02624    1.67  0.09495 .
## Year2010            0.01795    0.02631    0.68  0.49500
## Year2011            0.01268    0.02645    0.48  0.63170
## Year2012            0.00214    0.02693    0.08  0.93668
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

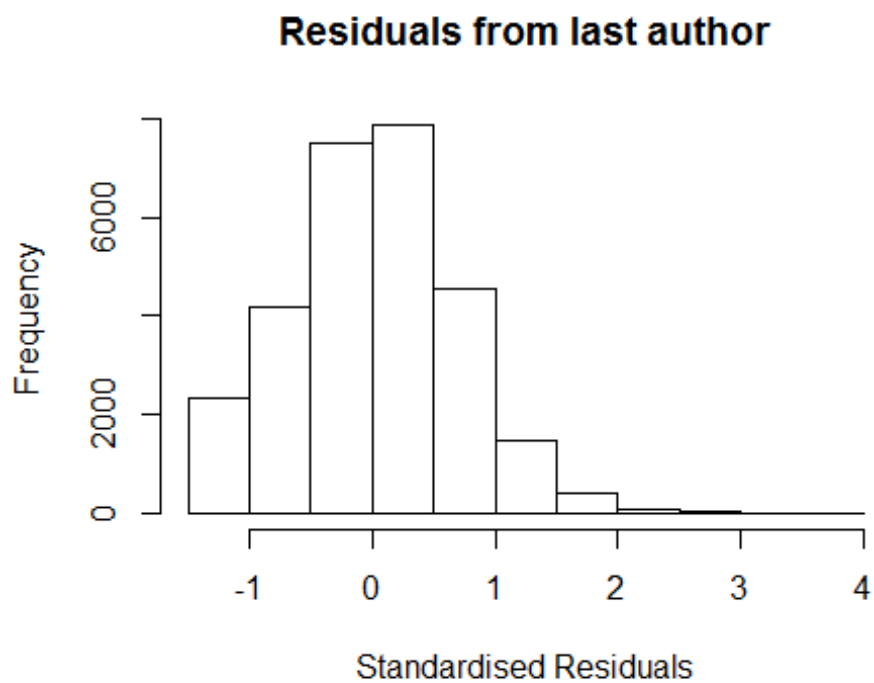
```
##
## Robust residual standard error: 0.675
## Multiple R-squared:  0.011,  Adjusted R-squared:  0.0104
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  4 observations c(2533,7736,8804,15413)
##   are outliers with |weight| = 0 ( < 3.5e-06);
##  2426 weights are ~= 1. The remaining 25991 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0031  0.8690  0.9500  0.9080  0.9860  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          3.52e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##                     GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.004  1           1.002
## Year              1.004 16           1.000
```




```
## [1] "List of  23 outliers with residuals above  2.5"
##          ScopusId  NLCS Year OneField Fields residuals
## 3756   0032047505 3.905 1998     1706      3     2.660
## 5980  27144489164 4.703 1998     1705      3     3.458
## 6323   0031773680 3.959 1998     1303      6     2.714
## 13289  9744248703 3.506 2001     1404      4     2.549
## 18682 67549103767 3.604 2002     1401      2     2.550
## 18932 33646589837 4.485 2002     1706      3     3.431
## 19149  0036856989 3.589 2002     1702      5     2.535
## 19654  0036670170 3.701 2002     1705      4     2.607
## 20084  0036611602 3.834 2002     1706      3     2.740
## 21430  0036688074 4.740 2002     1705      3     3.646
## 27598  4644244041 4.109 2004     1706      3     2.834
## 30191 29144439194 3.812 2005     1706      3     2.556
## 34404 33947416035 3.963 2006     1706      3     2.678
## 37011 33645712892 4.593 2006     1706      3     3.308
## 37366 33644997478 3.797 2006     1706      3     2.512
## 37629 31744440684 4.357 2006     1706      3     3.072
## 38895 64649083745 4.114 2007     1706      3     2.929
## 39498 37249053884 4.031 2007     1705      2     2.886
## 40665 34548285057 3.838 2007     1705      2     2.622
## 45447 48849113556 3.758 2008     1706      3     2.533
## 51504 65449136284 3.759 2009     1303      7     2.546
## 66318 84867288643 3.693 2012     1606      4     2.522
## 69028 82455167807 3.570 2012     1705      3     2.510
##
## Call:
```

```
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.4837 -0.4540  0.0125  0.4575  3.6514
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.163370   0.022814   50.99  < 2e-16 ***
## FirstAuthorFemale1 -0.064681   0.010998   -5.88  4.1e-09 ***
## Year1997            0.320344   0.085919    3.73  0.00019 ***
## Year1998            0.075363   0.031895    2.36  0.01814 *
## Year1999            0.088945   0.031461    2.83  0.00470 **
## Year2000            0.041062   0.030194    1.36  0.17385
## Year2001           -0.099453   0.029437   -3.38  0.00073 ***
## Year2002           -0.074740   0.027867   -2.68  0.00732 **
## Year2003            0.060745   0.029449    2.06  0.03915 *
## Year2004            0.105247   0.028760    3.66  0.00025 ***
## Year2005            0.086410   0.028464    3.04  0.00240 **
## Year2006            0.115171   0.028022    4.11  4.0e-05 ***
## Year2007            0.085215   0.026705    3.19  0.00142 **
## Year2008            0.054158   0.026590    2.04  0.04168 *
## Year2009            0.042102   0.026271    1.60  0.10903
## Year2010            0.016722   0.026336    0.63  0.52547
## Year2011            0.010427   0.026490    0.39  0.69386
## Year2012           -0.000377   0.026973   -0.01  0.98885
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.676
## Multiple R-squared:  0.00974,    Adjusted R-squared:  0.00915
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  4 observations c(2533,7736,8804,15413)
##   are outliers with |weight| = 0 ( < 3.5e-06);
##  2387 weights are ~= 1. The remaining 26030 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0029  0.8700  0.9510  0.9080  0.9860  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          3.52e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
```

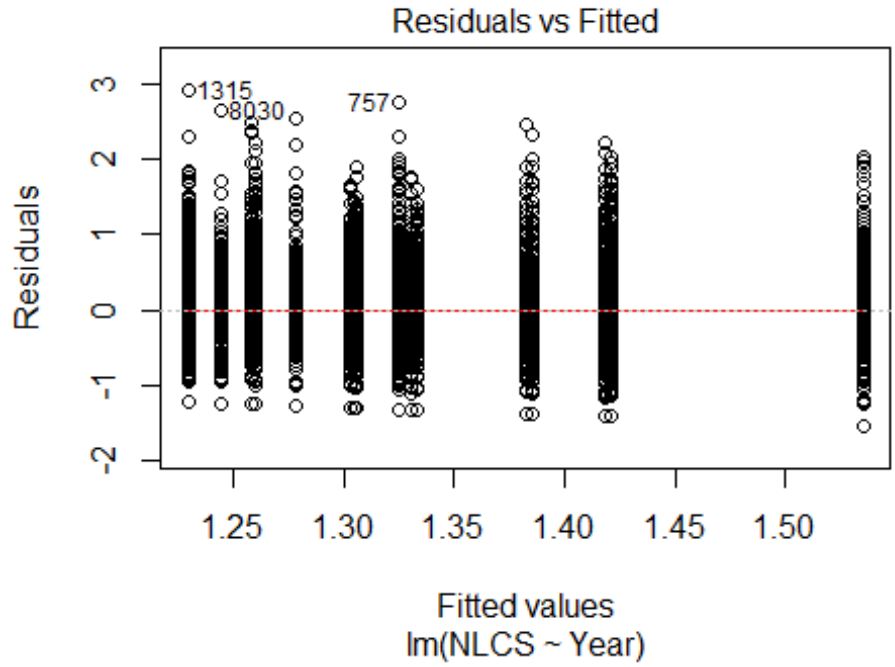
```
##   trace.lev         mts  compute.rd
##           0        1000           0
##                    psi           subsampling                    cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                   "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.003  1           1.001
## Year             1.003 16           1.000
```



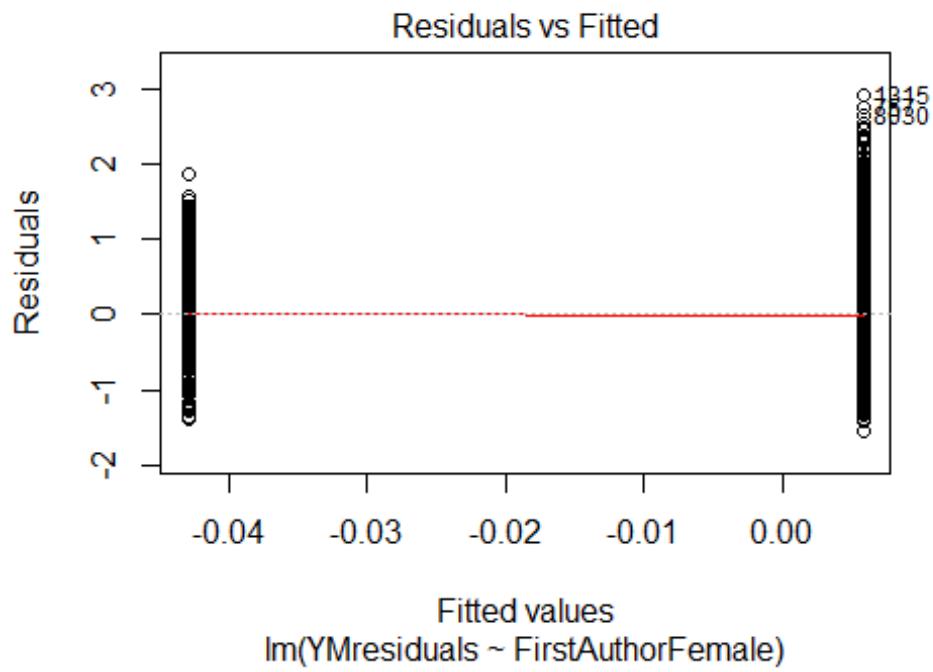

```
## [1] "List of  23 outliers with residuals above  2.5"
##          ScopusId  NLCS Year OneField Fields residuals
## 3756   0032047505 3.905 1998     1706      3     2.660
## 5980  27144489164 4.703 1998     1705      3     3.458
## 6323   0031773680 3.959 1998     1303      6     2.714
## 13289  9744248703 3.506 2001     1404      4     2.549
## 18682 67549103767 3.604 2002     1401      2     2.550
## 18932 33646589837 4.485 2002     1706      3     3.431
## 19149  0036856989 3.589 2002     1702      5     2.535
## 19654  0036670170 3.701 2002     1705      4     2.607
## 20084  0036611602 3.834 2002     1706      3     2.740
## 21430  0036688074 4.740 2002     1705      3     3.646
## 27598  4644244041 4.109 2004     1706      3     2.834
## 30191 29144439194 3.812 2005     1706      3     2.556
## 34404 33947416035 3.963 2006     1706      3     2.678
```

```
## 37011 33645712892 4.593 2006     1706      3     3.308
## 37366 33644997478 3.797 2006     1706      3     2.512
## 37629 31744440684 4.357 2006     1706      3     3.072
## 38895 64649083745 4.114 2007     1706      3     2.929
## 39498 37249053884 4.031 2007     1705      2     2.886
## 40665 34548285057 3.838 2007     1705      2     2.622
## 45447 48849113556 3.758 2008     1706      3     2.533
## 51504 65449136284 3.759 2009     1303      7     2.546
## 66318 84867288643 3.693 2012     1606      4     2.522
## 69028 82455167807 3.570 2012     1705      3     2.510
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.4836 -0.4542  0.0113  0.4568  3.6508
##
## Coefficients:
##                   Estimate Std. Error t value Pr(>|t|)
## (Intercept)        1.16500    0.02277   51.17  < 2e-16 ***
## LastAuthorFemale1 -0.08675    0.01165   -7.45  9.9e-14 ***
## Year1997           0.31862    0.08654    3.68  0.00023 ***
## Year1998           0.07674    0.03186    2.41  0.01602 *
## Year1999           0.08938    0.03141    2.85  0.00443 **
## Year2000           0.04166    0.03015    1.38  0.16714
## Year2001          -0.10008    0.02944   -3.40  0.00068 ***
## Year2002          -0.07582    0.02785   -2.72  0.00648 **
## Year2003           0.06088    0.02944    2.07  0.03863 *
## Year2004           0.10499    0.02870    3.66  0.00025 ***
## Year2005           0.08660    0.02841    3.05  0.00230 **
## Year2006           0.11462    0.02799    4.09  4.2e-05 ***
## Year2007           0.08623    0.02668    3.23  0.00123 **
## Year2008           0.05541    0.02655    2.09  0.03690 *
## Year2009           0.04318    0.02624    1.65  0.09995 .
## Year2010           0.01682    0.02630    0.64  0.52258
## Year2011           0.01125    0.02646    0.43  0.67078
## Year2012           0.00132    0.02694    0.05  0.96106
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.675
## Multiple R-squared:  0.0106, Adjusted R-squared:  0.00997
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  4 observations c(2533,7736,8804,15413)
##   are outliers with |weight| = 0 ( < 3.5e-06);
```

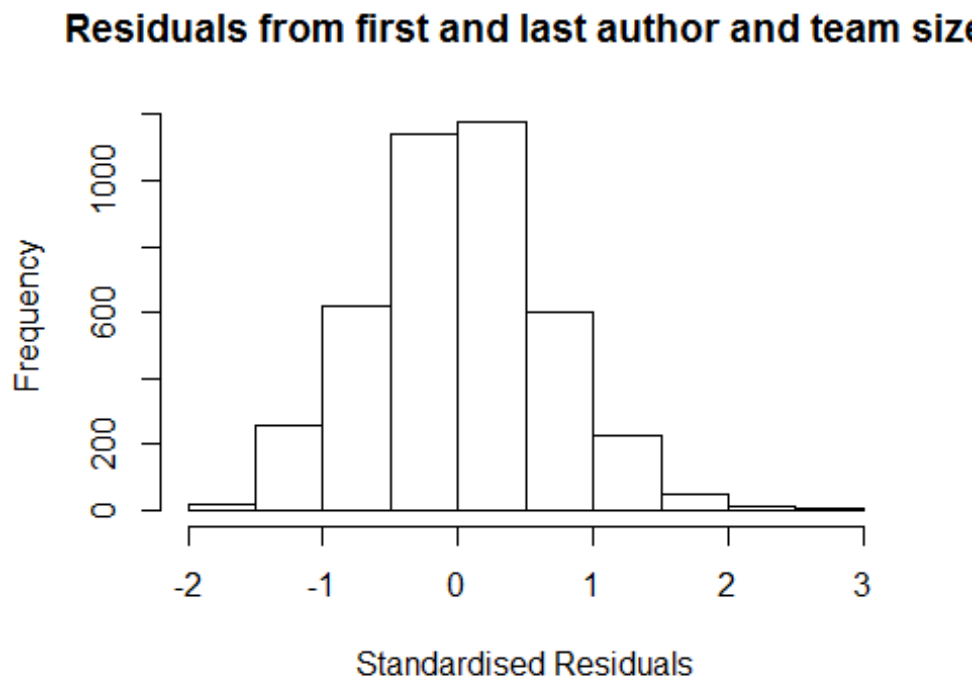
```
##  2408 weights are ~= 1. The remaining 26009 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0028  0.8690  0.9500  0.9080  0.9860  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          3.52e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Sample size for the above analysis:  28421"
## [1] ""
## [1] ""
## [1] "####################################"
## [1] "Analysis of AJSC 1707"
## [1] "####################################"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  555  619  573  545  632  533  520  439  408  467  491  524  398  431  382
## 2011 2012
##  425  401
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  339  406  388  339  392  296  312  271  241  273  285  307  236  240  200
## 2011 2012
##  251  226
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  270  340  328  274  325  240  273  219  187  221  239  250  197  202  154
## 2011 2012
##  211  173
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
##  Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 89, df = 16, p-value = 3e-12
```




```
##
##  Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 9.1, df = 1, p-value = 0.003
```

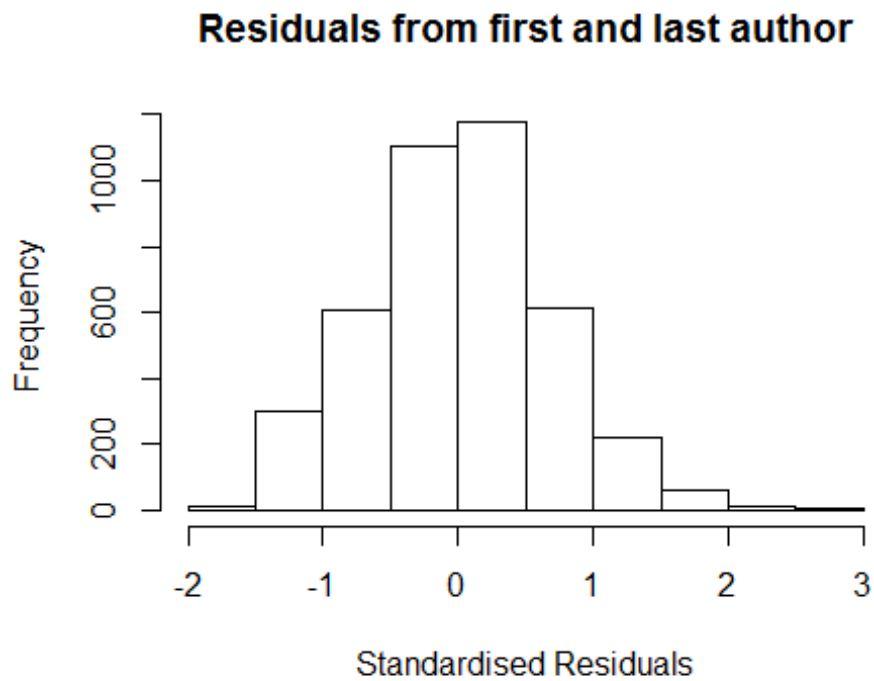



```
## [1] "Female first author team size 2018 geometric mean: 3.24021327989036"
## [1] "Male first author team size 2018 geometric mean: 2.81883435188996"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 2500, p-value = 0.4
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 2.59969696294812"
## [1] "Male last author team size 2018 geometric mean: 2.93460257740057"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 1500, p-value = 0.4
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.070  1           1.034
## LastAuthorFemale  1.063  1           1.031
## UniqueAuthors     1.108  4           1.013
## Year              1.126 16           1.004
```



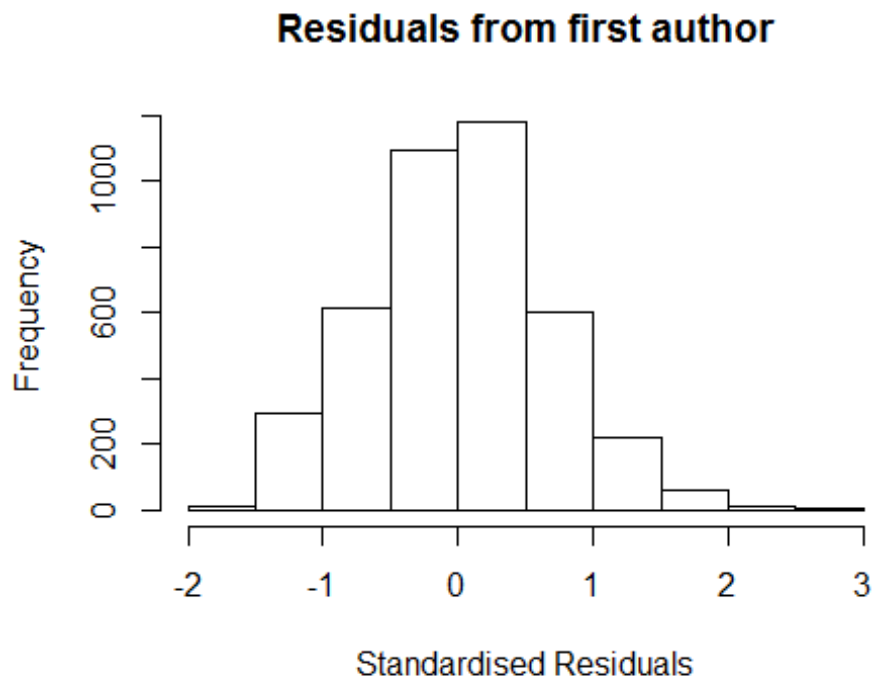

```
## [1] "List of  4 outliers with residuals above  2.5"
##         ScopusId  NLCS Year OneField Fields residuals
## 757   0031185845 4.068 1997     1702      5     2.684
## 1315  0032204063 4.140 1998     1702      4     2.845
## 6273 33646023117 3.828 2006     1702      4     2.652
## 9812 79953048649 3.731 2011     1702      5     2.518
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
UniqueAuthors +
##     Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##     k.max = 1000))
##  \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.57935 -0.42985  0.00469  0.43483  2.84523
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)           1.2170     0.0530   22.96  < 2e-16 ***
## FirstAuthorFemale1   -0.0264     0.0309   -0.85  0.39286
## LastAuthorFemale1    -0.1063     0.0335   -3.17  0.00153 **
## UniqueAuthors2        0.2937     0.0297    9.90  < 2e-16 ***
## UniqueAuthors3        0.3053     0.0336    9.09  < 2e-16 ***
## UniqueAuthors4        0.2508     0.0406    6.18  7.0e-10 ***
## UniqueAuthors5        0.2622     0.0430    6.09  1.2e-09 ***
```

```
## Year1997            -0.1381     0.0630   -2.19  0.02848 *
## Year1998            -0.2275     0.0624   -3.65  0.00027 ***
## Year1999            -0.1221     0.0652   -1.87  0.06107 .
## Year2000            -0.2242     0.0649   -3.45  0.00056 ***
## Year2001             0.0570     0.0668    0.85  0.39313
## Year2002            -0.0297     0.0670   -0.44  0.65783
## Year2003            -0.1464     0.0608   -2.41  0.01600 *
## Year2004            -0.1009     0.0685   -1.47  0.14088
## Year2005            -0.0935     0.0649   -1.44  0.14995
## Year2006            -0.0408     0.0631   -0.65  0.51785
## Year2007            -0.0851     0.0603   -1.41  0.15827
## Year2008            -0.1455     0.0639   -2.28  0.02275 *
## Year2009            -0.2033     0.0622   -3.27  0.00109 **
## Year2010            -0.2030     0.0684   -2.97  0.00301 **
## Year2011            -0.2551     0.0677   -3.77  0.00017 ***
## Year2012            -0.1912     0.0693   -2.76  0.00580 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.646
## Multiple R-squared:  0.0508, Adjusted R-squared:  0.0457
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  346 weights are ~= 1. The remaining 3757 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0136  0.8640  0.9510  0.9020  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          2.44e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                    psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.054  1           1.027
## LastAuthorFemale  1.052  1           1.026
## Year              1.024 16           1.001
```



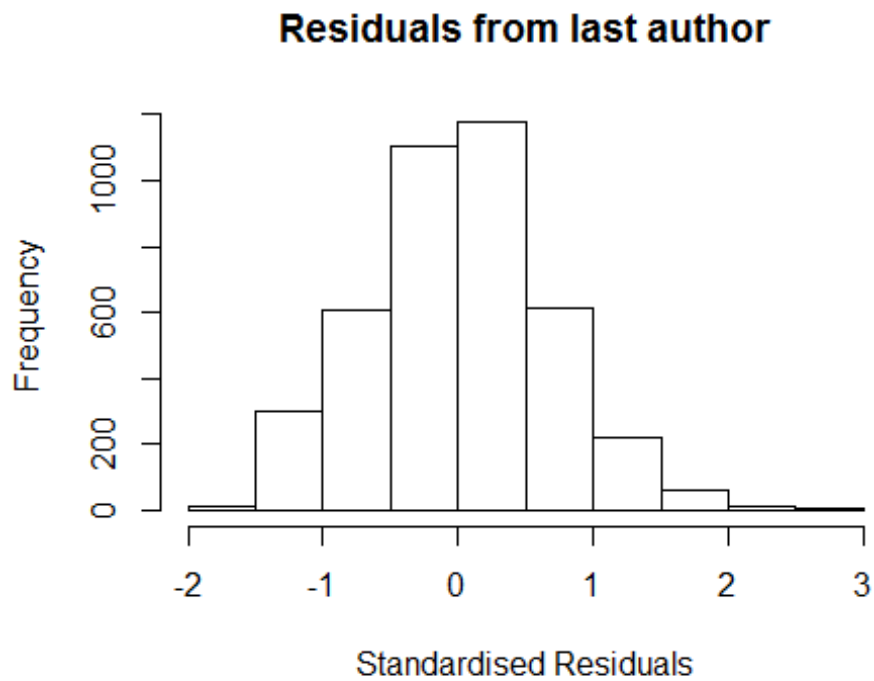

```
## [1] "List of  3 outliers with residuals above  2.5"
##         ScopusId  NLCS Year OneField Fields residuals
## 757   0031185845 4.068 1997     1702      5     2.785
## 1315  0032204063 4.140 1998     1702      4     2.934
## 9812 79953048649 3.731 2011     1702      5     2.551
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##    Min     1Q Median     3Q    Max
## -1.516 -0.444  0.011  0.436  2.934
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.44087    0.04887   29.48  < 2e-16 ***
## FirstAuthorFemale1 -0.00874    0.03125   -0.28  0.77975
## LastAuthorFemale1  -0.10060    0.03403   -2.96  0.00313 **
## Year1997           -0.15800    0.06349   -2.49  0.01287 *
## Year1998           -0.23531    0.06407   -3.67  0.00024 ***
## Year1999           -0.14625    0.06494   -2.25  0.02436 *
## Year2000           -0.21959    0.06479   -3.39  0.00071 ***
## Year2001            0.07536    0.06690    1.13  0.26004
## Year2002           -0.03808    0.06767   -0.56  0.57363
## Year2003           -0.14743    0.06185   -2.38  0.01719 *
```

```
## Year2004           -0.11073    0.07109   -1.56  0.11938
## Year2005           -0.09357    0.06630   -1.41  0.15823
## Year2006           -0.03417    0.06348   -0.54  0.59042
## Year2007           -0.07311    0.06087   -1.20  0.22982
## Year2008           -0.14422    0.06563   -2.20  0.02803 *
## Year2009           -0.18425    0.06294   -2.93  0.00344 **
## Year2010           -0.17959    0.06899   -2.60  0.00927 **
## Year2011           -0.26112    0.07027   -3.72  0.00021 ***
## Year2012           -0.18247    0.07012   -2.60  0.00930 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.657
## Multiple R-squared:  0.0195, Adjusted R-squared:  0.0152
## Convergence in 10 IRWLS iterations
##
## Robustness weights:
##  356 weights are ~= 1. The remaining 3747 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0081  0.8660  0.9510  0.9020  0.9860  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          2.44e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.013  1           1.006
## Year              1.013 16           1.000
```




```
## [1] "List of  3 outliers with residuals above  2.5"
##         ScopusId  NLCS Year OneField Fields residuals
## 757   0031185845 4.068 1997     1702      5     2.785
## 1315  0032204063 4.140 1998     1702      4     2.934
## 9812 79953048649 3.731 2011     1702      5     2.551
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.5099 -0.4418  0.0109  0.4404  2.9426
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)          1.4342     0.0487   29.45  < 2e-16 ***
## FirstAuthorFemale1  -0.0304     0.0310   -0.98  0.32651
## Year1997            -0.1590     0.0634   -2.51  0.01218 *
## Year1998            -0.2368     0.0642   -3.69  0.00023 ***
## Year1999            -0.1457     0.0648   -2.25  0.02462 *
## Year2000            -0.2186     0.0648   -3.37  0.00076 ***
## Year2001             0.0757     0.0668    1.13  0.25719
## Year2002            -0.0396     0.0678   -0.58  0.55891
## Year2003            -0.1481     0.0618   -2.40  0.01662 *
## Year2004            -0.1180     0.0708   -1.67  0.09556 .
```

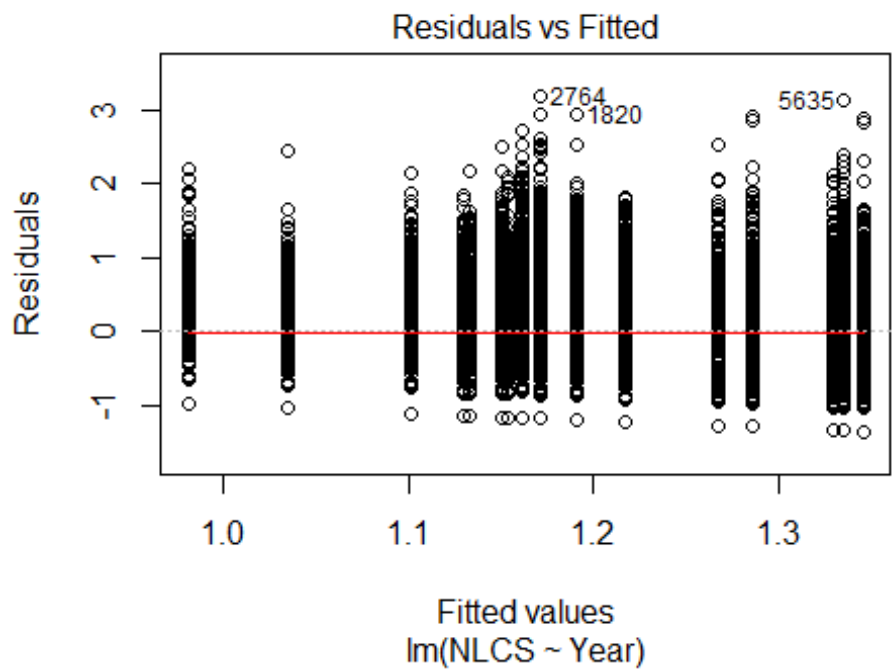
```
## Year2005            -0.0951     0.0664   -1.43  0.15189
## Year2006            -0.0368     0.0635   -0.58  0.56266
## Year2007            -0.0755     0.0609   -1.24  0.21485
## Year2008            -0.1458     0.0657   -2.22  0.02647 *
## Year2009            -0.1878     0.0628   -2.99  0.00280 **
## Year2010            -0.1841     0.0689   -2.67  0.00754 **
## Year2011            -0.2655     0.0704   -3.77  0.00017 ***
## Year2012            -0.1874     0.0704   -2.66  0.00784 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.657
## Multiple R-squared:  0.0174, Adjusted R-squared:  0.0133
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  360 weights are ~= 1. The remaining 3743 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0074  0.8660  0.9510  0.9020  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          2.44e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##                   GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.011  1           1.006
## Year             1.011 16           1.000
```



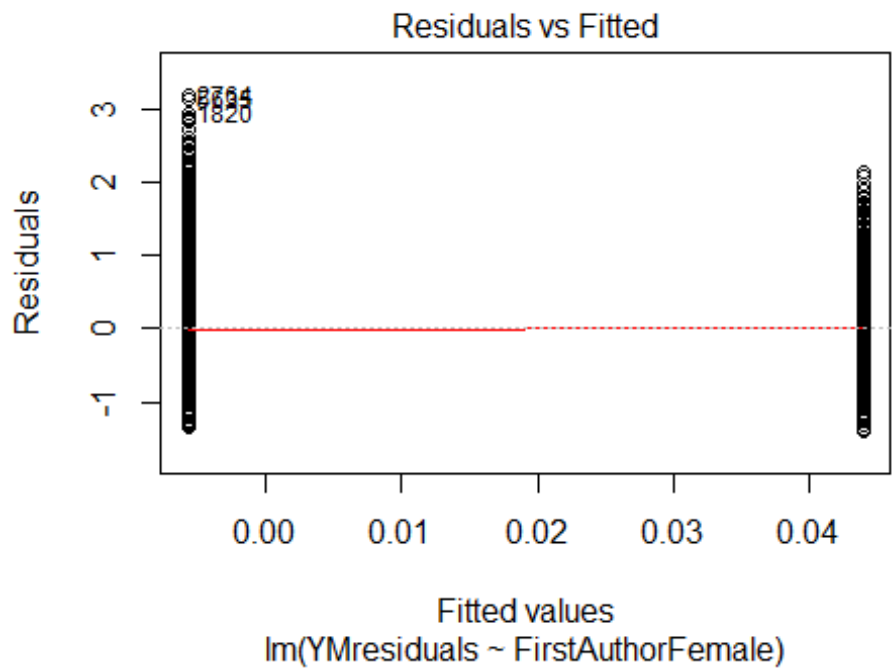

```
## [1] "List of  3 outliers with residuals above  2.5"
##         ScopusId  NLCS Year OneField Fields residuals
## 757   0031185845 4.068 1997     1702      5     2.785
## 1315  0032204063 4.140 1998     1702      4     2.934
## 9812 79953048649 3.731 2011     1702      5     2.551
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.5153 -0.4444  0.0122  0.4354  2.9352
##
## Coefficients:
##                   Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.4401     0.0488   29.49  < 2e-16 ***
## LastAuthorFemale1  -0.1027     0.0334   -3.08  0.00212 **
## Year1997           -0.1579     0.0635   -2.49  0.01291 *
## Year1998           -0.2352     0.0641   -3.67  0.00024 ***
## Year1999           -0.1462     0.0650   -2.25  0.02445 *
## Year2000           -0.2197     0.0648   -3.39  0.00070 ***
## Year2001            0.0753     0.0669    1.13  0.26049
## Year2002           -0.0382     0.0677   -0.56  0.57211
## Year2003           -0.1477     0.0618   -2.39  0.01695 *
## Year2004           -0.1108     0.0711   -1.56  0.11905
```

```
## Year2005           -0.0940     0.0662   -1.42  0.15584
## Year2006           -0.0340     0.0635   -0.54  0.59207
## Year2007           -0.0731     0.0609   -1.20  0.22957
## Year2008           -0.1443     0.0656   -2.20  0.02797 *
## Year2009           -0.1847     0.0629   -2.94  0.00334 **
## Year2010           -0.1797     0.0690   -2.61  0.00921 **
## Year2011           -0.2613     0.0703   -3.72  0.00020 ***
## Year2012           -0.1825     0.0701   -2.60  0.00928 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.657
## Multiple R-squared:  0.0195, Adjusted R-squared:  0.0154
## Convergence in 10 IRWLS iterations
##
## Robustness weights:
##  357 weights are ~= 1. The remaining 3746 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.008   0.866   0.951   0.902   0.986   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          2.44e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                  psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Sample size for the above analysis:  4103"
## [1] ""
## [1] ""
## [1] "##################################"
## [1] "Analysis of AJSC 1708"
## [1] "##################################"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  982  967  973  846 1034 1035  910  790  780  940  974 1007  990  945  858
## 2011 2012
##  903  945
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
```

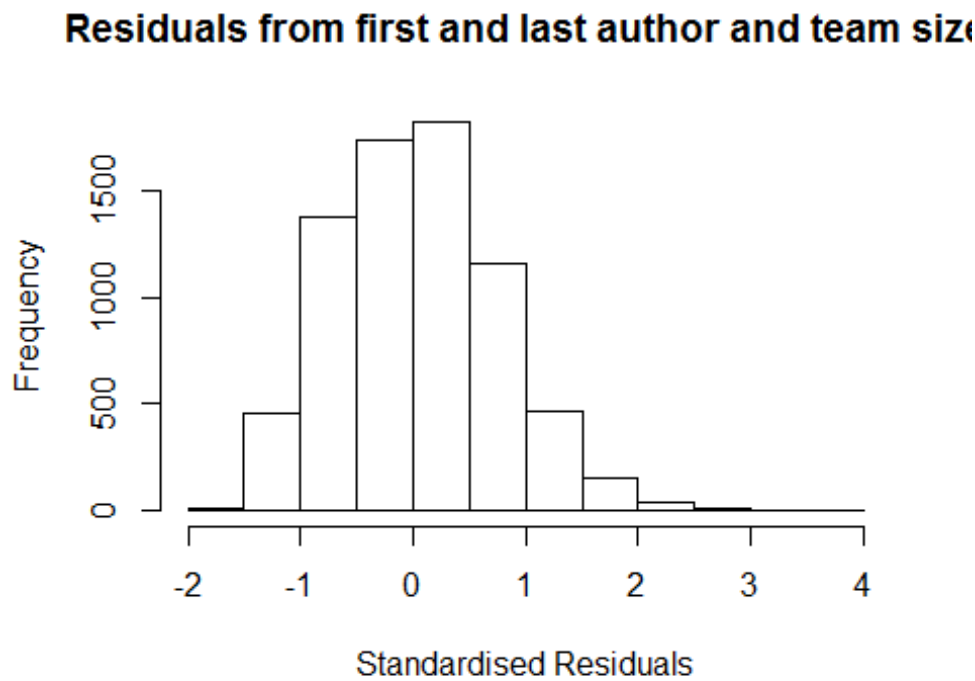
```
##  661  620  596  502  653  573  512  482  442  534  513  533  546  512  466
## 2011 2012
##  483  510
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  552  517  495  417  544  466  391  382  335  425  371  412  417  391  341
## 2011 2012
##  371  390
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
##  Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 200, df = 16, p-value <2e-16
```




```
##
##  Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 0.87, df = 1, p-value = 0.4
```



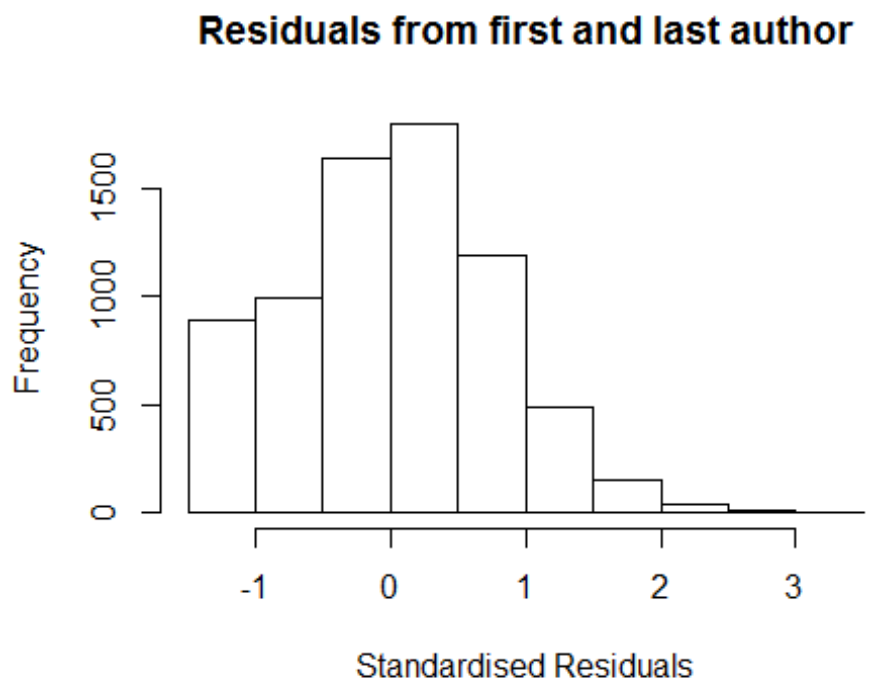

```
## [1] "Female first author team size 2018 geometric mean: 3.04189505258258"
## [1] "Male first author team size 2018 geometric mean: 2.73433839777702"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 2000, p-value = 0.6
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 2.94843711059142"
## [1] "Male last author team size 2018 geometric mean: 2.74674913004816"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 1900, p-value = 0.4
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.063  1           1.031
## LastAuthorFemale  1.063  1           1.031
## UniqueAuthors     1.066  4           1.008
## Year              1.071 16           1.002
```




```
## [1] "List of  10 outliers with residuals above  2.5"
##          ScopusId  NLCS Year OneField Fields residuals
## 1820   0031153727 4.130 1997     1708      3     2.905
## 2764  85045953475 4.346 1998     1705      2     3.511
## 3422   9744277690 3.588 1999     1705      3     2.749
## 4144   4243148480 4.090 1999     1702      5     3.251
## 5635   0035455653 4.472 2001     1708      3     3.065
## 7239   0036953758 4.202 2002     1708      1     3.329
## 8769   5444275208 3.799 2004     1705      3     2.559
## 9621  33846118079 3.662 2005     1708      3     2.691
## 10076 13944252629 4.211 2005     1705      3     2.786
## 10198 15844407150 4.159 2005     1708      2     2.575
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
UniqueAuthors +
##     Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##     k.max = 1000))
##  \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.58432 -0.51031  0.00893  0.50461  3.51062
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)          0.83200    0.03828   21.73  < 2e-16 ***
```

```
## FirstAuthorFemale1  0.05307    0.02901    1.83  0.06740 .
## LastAuthorFemale1  -0.01226    0.02922   -0.42  0.67490
## UniqueAuthors2      0.36508    0.02455   14.87  < 2e-16 ***
## UniqueAuthors3      0.45402    0.02697   16.84  < 2e-16 ***
## UniqueAuthors4      0.47759    0.03219   14.84  < 2e-16 ***
## UniqueAuthors5      0.61311    0.03365   18.22  < 2e-16 ***
## Year1997            0.02828    0.05126    0.55  0.58111
## Year1998            0.00337    0.05272    0.06  0.94900
## Year1999            0.00708    0.05349    0.13  0.89464
## Year2000           -0.02966    0.05021   -0.59  0.55470
## Year2001            0.12062    0.05345    2.26  0.02406 *
## Year2002            0.04123    0.05220    0.79  0.42970
## Year2003            0.02187    0.05155    0.42  0.67140
## Year2004            0.04306    0.05405    0.80  0.42566
## Year2005            0.13921    0.05172    2.69  0.00713 **
## Year2006            0.07520    0.05399    1.39  0.16373
## Year2007           -0.07546    0.04784   -1.58  0.11474
## Year2008           -0.06926    0.04692   -1.48  0.13994
## Year2009           -0.07755    0.04854   -1.60  0.11019
## Year2010           -0.17094    0.04950   -3.45  0.00056 ***
## Year2011           -0.25757    0.05049   -5.10  3.5e-07 ***
## Year2012           -0.14563    0.04934   -2.95  0.00317 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.737
## Multiple R-squared:  0.0841, Adjusted R-squared:  0.0813
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  observation 1337 is an outlier with |weight| = 0 ( < 1.4e-05);
##  593 weights are ~= 1. The remaining 6623 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0051  0.8780  0.9480  0.9110  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.39e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                    psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
```

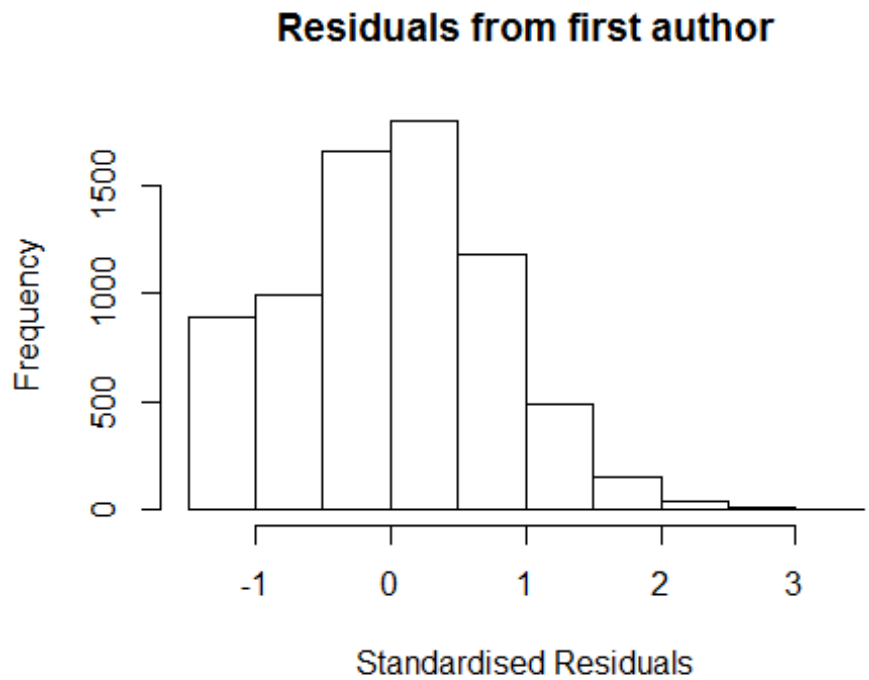
```
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##                     GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.050  1           1.024
## LastAuthorFemale  1.050  1           1.024
## Year              1.012 16           1.000
```




```
## [1] "List of  14 outliers with residuals above  2.5"
##          ScopusId  NLCS Year OneField Fields residuals
## 859    0030370740 3.688 1996     1705      4     2.604
## 1424   0031274649 3.708 1997     1705      4     2.587
## 1820   0031153727 4.130 1997     1708      3     3.009
## 2707   0002806618 3.648 1998     1702      5     2.521
## 2764  85045953475 4.346 1998     1705      2     3.219
## 3040   0032313923 3.699 1998     1705      4     2.572
## 3782   0032592464 3.678 1999     1705      3     2.561
## 4144   4243148480 4.090 1999     1702      5     2.973
## 5195   0034445661 3.652 2000     1708      3     2.536
## 5635   0035455653 4.472 2001     1708      3     3.166
## 7239   0036953758 4.202 2002     1708      1     2.983
## 8769   5444275208 3.799 2004     1705      3     2.594
## 10076 13944252629 4.211 2005     1705      3     2.891
## 10198 15844407150 4.159 2005     1708      2     2.839
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
```

```
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.3852 -0.5366  0.0222  0.5216  3.2192
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)          1.0842     0.0381   28.46  < 2e-16 ***
## FirstAuthorFemale1   0.0649     0.0296    2.19  0.02869 *
## LastAuthorFemale1   -0.0126     0.0300   -0.42  0.67562
## Year1997             0.0365     0.0540    0.68  0.49857
## Year1998             0.0426     0.0555    0.77  0.44209
## Year1999             0.0324     0.0560    0.58  0.56332
## Year2000             0.0318     0.0526    0.60  0.54564
## Year2001             0.2214     0.0555    3.99  6.8e-05 ***
## Year2002             0.1343     0.0549    2.45  0.01438 *
## Year2003             0.0792     0.0554    1.43  0.15296
## Year2004             0.1209     0.0569    2.12  0.03376 *
## Year2005             0.2362     0.0539    4.38  1.2e-05 ***
## Year2006             0.1991     0.0545    3.65  0.00026 ***
## Year2007             0.0194     0.0503    0.39  0.70007
## Year2008             0.0307     0.0492    0.62  0.53226
## Year2009             0.0376     0.0509    0.74  0.45965
## Year2010            -0.0414     0.0509   -0.81  0.41653
## Year2011            -0.1514     0.0521   -2.90  0.00370 **
## Year2012            -0.0257     0.0522   -0.49  0.62284
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.783
## Multiple R-squared:  0.0156, Adjusted R-squared:  0.0131
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  600 weights are ~= 1. The remaining 6617 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0532  0.8650  0.9500  0.9150  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.39e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                    psi           subsampling                   cov
```

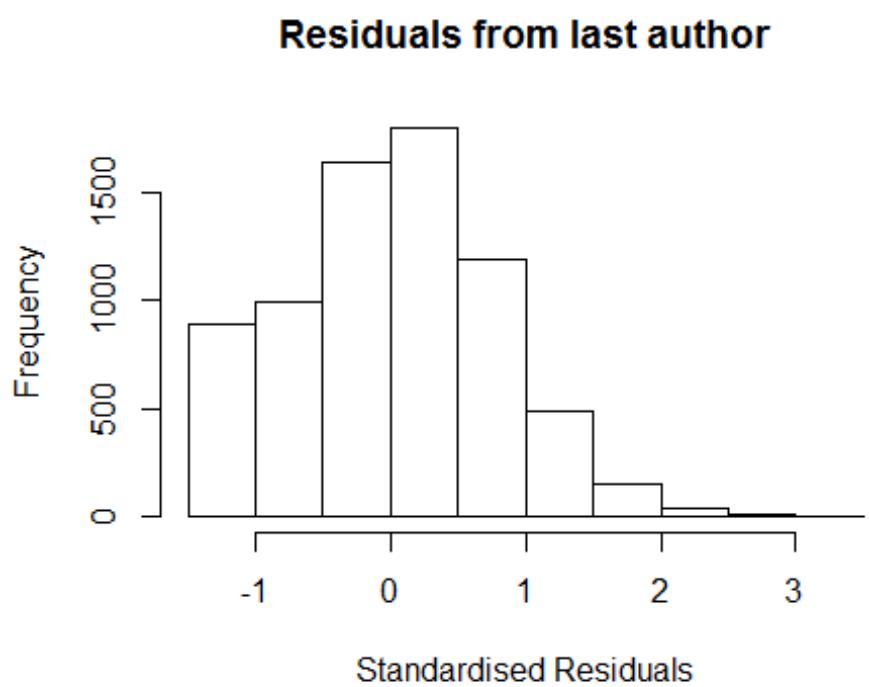
```
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.006  1           1.003
## Year              1.006 16           1.000
```




```
## [1] "List of  14 outliers with residuals above  2.5"
##          ScopusId  NLCS Year OneField Fields residuals
## 859    0030370740 3.688 1996     1705      4     2.604
## 1424   0031274649 3.708 1997     1705      4     2.587
## 1820   0031153727 4.130 1997     1708      3     3.009
## 2707   0002806618 3.648 1998     1702      5     2.521
## 2764  85045953475 4.346 1998     1705      2     3.219
## 3040   0032313923 3.699 1998     1705      4     2.572
## 3782   0032592464 3.678 1999     1705      3     2.561
## 4144   4243148480 4.090 1999     1702      5     2.973
## 5195   0034445661 3.652 2000     1708      3     2.536
## 5635   0035455653 4.472 2001     1708      3     3.166
## 7239   0036953758 4.202 2002     1708      1     2.983
## 8769   5444275208 3.799 2004     1705      3     2.594
## 10076 13944252629 4.211 2005     1705      3     2.891
## 10198 15844407150 4.159 2005     1708      2     2.839
##
## Call:
```

```
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.3808 -0.5372  0.0218  0.5218  3.2200
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)          1.0832     0.0380   28.50  < 2e-16 ***
## FirstAuthorFemale1   0.0615     0.0291    2.11  0.03457 *
## Year1997             0.0365     0.0540    0.68  0.49952
## Year1998             0.0428     0.0555    0.77  0.44054
## Year1999             0.0323     0.0560    0.58  0.56482
## Year2000             0.0318     0.0526    0.60  0.54552
## Year2001             0.2216     0.0556    3.99  6.7e-05 ***
## Year2002             0.1341     0.0549    2.44  0.01454 *
## Year2003             0.0793     0.0554    1.43  0.15232
## Year2004             0.1211     0.0569    2.13  0.03349 *
## Year2005             0.2360     0.0539    4.38  1.2e-05 ***
## Year2006             0.1994     0.0545    3.66  0.00026 ***
## Year2007             0.0197     0.0502    0.39  0.69513
## Year2008             0.0310     0.0492    0.63  0.52905
## Year2009             0.0370     0.0508    0.73  0.46629
## Year2010            -0.0414     0.0509   -0.81  0.41630
## Year2011            -0.1516     0.0521   -2.91  0.00364 **
## Year2012            -0.0257     0.0522   -0.49  0.62216
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.783
## Multiple R-squared:  0.0156, Adjusted R-squared:  0.0132
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  603 weights are ~= 1. The remaining 6614 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0529  0.8650  0.9500  0.9150  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.39e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
```

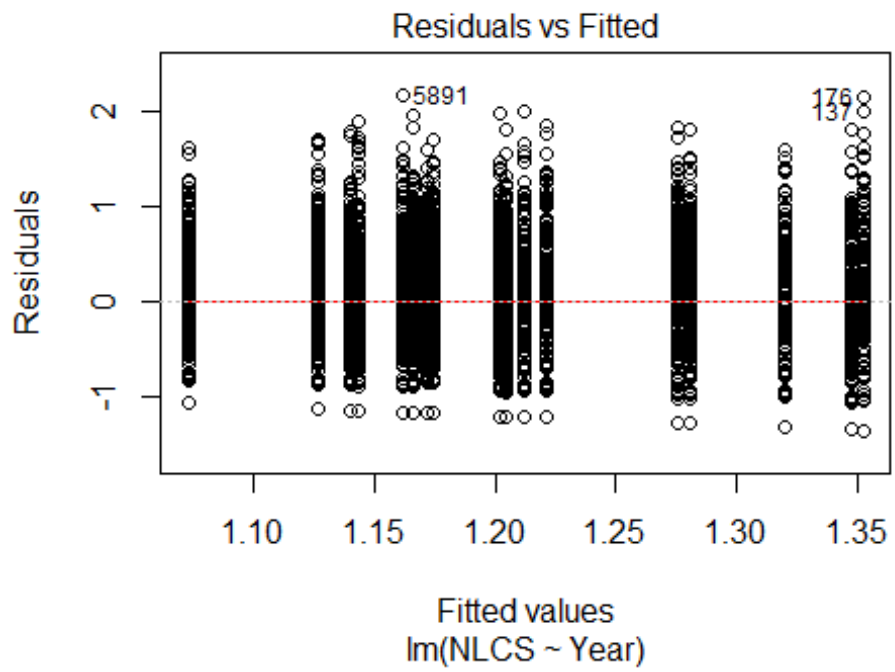
```
##                    psi           subsampling                    cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                   "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##                   GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.005  1           1.003
## Year             1.005 16           1.000
```



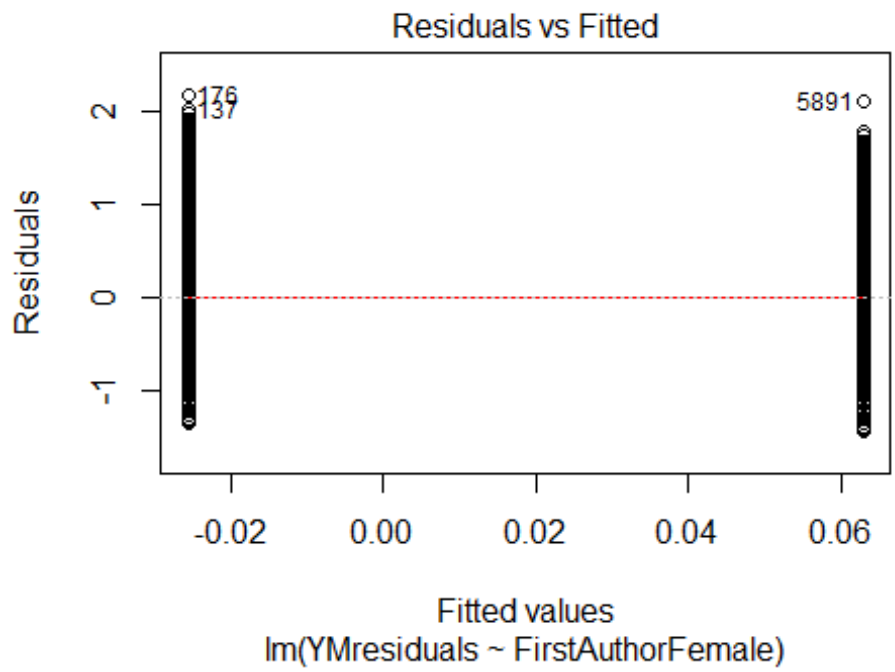

```
## [1] "List of  14 outliers with residuals above  2.5"
##          ScopusId  NLCS Year OneField Fields residuals
## 859    0030370740 3.688 1996     1705      4     2.604
## 1424   0031274649 3.708 1997     1705      4     2.587
## 1820   0031153727 4.130 1997     1708      3     3.009
## 2707   0002806618 3.648 1998     1702      5     2.521
## 2764  85045953475 4.346 1998     1705      2     3.219
## 3040   0032313923 3.699 1998     1705      4     2.572
## 3782   0032592464 3.678 1999     1705      3     2.561
## 4144   4243148480 4.090 1999     1702      5     2.973
## 5195   0034445661 3.652 2000     1708      3     2.536
## 5635   0035455653 4.472 2001     1708      3     3.166
## 7239   0036953758 4.202 2002     1708      1     2.983
## 8769   5444275208 3.799 2004     1705      3     2.594
## 10076 13944252629 4.211 2005     1705      3     2.891
## 10198 15844407150 4.159 2005     1708      2     2.839
##
```

```
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.3290 -0.5357  0.0246  0.5240  3.2132
##
## Coefficients:
##                   Estimate Std. Error t value Pr(>|t|)
## (Intercept)        1.08890    0.03804   28.62  < 2e-16 ***
## LastAuthorFemale1  0.00566    0.02931    0.19  0.84688
## Year1997           0.03583    0.05401    0.66  0.50702
## Year1998           0.04390    0.05549    0.79  0.42890
## Year1999           0.03396    0.05606    0.61  0.54469
## Year2000           0.03216    0.05267    0.61  0.54151
## Year2001           0.22343    0.05561    4.02  5.9e-05 ***
## Year2002           0.13543    0.05494    2.47  0.01372 *
## Year2003           0.08128    0.05560    1.46  0.14380
## Year2004           0.12115    0.05707    2.12  0.03380 *
## Year2005           0.23442    0.05388    4.35  1.4e-05 ***
## Year2006           0.20032    0.05451    3.67  0.00024 ***
## Year2007           0.02120    0.05037    0.42  0.67393
## Year2008           0.03253    0.04932    0.66  0.50952
## Year2009           0.03737    0.05088    0.73  0.46263
## Year2010          -0.04156    0.05100   -0.81  0.41520
## Year2011          -0.15007    0.05213   -2.88  0.00401 **
## Year2012          -0.02492    0.05222   -0.48  0.63327
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.784
## Multiple R-squared:  0.0149, Adjusted R-squared:  0.0126
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  606 weights are ~= 1. The remaining 6611 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.055   0.867   0.950   0.915   0.985   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.39e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
```

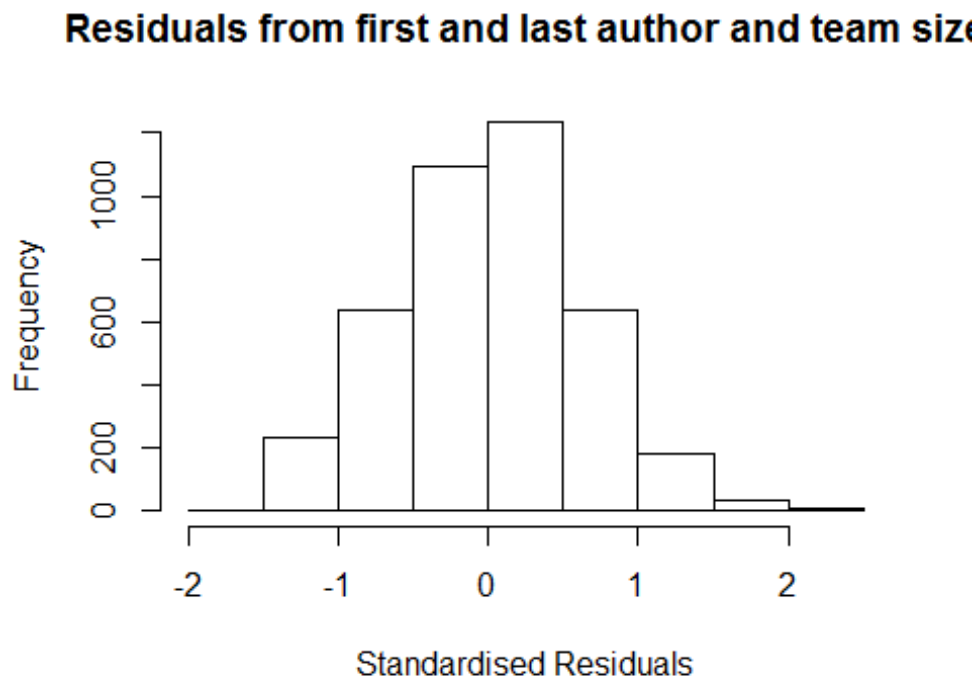
```
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Sample size for the above analysis:  7217"
## [1] ""
## [1] ""
## [1] "##################################"
## [1] "Analysis of AJSC 1709"
## [1] "##################################"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  221  219  275  285  346  315  345  276  357  361  439  499  529  540  539
## 2011 2012
##  569  566
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  172  142  195  170  214  213  250  197  284  245  322  333  367  395  383
## 2011 2012
##  401  390
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  150  122  175  148  193  193  215  168  248  213  280  287  314  339  326
## 2011 2012
##  357  328
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
##  Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 48, df = 16, p-value = 5e-05
```




```
##
##  Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 3.6, df = 1, p-value = 0.06
```

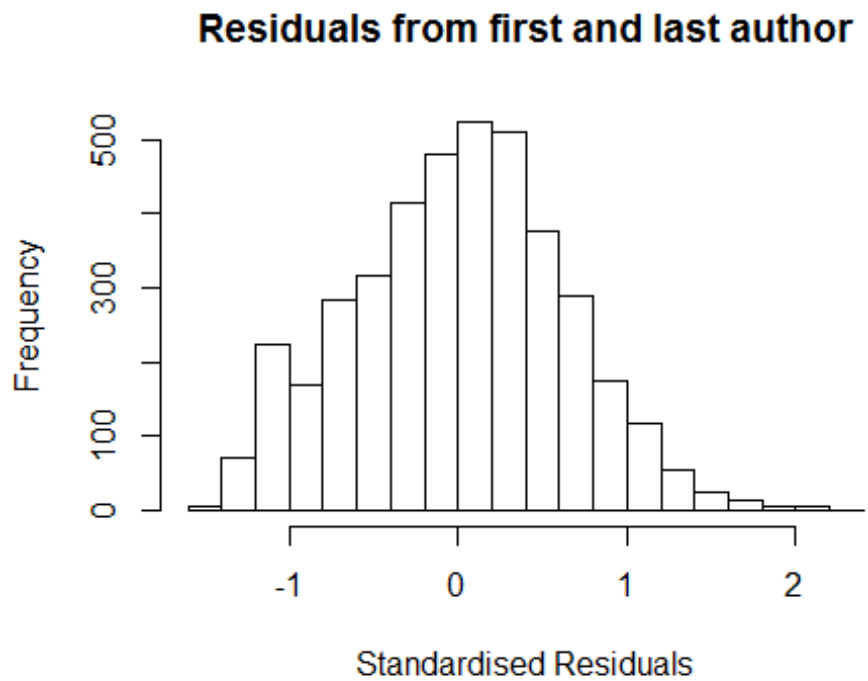



```
## [1] "Female first author team size 2018 geometric mean: 2.38241871757647"
## [1] "Male first author team size 2018 geometric mean: 2.39408278320984"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 8100, p-value = 0.9
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 2.13701530638322"
## [1] "Male last author team size 2018 geometric mean: 2.58199269539256"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 7000, p-value = 0.04
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##                     GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.159  1           1.077
## LastAuthorFemale  1.137  1           1.066
## UniqueAuthors     1.083  4           1.010
## Year              1.117 16           1.003
```



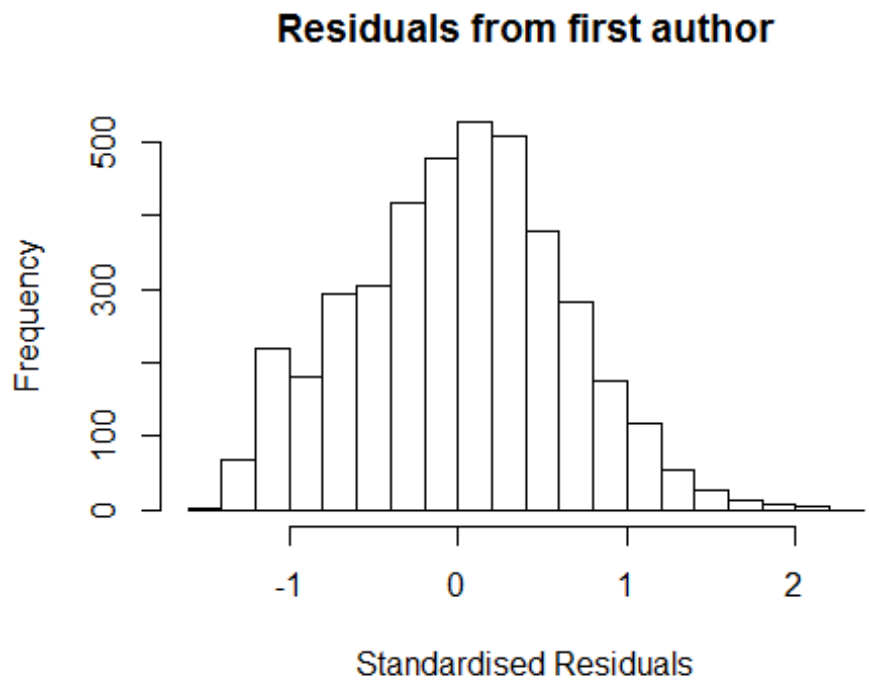

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
UniqueAuthors +
##     Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##     k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.5276 -0.4236  0.0256  0.4253  2.3413
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)          1.1467     0.0655   17.50  < 2e-16 ***
## FirstAuthorFemale1   0.0812     0.0233    3.48  0.00050 ***
## LastAuthorFemale1    0.0509     0.0242    2.10  0.03552 *
## UniqueAuthors2       0.1753     0.0258    6.79  1.3e-11 ***
## UniqueAuthors3       0.2448     0.0301    8.14  5.3e-16 ***
## UniqueAuthors4       0.2776     0.0349    7.95  2.3e-15 ***
## UniqueAuthors5       0.3407     0.0368    9.25  < 2e-16 ***
## Year1997             0.0550     0.0964    0.57  0.56847
## Year1998            -0.1116     0.0872   -1.28  0.20080
## Year1999            -0.1353     0.0834   -1.62  0.10489
```

```
## Year2000            -0.0531     0.0794   -0.67  0.50391
## Year2001             0.0296     0.0792    0.37  0.70836
## Year2002            -0.0565     0.0772   -0.73  0.46420
## Year2003            -0.2296     0.0799   -2.87  0.00409 **
## Year2004            -0.2825     0.0755   -3.74  0.00018 ***
## Year2005            -0.1521     0.0788   -1.93  0.05371 .
## Year2006            -0.2138     0.0751   -2.85  0.00443 **
## Year2007            -0.1164     0.0736   -1.58  0.11401
## Year2008            -0.2170     0.0704   -3.08  0.00206 **
## Year2009            -0.2220     0.0713   -3.11  0.00186 **
## Year2010            -0.1960     0.0714   -2.74  0.00610 **
## Year2011            -0.1952     0.0711   -2.74  0.00608 **
## Year2012            -0.2099     0.0723   -2.90  0.00369 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.631
## Multiple R-squared:  0.0538, Adjusted R-squared:  0.0486
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  325 weights are ~= 1. The remaining 3731 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.138   0.870   0.951   0.911   0.986   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          2.47e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.152  1           1.073
## LastAuthorFemale  1.135  1           1.065
## Year              1.038 16           1.001
```



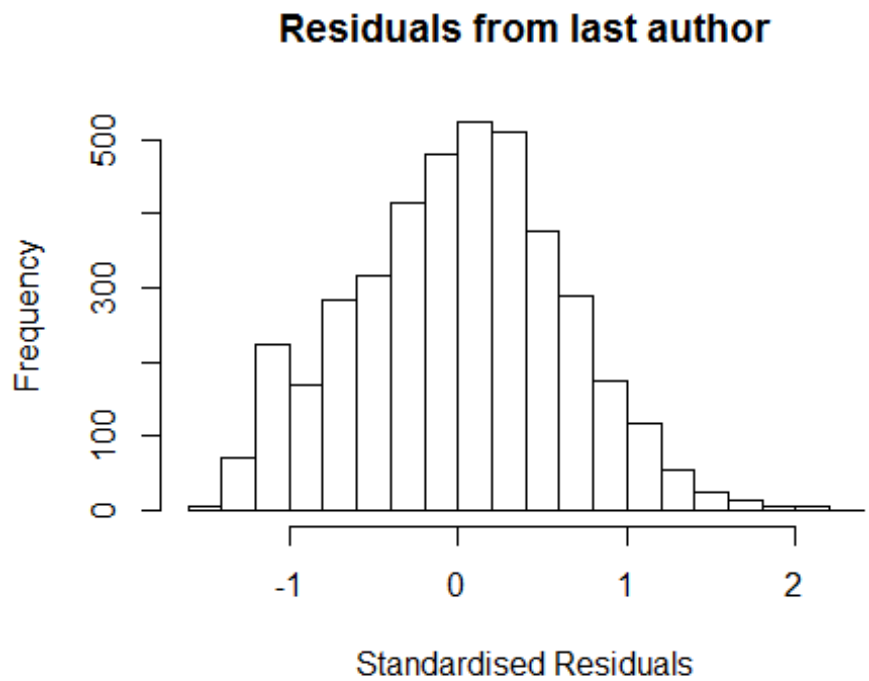

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.4271 -0.4286  0.0279  0.4214  2.2058
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)           1.2822     0.0644   19.92  < 2e-16 ***
## FirstAuthorFemale1    0.0918     0.0237    3.87  0.00011 ***
## LastAuthorFemale1     0.0389     0.0246    1.58  0.11418
## Year1997              0.0365     0.0966    0.38  0.70576
## Year1998             -0.1092     0.0872   -1.25  0.21029
## Year1999             -0.1378     0.0847   -1.63  0.10382
## Year2000             -0.0535     0.0813   -0.66  0.51040
## Year2001              0.0143     0.0797    0.18  0.85766
## Year2002             -0.0425     0.0775   -0.55  0.58296
## Year2003             -0.2144     0.0814   -2.63  0.00851 **
## Year2004             -0.2731     0.0768   -3.55  0.00038 ***
## Year2005             -0.1310     0.0802   -1.63  0.10248
```

```
## Year2006            -0.2261     0.0767   -2.95  0.00322 **
## Year2007            -0.1091     0.0749   -1.46  0.14519
## Year2008            -0.1984     0.0715   -2.77  0.00558 **
## Year2009            -0.1856     0.0721   -2.57  0.01010 *
## Year2010            -0.1711     0.0724   -2.36  0.01823 *
## Year2011            -0.1617     0.0722   -2.24  0.02527 *
## Year2012            -0.1833     0.0738   -2.48  0.01302 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.639
## Multiple R-squared:  0.0207, Adjusted R-squared:  0.0163
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  321 weights are ~= 1. The remaining 3735 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.208   0.869   0.951   0.910   0.986   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          2.47e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.026  1           1.013
## Year              1.026 16           1.001
```




```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##    Min     1Q Median     3Q    Max
## -1.429 -0.426  0.028  0.424  2.204
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)           1.2842     0.0642   20.00  < 2e-16 ***
## FirstAuthorFemale1    0.1078     0.0224    4.81  1.6e-06 ***
## Year1997              0.0367     0.0964    0.38  0.70346
## Year1998             -0.1071     0.0871   -1.23  0.21888
## Year1999             -0.1347     0.0846   -1.59  0.11126
## Year2000             -0.0512     0.0813   -0.63  0.52881
## Year2001              0.0167     0.0796    0.21  0.83387
## Year2002             -0.0394     0.0774   -0.51  0.61097
## Year2003             -0.2085     0.0813   -2.56  0.01037 *
## Year2004             -0.2692     0.0767   -3.51  0.00045 ***
## Year2005             -0.1282     0.0801   -1.60  0.10942
## Year2006             -0.2218     0.0766   -2.90  0.00379 **
```

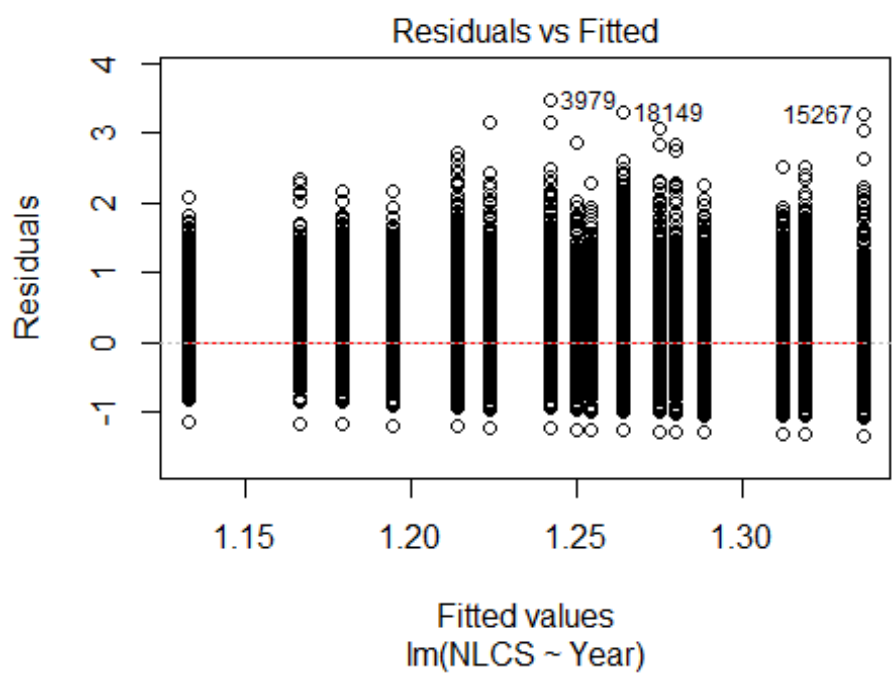
```
## Year2007            -0.1053     0.0748   -1.41  0.15925
## Year2008            -0.1952     0.0714   -2.73  0.00631 **
## Year2009            -0.1824     0.0721   -2.53  0.01140 *
## Year2010            -0.1662     0.0723   -2.30  0.02158 *
## Year2011            -0.1565     0.0721   -2.17  0.02999 *
## Year2012            -0.1772     0.0736   -2.41  0.01607 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.639
## Multiple R-squared:  0.02,   Adjusted R-squared:  0.0159
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  328 weights are ~= 1. The remaining 3728 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.210   0.870   0.951   0.910   0.986   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          2.47e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##                   GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.012  1           1.006
## Year             1.012 16           1.000
```



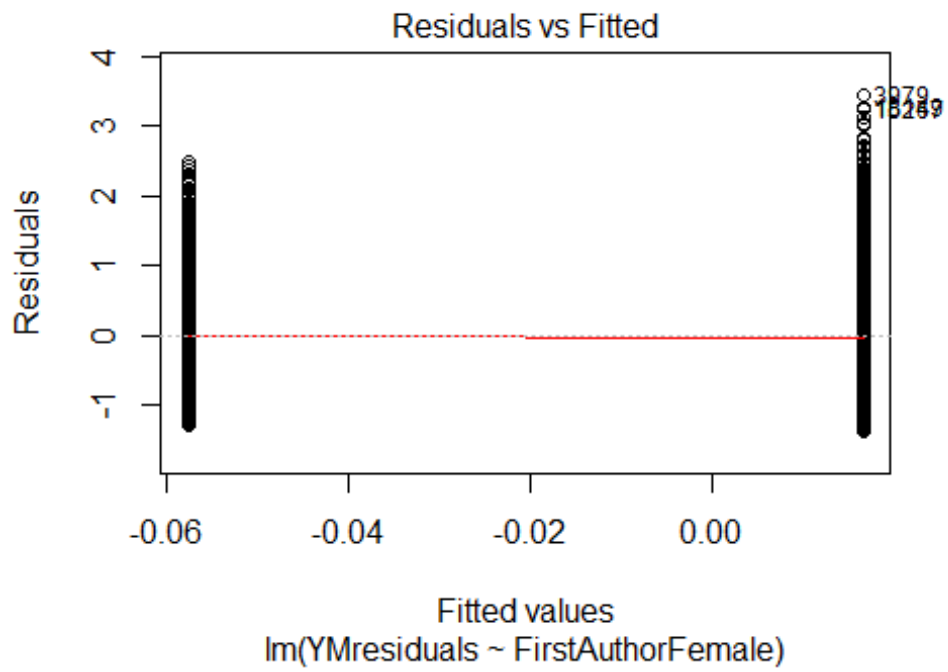

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.3922 -0.4291  0.0276  0.4223  2.1901
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)          1.2979     0.0646   20.09  < 2e-16 ***
## LastAuthorFemale1    0.0791     0.0232    3.41  0.00067 ***
## Year1997             0.0299     0.0965    0.31  0.75702
## Year1998            -0.1115     0.0875   -1.27  0.20252
## Year1999            -0.1442     0.0850   -1.69  0.09016 .
## Year2000            -0.0492     0.0814   -0.60  0.54567
## Year2001             0.0152     0.0800    0.19  0.84956
## Year2002            -0.0327     0.0778   -0.42  0.67436
## Year2003            -0.2199     0.0816   -2.70  0.00705 **
## Year2004            -0.2696     0.0771   -3.49  0.00048 ***
## Year2005            -0.1293     0.0805   -1.61  0.10811
## Year2006            -0.2283     0.0772   -2.96  0.00310 **
```

```
## Year2007           -0.1075     0.0753   -1.43  0.15363
## Year2008           -0.1997     0.0720   -2.77  0.00559 **
## Year2009           -0.1800     0.0724   -2.49  0.01297 *
## Year2010           -0.1738     0.0728   -2.39  0.01706 *
## Year2011           -0.1619     0.0727   -2.23  0.02597 *
## Year2012           -0.1825     0.0742   -2.46  0.01397 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.64
## Multiple R-squared:  0.0174, Adjusted R-squared:  0.0132
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  333 weights are ~= 1. The remaining 3723 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.217   0.868   0.951   0.910   0.986   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          2.47e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Sample size for the above analysis:  4056"
## [1] ""
## [1] ""
## [1] "####################################"
## [1] "Analysis of AJSC 1710"
## [1] "####################################"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 1309 1234 1213 1228 1338 1351 1303 1051 1113 1269 1447 1461 1481 1567 1421
## 2011 2012
## 1444 1566
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  903  695  838  849  917  842  888  732  741  822  882  932  879 1015  907
## 2011 2012
```

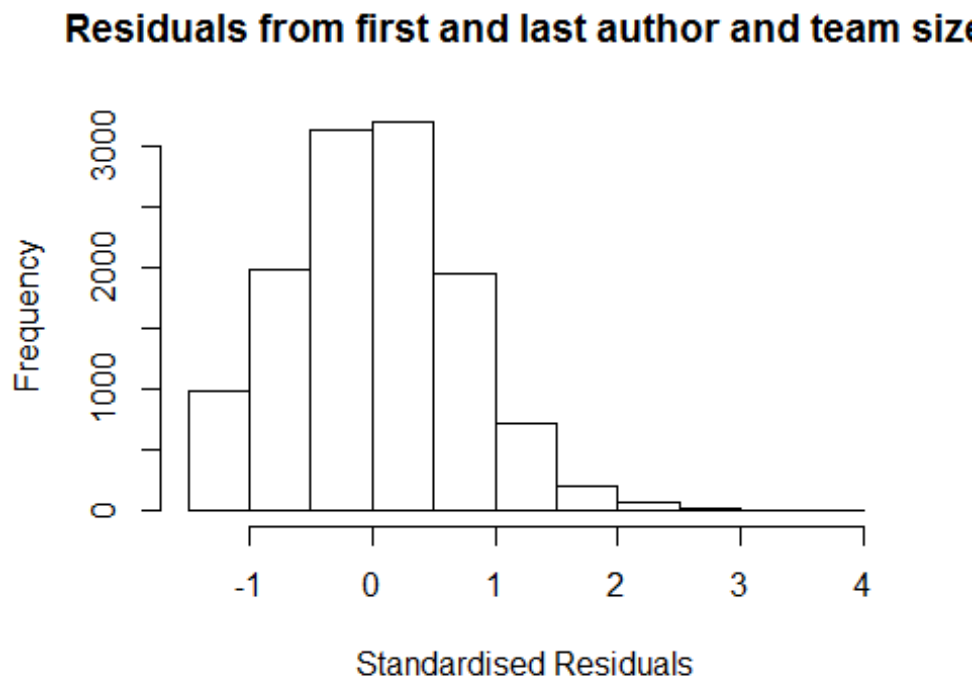
```
##  905  994
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  787  594  739  736  784  716  756  614  618  678  693  763  715  832  727
## 2011 2012
##  733  799
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
##  Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 200, df = 16, p-value <2e-16
```




```
##
##  Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 3.3, df = 1, p-value = 0.07
```



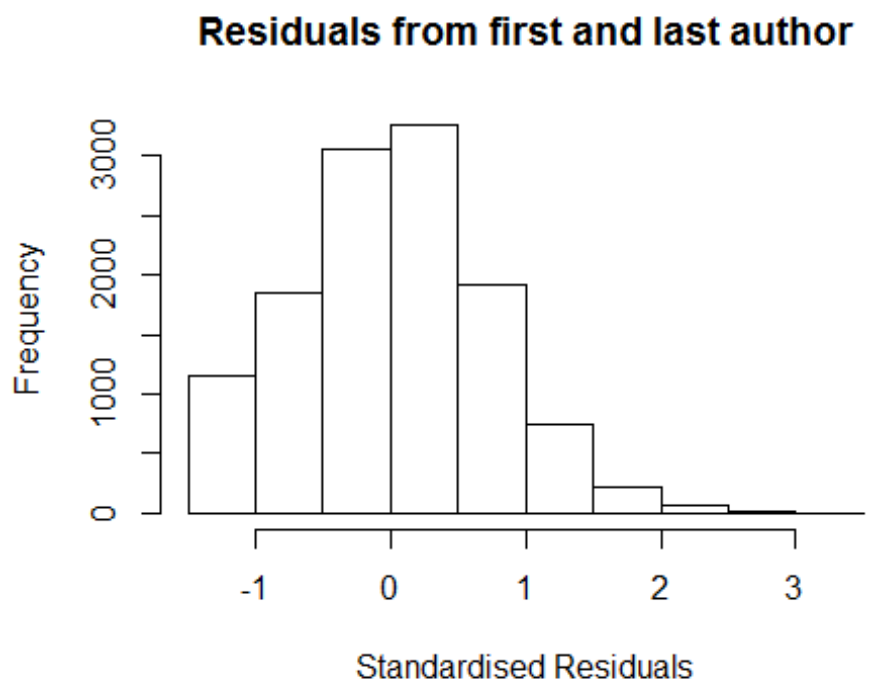

```
## [1] "Female first author team size 2018 geometric mean: 2.13039558041416"
## [1] "Male first author team size 2018 geometric mean: 2.57138902924946"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 32000, p-value = 0.002
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 2.02246609411603"
## [1] "Male last author team size 2018 geometric mean: 2.63029021008777"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 30000, p-value = 4e-06
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.282  1           1.132
## LastAuthorFemale  1.292  1           1.137
## UniqueAuthors     1.062  4           1.008
## Year              1.069 16           1.002
```




```
## [1] "List of  14 outliers with residuals above  2.5"
##          ScopusId  NLCS Year OneField Fields residuals
## 512    0030240869 3.699 1996     1710      2     2.704
## 2382   0031246031 3.912 1997     1705      3     2.559
## 2536  21744462998 3.531 1997     1705      2     2.551
## 2544  27144463192 3.637 1997     1705      2     2.657
## 2555   4344578226 3.483 1997     1705      2     2.503
## 3979  27144489164 4.703 1998     1705      3     3.664
## 4505   9744277690 3.588 1999     1705      3     2.529
## 5542   4243148480 4.090 1999     1702      5     3.031
## 9006   0036811662 3.649 2002     1702      4     2.593
## 14328 33947416035 3.963 2006     1706      3     2.606
## 15267 33645712892 4.593 2006     1706      3     3.481
## 15516 31744440684 4.357 2006     1706      3     2.930
## 15983 64649083745 4.114 2007     1706      3     2.884
## 18149 48449095896 4.549 2008     1701      2     3.299
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
UniqueAuthors +
##     Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##     k.max = 1000))
##  \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.47751 -0.47840  0.00553  0.48123  3.66448
```

```
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.06785    0.03282   32.53  < 2e-16 ***
## FirstAuthorFemale1 -0.03812    0.01754   -2.17  0.02980 *
## LastAuthorFemale1  -0.03487    0.01826   -1.91  0.05621 .
## UniqueAuthors2      0.24586    0.01625   15.13  < 2e-16 ***
## UniqueAuthors3      0.31547    0.01855   17.00  < 2e-16 ***
## UniqueAuthors4      0.36592    0.02535   14.43  < 2e-16 ***
## UniqueAuthors5      0.37293    0.02870   13.00  < 2e-16 ***
## Year1997           -0.08790    0.04621   -1.90  0.05717 .
## Year1998           -0.02933    0.04274   -0.69  0.49253
## Year1999           -0.00888    0.04271   -0.21  0.83533
## Year2000           -0.03143    0.04210   -0.75  0.45538
## Year2001            0.02605    0.04233    0.62  0.53829
## Year2002            0.06105    0.04101    1.49  0.13661
## Year2003            0.01440    0.04171    0.35  0.72985
## Year2004           -0.02197    0.04214   -0.52  0.60221
## Year2005            0.03463    0.04233    0.82  0.41324
## Year2006            0.04375    0.04142    1.06  0.29092
## Year2007           -0.04896    0.03919   -1.25  0.21158
## Year2008           -0.02842    0.04011   -0.71  0.47872
## Year2009           -0.08434    0.03900   -2.16  0.03059 *
## Year2010           -0.13533    0.04013   -3.37  0.00075 ***
## Year2011           -0.09317    0.04008   -2.32  0.02012 *
## Year2012           -0.15257    0.04091   -3.73  0.00019 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.711
## Multiple R-squared:  0.0471, Adjusted R-squared:  0.0453
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  2 observations c(2013,7465) are outliers with |weight| = 0 ( < 8.1e-06);
##  1052 weights are ~= 1. The remaining 11230 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0003  0.8700  0.9500  0.9090  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          8.14e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                    psi           subsampling                   cov
```

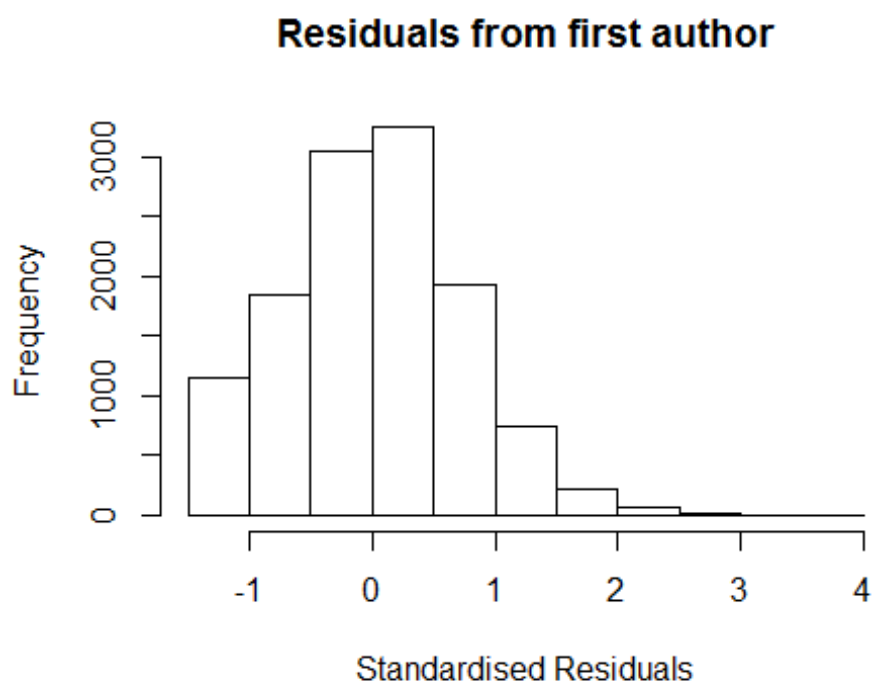
```
##             "bisquare"          "nonsingular"          ".vcov.avar1"
## compute.outlier.stats
##                   "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##                     GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.268  1           1.126
## LastAuthorFemale  1.273  1           1.128
## Year              1.013 16           1.000
```




```
## [1] "List of  15 outliers with residuals above  2.5"
##          ScopusId  NLCS Year OneField Fields residuals
## 512    0030240869 3.699 1996     1710      2     2.573
## 1906   0031274649 3.708 1997     1705      4     2.550
## 2382   0031246031 3.912 1997     1705      3     2.754
## 2532  21744433274 3.850 1997     1705      2     2.692
## 3914   0032447095 3.733 1998     1710      3     2.528
## 3979  27144489164 4.703 1998     1705      3     3.498
## 5542   4243148480 4.090 1999     1702      5     2.868
## 9354   0036611602 3.834 2002     1706      3     2.517
## 12767 29144439194 3.812 2005     1706      3     2.509
## 14328 33947416035 3.963 2006     1706      3     2.631
## 15267 33645712892 4.593 2006     1706      3     3.261
## 15516 31744440684 4.357 2006     1706      3     3.025
## 15983 64649083745 4.114 2007     1706      3     2.941
## 18149 48449095896 4.549 2008     1701      2     3.366
```

```
## 18197 48849113556 3.758 2008     1706      3     2.515
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.3324 -0.4896  0.0112  0.4826  3.4984
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.22670    0.03213   38.18   <2e-16 ***
## FirstAuthorFemale1 -0.04148    0.01774   -2.34   0.0194 *
## LastAuthorFemale1  -0.05923    0.01845   -3.21   0.0013 **
## Year1997           -0.06895    0.04695   -1.47   0.1419
## Year1998           -0.02208    0.04338   -0.51   0.6107
## Year1999           -0.00449    0.04342   -0.10   0.9177
## Year2000           -0.00369    0.04286   -0.09   0.9314
## Year2001            0.04284    0.04306    0.99   0.3198
## Year2002            0.09013    0.04169    2.16   0.0307 *
## Year2003            0.05096    0.04230    1.20   0.2283
## Year2004            0.01581    0.04324    0.37   0.7147
## Year2005            0.07604    0.04313    1.76   0.0779 .
## Year2006            0.10567    0.04178    2.53   0.0114 *
## Year2007            0.00509    0.03978    0.13   0.8981
## Year2008            0.01594    0.04100    0.39   0.6974
## Year2009           -0.02656    0.03948   -0.67   0.5011
## Year2010           -0.07831    0.04068   -1.93   0.0542 .
## Year2011           -0.02623    0.04024   -0.65   0.5145
## Year2012           -0.08479    0.04157   -2.04   0.0414 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.721
## Multiple R-squared:  0.00823,    Adjusted R-squared:  0.00678
## Convergence in 10 IRWLS iterations
##
## Robustness weights:
##  observation 2013 is an outlier with |weight| = 0 ( < 8.1e-06);
##  1103 weights are ~= 1. The remaining 11180 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0001  0.8690  0.9500  0.9080  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          8.14e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
```

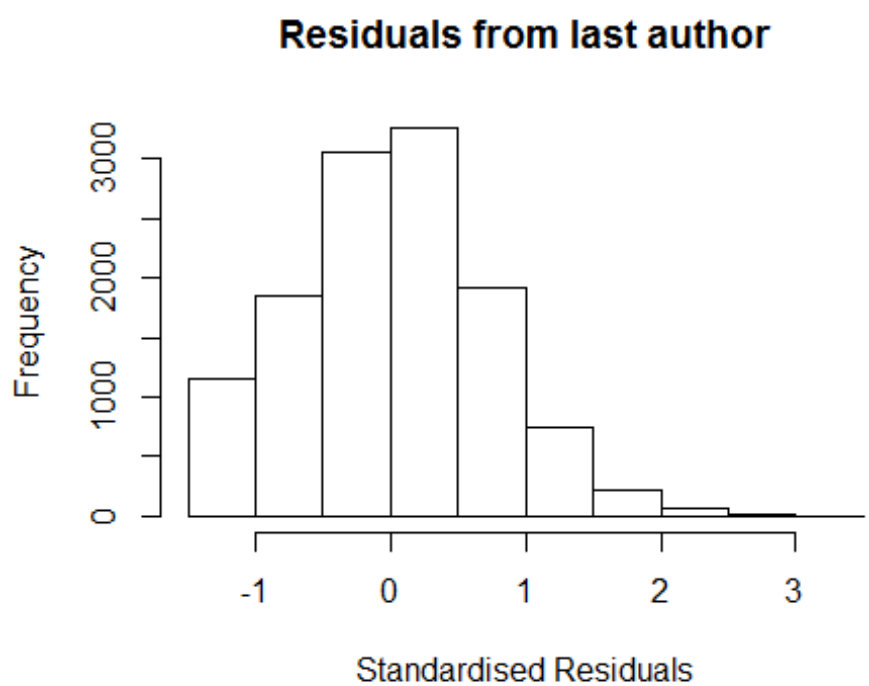
```
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.004  1           1.002
## Year              1.004 16           1.000
```




```
## [1] "List of  15 outliers with residuals above  2.5"
##          ScopusId  NLCS Year OneField Fields residuals
## 512    0030240869 3.699 1996     1710      2     2.573
## 1906   0031274649 3.708 1997     1705      4     2.550
## 2382   0031246031 3.912 1997     1705      3     2.754
## 2532  21744433274 3.850 1997     1705      2     2.692
## 3914   0032447095 3.733 1998     1710      3     2.528
## 3979  27144489164 4.703 1998     1705      3     3.498
## 5542   4243148480 4.090 1999     1702      5     2.868
## 9354   0036611602 3.834 2002     1706      3     2.517
## 12767 29144439194 3.812 2005     1706      3     2.509
## 14328 33947416035 3.963 2006     1706      3     2.631
```

```
## 15267 33645712892 4.593 2006     1706      3     3.261
## 15516 31744440684 4.357 2006     1706      3     3.025
## 15983 64649083745 4.114 2007     1706      3     2.941
## 18149 48449095896 4.549 2008     1701      2     3.366
## 18197 48849113556 3.758 2008     1706      3     2.515
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.3273 -0.4902  0.0112  0.4837  3.5042
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.22235    0.03212   38.05  < 2e-16 ***
## FirstAuthorFemale1 -0.06953    0.01586   -4.38  1.2e-05 ***
## Year1997           -0.07030    0.04698   -1.50    0.135
## Year1998           -0.02356    0.04345   -0.54    0.588
## Year1999           -0.00727    0.04344   -0.17    0.867
## Year2000           -0.00493    0.04287   -0.12    0.908
## Year2001            0.04265    0.04306    0.99    0.322
## Year2002            0.08875    0.04173    2.13    0.033 *
## Year2003            0.05035    0.04229    1.19    0.234
## Year2004            0.01403    0.04321    0.32    0.745
## Year2005            0.07323    0.04316    1.70    0.090 .
## Year2006            0.10494    0.04176    2.51    0.012 *
## Year2007            0.00350    0.03976    0.09    0.930
## Year2008            0.01438    0.04098    0.35    0.726
## Year2009           -0.02990    0.03944   -0.76    0.448
## Year2010           -0.08257    0.04064   -2.03    0.042 *
## Year2011           -0.02929    0.04022   -0.73    0.466
## Year2012           -0.08800    0.04160   -2.12    0.034 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.721
## Multiple R-squared:  0.00738,    Adjusted R-squared:  0.00601
## Convergence in 10 IRWLS iterations
##
## Robustness weights:
##  observation 2013 is an outlier with |weight| = 0 ( < 8.1e-06);
##  1072 weights are ~= 1. The remaining 11211 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0015  0.8700  0.9500  0.9080  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
```

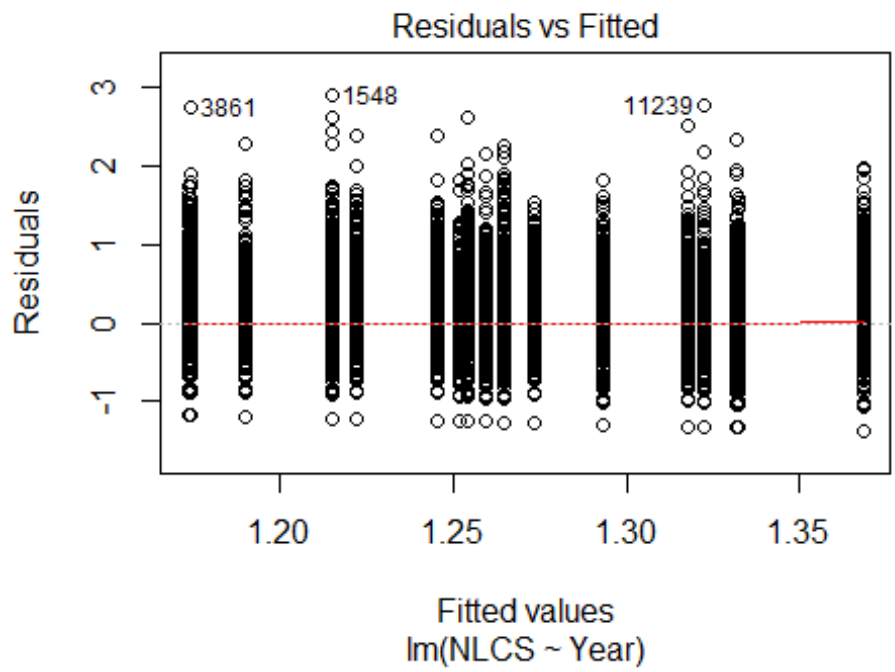
```
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          8.14e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##                   GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.008  1           1.004
## Year             1.008 16           1.000
```



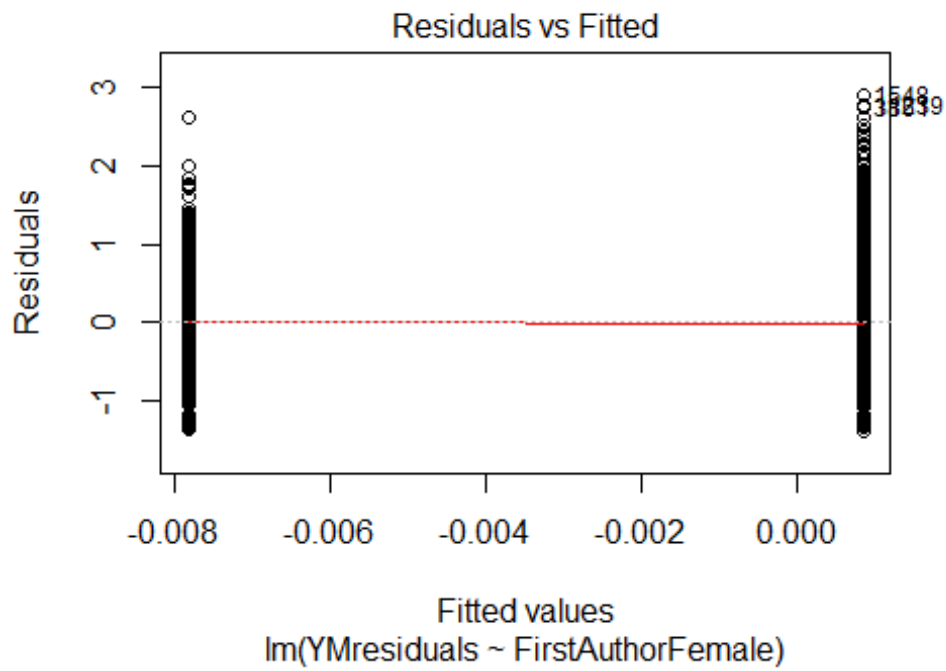

```
## [1] "List of  15 outliers with residuals above  2.5"
##          ScopusId  NLCS Year OneField Fields residuals
## 512    0030240869 3.699 1996     1710      2     2.573
## 1906   0031274649 3.708 1997     1705      4     2.550
## 2382   0031246031 3.912 1997     1705      3     2.754
## 2532  21744433274 3.850 1997     1705      2     2.692
## 3914   0032447095 3.733 1998     1710      3     2.528
## 3979  27144489164 4.703 1998     1705      3     3.498
## 5542   4243148480 4.090 1999     1702      5     2.868
```

```
## 9354   0036611602 3.834 2002     1706      3     2.517
## 12767 29144439194 3.812 2005     1706      3     2.509
## 14328 33947416035 3.963 2006     1706      3     2.631
## 15267 33645712892 4.593 2006     1706      3     3.261
## 15516 31744440684 4.357 2006     1706      3     3.025
## 15983 64649083745 4.114 2007     1706      3     2.941
## 18149 48449095896 4.549 2008     1701      2     3.366
## 18197 48849113556 3.758 2008     1706      3     2.515
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##    Min     1Q Median     3Q    Max
## -1.327 -0.489  0.011  0.482  3.502
##
## Coefficients:
##                   Estimate Std. Error t value Pr(>|t|)
## (Intercept)        1.22138    0.03203   38.13  < 2e-16 ***
## LastAuthorFemale1 -0.07999    0.01648   -4.86  1.2e-06 ***
## Year1997          -0.06728    0.04691   -1.43    0.151
## Year1998          -0.02032    0.04335   -0.47    0.639
## Year1999          -0.00340    0.04344   -0.08    0.938
## Year2000          -0.00325    0.04288   -0.08    0.940
## Year2001           0.04275    0.04308    0.99    0.321
## Year2002           0.08956    0.04169    2.15    0.032 *
## Year2003           0.05095    0.04227    1.21    0.228
## Year2004           0.01539    0.04328    0.36    0.722
## Year2005           0.07620    0.04316    1.77    0.078 .
## Year2006           0.10555    0.04180    2.52    0.012 *
## Year2007           0.00520    0.03979    0.13    0.896
## Year2008           0.01643    0.04102    0.40    0.689
## Year2009          -0.02633    0.03952   -0.67    0.505
## Year2010          -0.07874    0.04068   -1.94    0.053 .
## Year2011          -0.02706    0.04026   -0.67    0.502
## Year2012          -0.08542    0.04159   -2.05    0.040 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.721
## Multiple R-squared:  0.00777,    Adjusted R-squared:  0.00639
## Convergence in 10 IRWLS iterations
##
## Robustness weights:
##  2 observations c(2013,8750) are outliers with |weight| = 0 ( < 8.1e-06);
##  1078 weights are ~= 1. The remaining 11204 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0042  0.8700  0.9500  0.9080  0.9850  0.9990
```

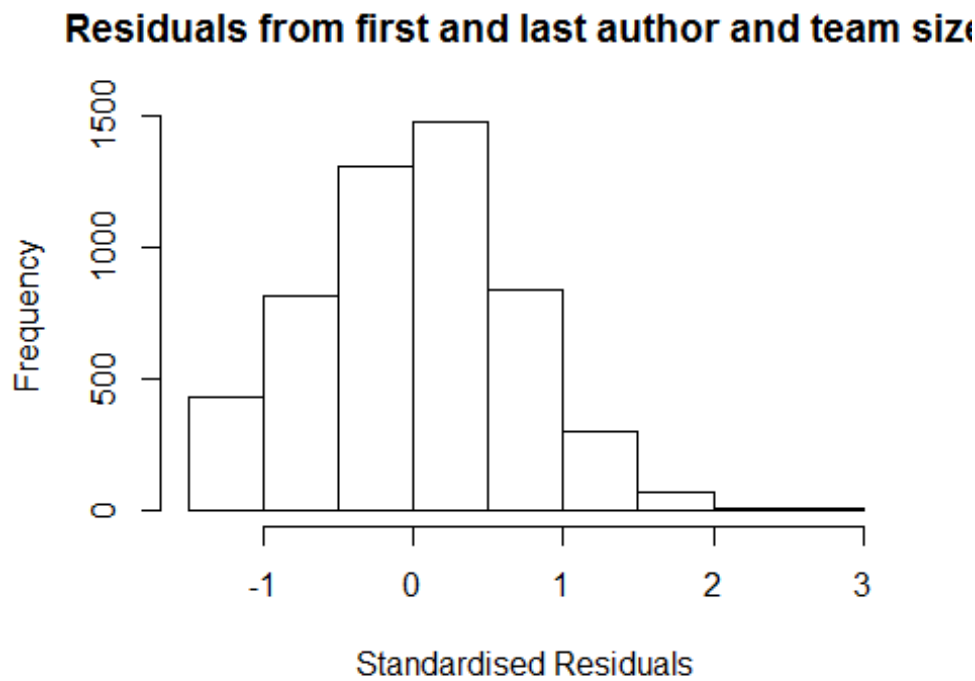
```
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          8.14e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Sample size for the above analysis:  12284"
## [1] ""
## [1] ""
## [1] "####################################"
## [1] "Analysis of AJSC 1711"
## [1] "####################################"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  798  860  824  671  811  857  888  533  560  619  733  831  752  722  777
## 2011 2012
##  704  759
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  483  474  485  389  487  439  492  300  304  331  373  423  382  345  389
## 2011 2012
##  364  385
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  386  381  378  297  391  331  366  215  231  250  277  325  284  274  305
## 2011 2012
##  290  283
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
##  Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 120, df = 16, p-value <2e-16
```




```
##
##  Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 0.046, df = 1, p-value = 0.8
```



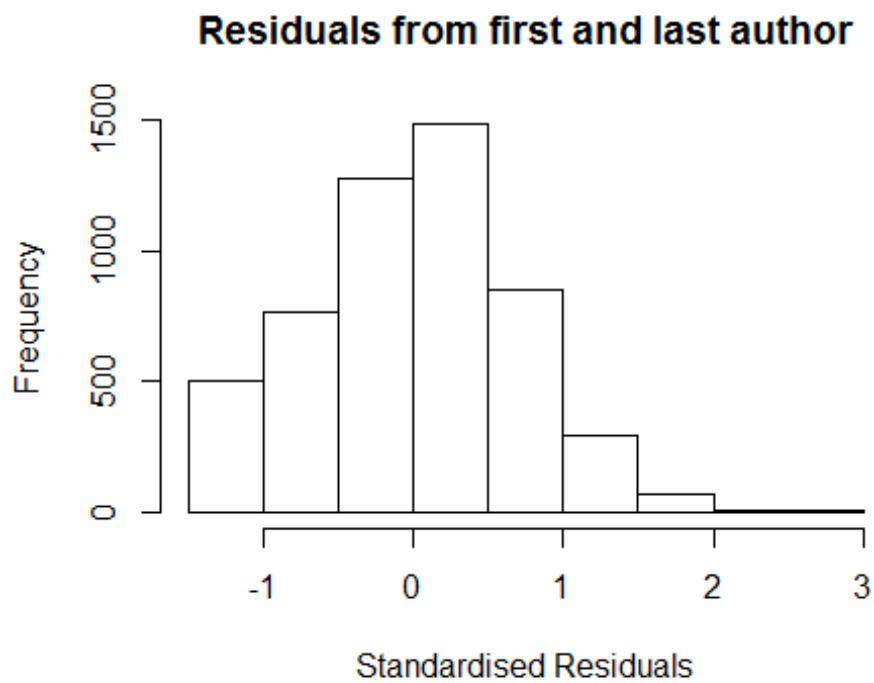

```
## [1] "Female first author team size 2018 geometric mean: 3.1942414839668"
## [1] "Male first author team size 2018 geometric mean: 2.63306409750109"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 7000, p-value = 0.02
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 2.77534465208727"
## [1] "Male last author team size 2018 geometric mean: 2.71052231922578"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 3600, p-value = 0.6
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.038  1           1.019
## LastAuthorFemale  1.033  1           1.016
## UniqueAuthors     1.137  4           1.016
## Year              1.143 16           1.004
```




```
## [1] "List of  9 outliers with residuals above  2.5"
##          ScopusId  NLCS Year OneField Fields residuals
## 782    0030287048 3.862 1996     1711      3     2.840
## 930    0031102203 3.829 1997     1711      2     2.585
## 979    0031165386 3.657 1997     1711      2     2.663
## 1548   0031078854 4.109 1997     1711      2     2.825
## 3861   0033884858 3.916 2000     1711      2     2.674
## 8915  33646023117 3.828 2006     1702      4     2.743
## 10336 85032751965 3.652 2007     1711      3     2.534
## 11239 85032750937 4.080 2008     1711      3     2.759
## 12651 77950369345 3.468 2010     1702      4     2.501
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
UniqueAuthors +
##     Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##     k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.4256 -0.4703  0.0235  0.4660  2.8403
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)          1.02165    0.04643   22.01  < 2e-16 ***
## FirstAuthorFemale1 -0.00943    0.03159   -0.30  0.76537
```

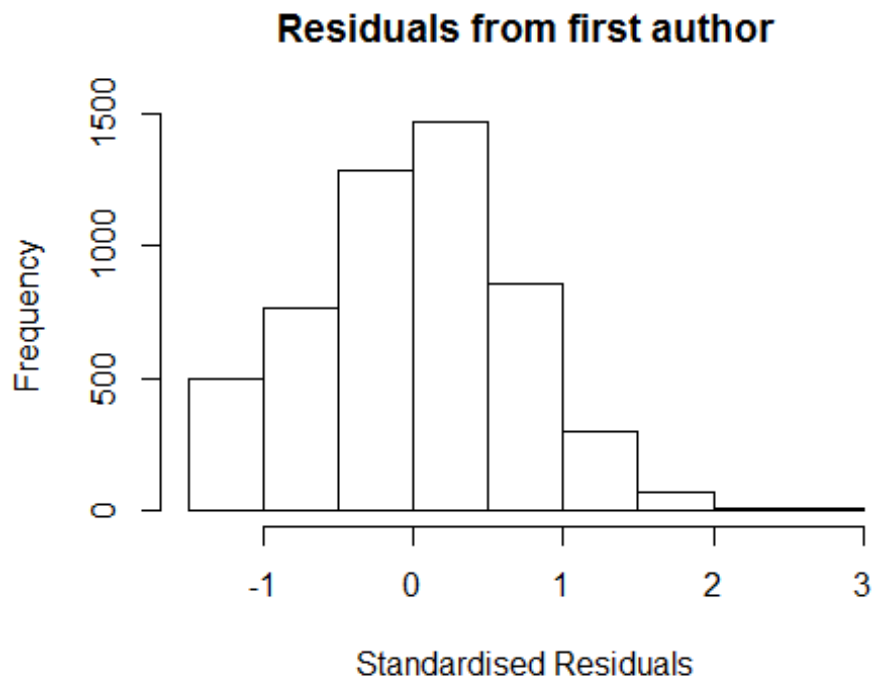
```
## LastAuthorFemale1  -0.11750    0.03442   -3.41  0.00064 ***
## UniqueAuthors2      0.25901    0.02673    9.69  < 2e-16 ***
## UniqueAuthors3      0.29486    0.02976    9.91  < 2e-16 ***
## UniqueAuthors4      0.28927    0.03991    7.25  4.8e-13 ***
## UniqueAuthors5      0.28383    0.04702    6.04  1.7e-09 ***
## Year1997           -0.02725    0.06322   -0.43  0.66653
## Year1998            0.03312    0.05804    0.57  0.56826
## Year1999           -0.00764    0.05928   -0.13  0.89750
## Year2000           -0.07496    0.05855   -1.28  0.20055
## Year2001            0.10912    0.06004    1.82  0.06921 .
## Year2002           -0.03361    0.05800   -0.58  0.56233
## Year2003            0.05258    0.05801    0.91  0.36476
## Year2004            0.09263    0.06064    1.53  0.12670
## Year2005            0.03771    0.06037    0.62  0.53221
## Year2006            0.06287    0.05793    1.09  0.27780
## Year2007            0.09624    0.05389    1.79  0.07418 .
## Year2008            0.04038    0.05661    0.71  0.47574
## Year2009            0.02592    0.05595    0.46  0.64317
## Year2010           -0.05454    0.05539   -0.98  0.32486
## Year2011           -0.02342    0.06007   -0.39  0.69658
## Year2012            0.02452    0.05871    0.42  0.67617
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.687
## Multiple R-squared:  0.0367, Adjusted R-squared:  0.0327
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
##  437 weights are ~= 1. The remaining 4827 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0491  0.8600  0.9490  0.9080  0.9860  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.90e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
```

```
##                     GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.022  1           1.011
## LastAuthorFemale  1.020  1           1.010
## Year              1.018 16           1.001
```



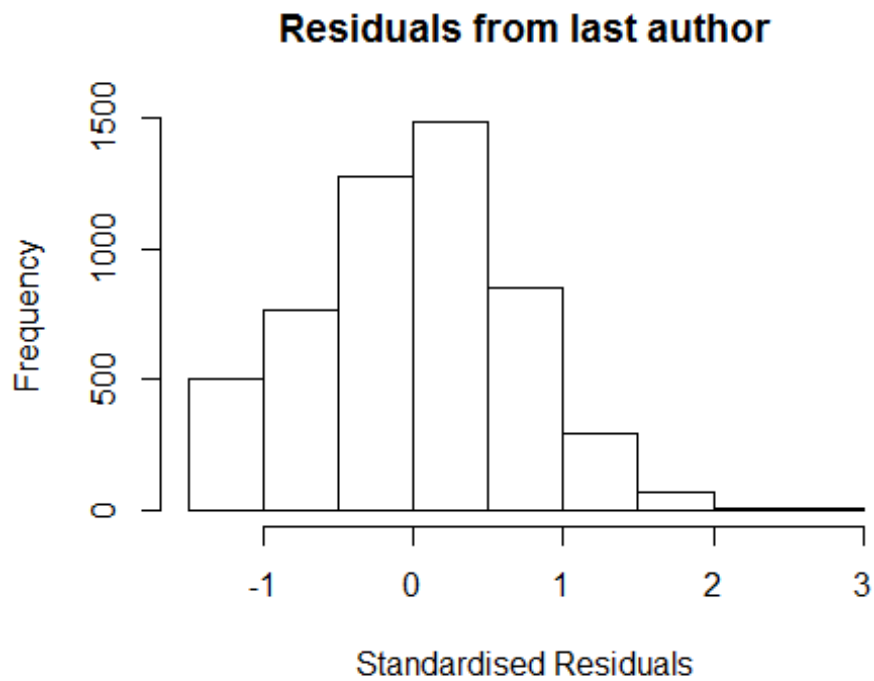

```
## [1] "List of  6 outliers with residuals above  2.5"
##          ScopusId  NLCS Year OneField Fields residuals
## 782    0030287048 3.862 1996     1711      3     2.648
## 930    0031102203 3.829 1997     1711      2     2.626
## 1548   0031078854 4.109 1997     1711      2     2.917
## 3861   0033884858 3.916 2000     1711      2     2.769
## 8915  33646023117 3.828 2006     1702      4     2.535
## 11239 85032750937 4.080 2008     1711      3     2.793
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.3546 -0.4729  0.0267  0.4698  2.9167
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)          1.21432    0.04392   27.65   <2e-16 ***
## FirstAuthorFemale1  0.01026    0.03198    0.32    0.748
```

```
## LastAuthorFemale1  -0.11522    0.03502   -3.29    0.001 **
## Year1997           -0.02206    0.06397   -0.34    0.730
## Year1998            0.00854    0.05919    0.14    0.885
## Year1999           -0.01188    0.06103   -0.19    0.846
## Year2000           -0.06688    0.05893   -1.13    0.257
## Year2001            0.13000    0.06064    2.14    0.032 *
## Year2002           -0.00446    0.05895   -0.08    0.940
## Year2003            0.05265    0.05983    0.88    0.379
## Year2004            0.10284    0.06256    1.64    0.100
## Year2005            0.04103    0.06281    0.65    0.514
## Year2006            0.07867    0.05956    1.32    0.187
## Year2007            0.12724    0.05548    2.29    0.022 *
## Year2008            0.07248    0.05810    1.25    0.212
## Year2009            0.05060    0.05772    0.88    0.381
## Year2010           -0.02409    0.05689   -0.42    0.672
## Year2011           -0.00168    0.06242   -0.03    0.978
## Year2012            0.06163    0.06002    1.03    0.305
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.698
## Multiple R-squared:  0.00818,    Adjusted R-squared:  0.00478
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
##  438 weights are ~= 1. The remaining 4826 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.042   0.867   0.950   0.908   0.985   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.90e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##                   GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.01  1           1.005
## Year              1.01 16           1.000
```




```
## [1] "List of  6 outliers with residuals above  2.5"
##          ScopusId  NLCS Year OneField Fields residuals
## 782    0030287048 3.862 1996     1711      3     2.648
## 930    0031102203 3.829 1997     1711      2     2.626
## 1548   0031078854 4.109 1997     1711      2     2.917
## 3861   0033884858 3.916 2000     1711      2     2.769
## 8915  33646023117 3.828 2006     1702      4     2.535
## 11239 85032750937 4.080 2008     1711      3     2.793
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.3398 -0.4707  0.0279  0.4680  2.9221
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)          1.20979    0.04392   27.54   <2e-16 ***
## FirstAuthorFemale1 -0.01230    0.03239   -0.38    0.704
## Year1997           -0.02290    0.06392   -0.36    0.720
## Year1998            0.00372    0.05921    0.06    0.950
## Year1999           -0.01443    0.06120   -0.24    0.814
## Year2000           -0.06759    0.05907   -1.14    0.253
## Year2001            0.13001    0.06072    2.14    0.032 *
```

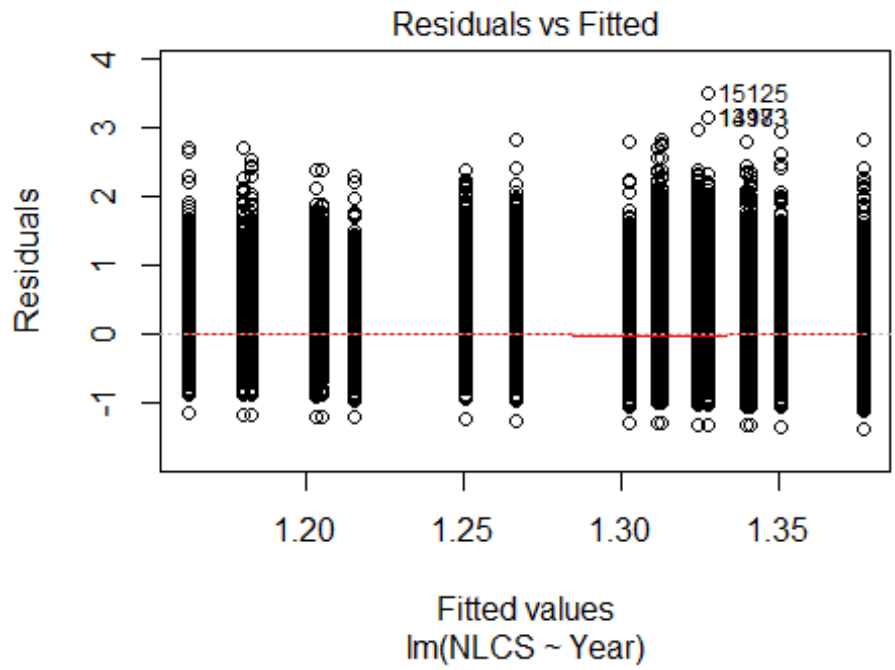
```
## Year2002           -0.00715    0.05910   -0.12    0.904
## Year2003            0.05158    0.05993    0.86    0.389
## Year2004            0.10005    0.06254    1.60    0.110
## Year2005            0.03716    0.06286    0.59    0.554
## Year2006            0.07739    0.05966    1.30    0.195
## Year2007            0.12272    0.05545    2.21    0.027 *
## Year2008            0.07030    0.05801    1.21    0.226
## Year2009            0.04342    0.05778    0.75    0.452
## Year2010           -0.02671    0.05689   -0.47    0.639
## Year2011           -0.00608    0.06258   -0.10    0.923
## Year2012            0.05871    0.06018    0.98    0.329
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.699
## Multiple R-squared:  0.00634,    Adjusted R-squared:  0.00312
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
##  434 weights are ~= 1. The remaining 4830 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0418  0.8670  0.9500  0.9080  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.90e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##                   GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.008  1           1.004
## Year             1.008 16           1.000
```



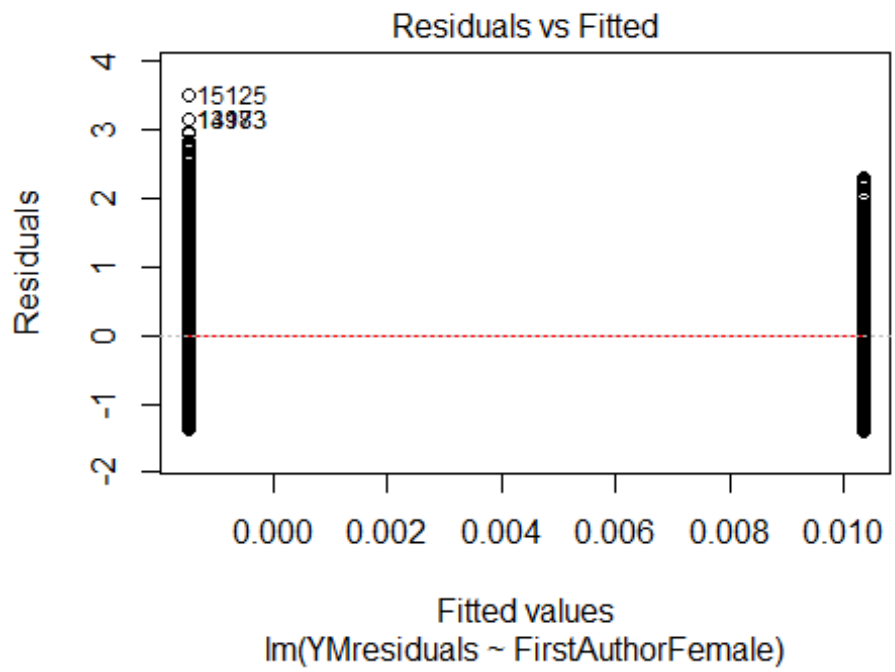

```
## [1] "List of  6 outliers with residuals above  2.5"
##          ScopusId  NLCS Year OneField Fields residuals
## 782    0030287048 3.862 1996     1711      3     2.648
## 930    0031102203 3.829 1997     1711      2     2.626
## 1548   0031078854 4.109 1997     1711      2     2.917
## 3861   0033884858 3.916 2000     1711      2     2.769
## 8915  33646023117 3.828 2006     1702      4     2.535
## 11239 85032750937 4.080 2008     1711      3     2.793
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.3451 -0.4732  0.0267  0.4695  2.9161
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.21497    0.04387   27.70   <2e-16 ***
## LastAuthorFemale1 -0.11277    0.03498   -3.22   0.0013 **
## Year1997          -0.02203    0.06397   -0.34   0.7306
## Year1998           0.00874    0.05917    0.15   0.8826
## Year1999          -0.01199    0.06103   -0.20   0.8443
## Year2000          -0.06690    0.05893   -1.14   0.2563
## Year2001           0.13014    0.06063    2.15   0.0319 *
```

```
## Year2002          -0.00437    0.05895   -0.07   0.9409
## Year2003           0.05309    0.05979    0.89   0.3746
## Year2004           0.10287    0.06255    1.64   0.1001
## Year2005           0.04122    0.06279    0.66   0.5115
## Year2006           0.07901    0.05957    1.33   0.1848
## Year2007           0.12743    0.05547    2.30   0.0216 *
## Year2008           0.07299    0.05805    1.26   0.2087
## Year2009           0.05104    0.05768    0.88   0.3762
## Year2010          -0.02385    0.05687   -0.42   0.6749
## Year2011          -0.00135    0.06244   -0.02   0.9828
## Year2012           0.06202    0.05999    1.03   0.3013
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.698
## Multiple R-squared:  0.00817,    Adjusted R-squared:  0.00496
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
##  438 weights are ~= 1. The remaining 4826 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0421  0.8670  0.9500  0.9080  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.90e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Sample size for the above analysis:  5264"
## [1] ""
## [1] ""
## [1] "####################################"
## [1] "Analysis of AJSC 1712"
## [1] "####################################"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 2415 2446 2229 2332 2597 2858 2386 1874 1965 2268 2473 2474 2368 2529 2257
## 2011 2012
```

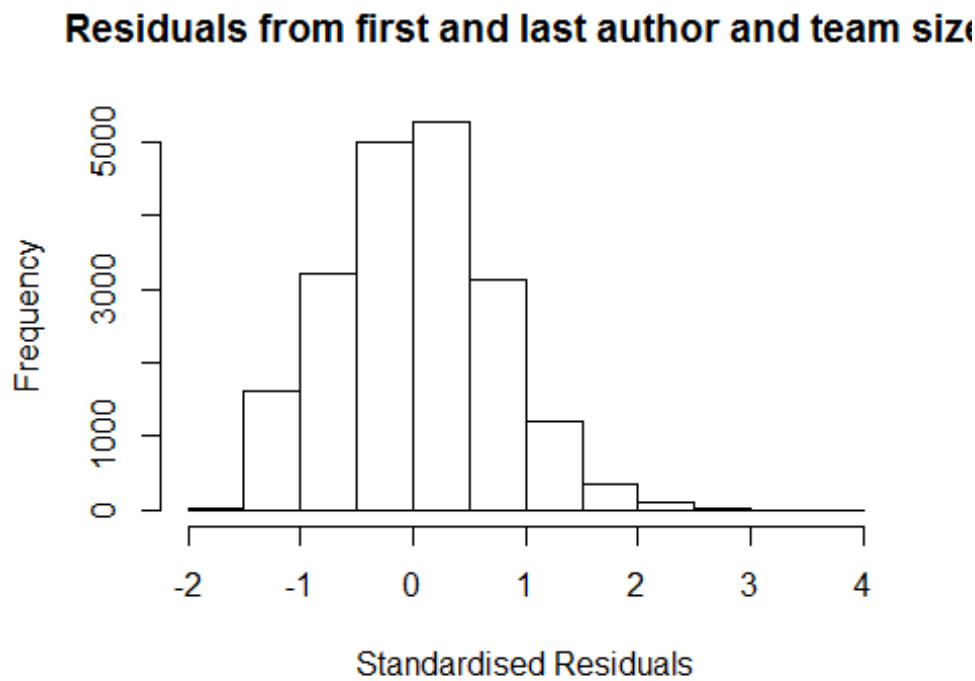
```
## 2160 2203
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 1608 1593 1484 1486 1677 1717 1478 1156 1210 1374 1457 1451 1402 1491 1339
## 2011 2012
## 1267 1248
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 1335 1348 1247 1248 1424 1444 1231  931  966 1116 1164 1169 1108 1199 1042
## 2011 2012
##  996  965
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
##  Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 350, df = 16, p-value <2e-16
```




```
##
##  Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 18, df = 1, p-value = 2e-05
```



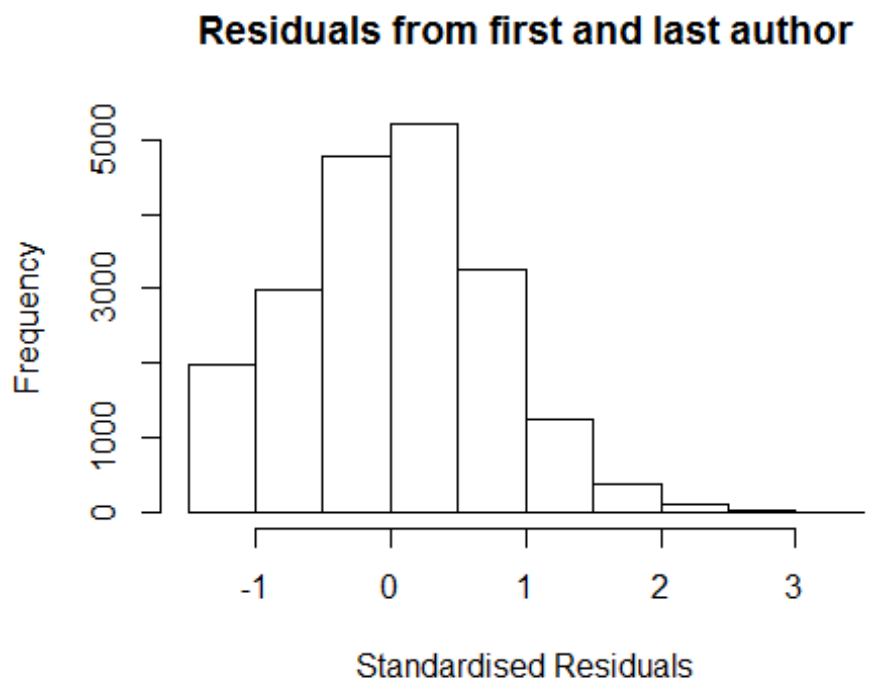

```
## [1] "Female first author team size 2018 geometric mean: 2.71520370865536"
## [1] "Male first author team size 2018 geometric mean: 2.64000731724081"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 27000, p-value = 0.6
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 2.55734587718235"
## [1] "Male last author team size 2018 geometric mean: 2.66758417679071"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 26000, p-value = 0.7
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.111  1           1.054
## LastAuthorFemale  1.112  1           1.054
## UniqueAuthors     1.038  4           1.005
## Year              1.041 16           1.001
```




```
## [1] "List of  20 outliers with residuals above  2.5"
##          ScopusId  NLCS Year OneField Fields residuals
## 2762  18944390941 3.710 1997     1712      1     2.704
## 2963   0031118203 4.089 1997     1703      3     2.728
## 3237   0031383380 3.875 1997     1705      4     2.514
## 3339   0031185845 4.068 1997     1702      5     2.627
## 4580   0031153727 4.130 1997     1708      3     2.769
## 6810   0032251894 4.293 1998     1712      2     2.816
## 10328  4243148480 4.090 1999     1702      5     3.122
## 13973  0035478854 4.469 2001     1702      2     3.501
## 14183  0035455653 4.472 2001     1708      3     3.068
## 17302  0036811662 3.649 2002     1702      4     2.619
## 20611  0141607824 4.104 2003     1702      4     2.685
## 22218 10944266504 4.134 2004     1705      3     2.773
## 22587  3843102623 3.655 2004     1704      3     2.647
## 24427 28244441640 3.576 2005     1704      2     2.542
## 24538 33846118079 3.662 2005     1708      3     2.628
## 25594 13944252629 4.211 2005     1705      3     2.742
## 28243 33646023117 3.828 2006     1702      4     2.824
## 28967 31144478351 3.964 2006     1704      2     2.605
## 39612 77950369345 3.468 2010     1702      4     2.670
## 40880 77950537175 3.862 2010     1712      3     2.628
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
UniqueAuthors +
##     Year, data = AllScopusDataOlderFirstLastGendered, control =
```

```
lmrob.control(fast.s.large.n = Inf,
##     k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.5733 -0.4863  0.0104  0.4805  3.5008
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.00688    0.02474   40.69  < 2e-16 ***
## FirstAuthorFemale1  0.02358    0.01608    1.47    0.142
## LastAuthorFemale1  -0.00246    0.01637   -0.15    0.880
## UniqueAuthors2      0.35524    0.01386   25.63  < 2e-16 ***
## UniqueAuthors3      0.43561    0.01550   28.11  < 2e-16 ***
## UniqueAuthors4      0.46676    0.01880   24.83  < 2e-16 ***
## UniqueAuthors5      0.51610    0.01988   25.96  < 2e-16 ***
## Year1997           -0.00130    0.03274   -0.04    0.968
## Year1998            0.00321    0.03229    0.10    0.921
## Year1999           -0.03860    0.03309   -1.17    0.243
## Year2000           -0.06278    0.03094   -2.03    0.042 *
## Year2001           -0.03863    0.03198   -1.21    0.227
## Year2002            0.00179    0.03217    0.06    0.956
## Year2003           -0.02389    0.03256   -0.73    0.463
## Year2004            0.00116    0.03266    0.04    0.972
## Year2005            0.02671    0.03150    0.85    0.396
## Year2006           -0.00266    0.03148   -0.08    0.933
## Year2007           -0.12773    0.02998   -4.26  2.0e-05 ***
## Year2008           -0.14473    0.03037   -4.77  1.9e-06 ***
## Year2009           -0.16288    0.02986   -5.45  5.0e-08 ***
## Year2010           -0.20842    0.03069   -6.79  1.1e-11 ***
## Year2011           -0.18862    0.03177   -5.94  3.0e-09 ***
## Year2012           -0.16933    0.03171   -5.34  9.4e-08 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.724
## Multiple R-squared:  0.0724, Adjusted R-squared:  0.0714
## Convergence in 10 IRWLS iterations
##
## Robustness weights:
##  observation 6779 is an outlier with |weight| = 0 ( < 5e-06);
##  1703 weights are ~= 1. The remaining 18229 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0235  0.8700  0.9510  0.9100  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          5.02e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
```

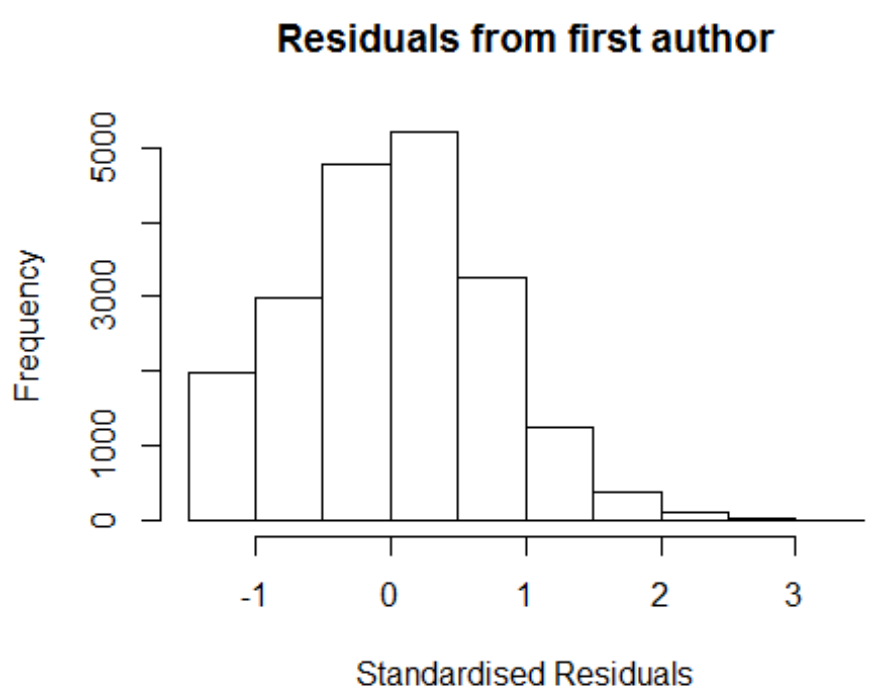
```
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                    psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.091  1           1.044
## LastAuthorFemale  1.090  1           1.044
## Year              1.006 16           1.000
```




```
## [1] "List of  16 outliers with residuals above  2.5"
##          ScopusId  NLCS Year OneField Fields residuals
## 2963   0031118203 4.089 1997     1703      3     2.825
## 3237   0031383380 3.875 1997     1705      4     2.611
## 3339   0031185845 4.068 1997     1702      5     2.804
## 4580   0031153727 4.130 1997     1708      3     2.866
## 6810   0032251894 4.293 1998     1712      2     3.010
## 10328  4243148480 4.090 1999     1702      5     2.867
## 13973  0035478854 4.469 2001     1702      2     3.192
## 14183  0035455653 4.472 2001     1708      3     3.195
```

```
## 20611  0141607824 4.104 2003     1702      4     2.829
## 22218 10944266504 4.134 2004     1705      3     2.830
## 25594 13944252629 4.211 2005     1705      3     2.869
## 27067 33750369245 3.834 2006     1704      2     2.515
## 28243 33646023117 3.828 2006     1702      4     2.509
## 28967 31144478351 3.964 2006     1704      2     2.645
## 40880 77950537175 3.862 2010     1712      3     2.731
## 43126 79953048649 3.731 2011     1702      5     2.590
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.3790 -0.4969  0.0223  0.5003  3.1952
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.260609   0.024371   51.73  < 2e-16 ***
## FirstAuthorFemale1  0.036860   0.016414    2.25  0.02474 *
## LastAuthorFemale1  -0.000525   0.016691   -0.03  0.97489
## Year1997            0.003336   0.033971    0.10  0.92177
## Year1998            0.022535   0.033548    0.67  0.50177
## Year1999           -0.037364   0.034275   -1.09  0.27567
## Year2000           -0.035465   0.032127   -1.10  0.26965
## Year2001            0.016221   0.033145    0.49  0.62457
## Year2002            0.046522   0.032992    1.41  0.15852
## Year2003            0.014059   0.033881    0.41  0.67818
## Year2004            0.043698   0.034137    1.28  0.20052
## Year2005            0.081570   0.033043    2.47  0.01357 *
## Year2006            0.058667   0.032685    1.79  0.07268 .
## Year2007           -0.064206   0.031171   -2.06  0.03943 *
## Year2008           -0.092210   0.031973   -2.88  0.00393 **
## Year2009           -0.096067   0.031149   -3.08  0.00204 **
## Year2010           -0.129277   0.032002   -4.04  5.4e-05 ***
## Year2011           -0.119524   0.032990   -3.62  0.00029 ***
## Year2012           -0.082950   0.032879   -2.52  0.01165 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.746
## Multiple R-squared:  0.00718,    Adjusted R-squared:  0.00629
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  1715 weights are ~= 1. The remaining 18218 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0269  0.8680  0.9500  0.9090  0.9850  0.9990
```

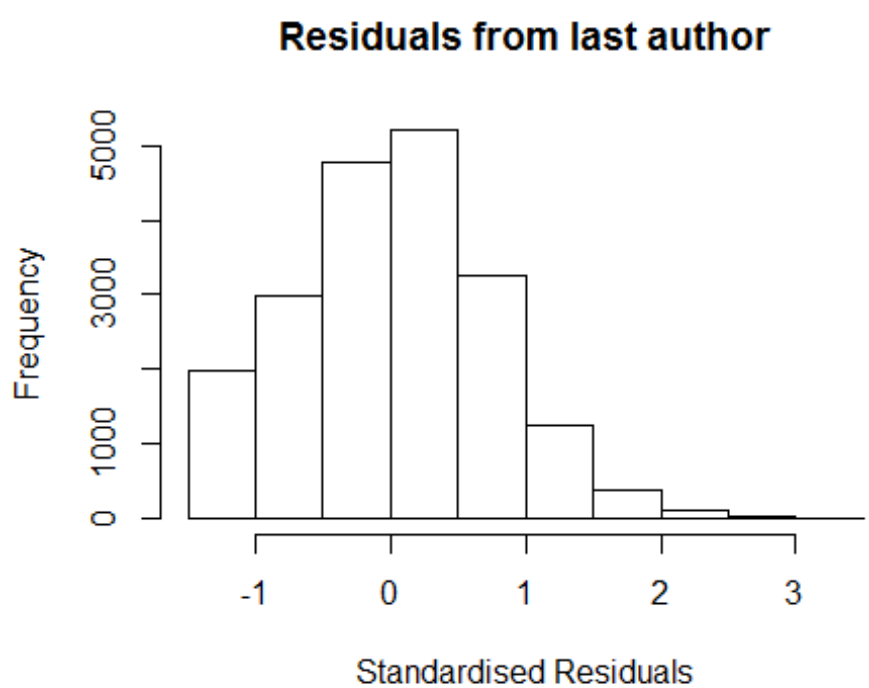
```
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          5.02e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.003  1           1.002
## Year              1.003 16           1.000
```




```
## [1] "List of  16 outliers with residuals above  2.5"
##          ScopusId  NLCS Year OneField Fields residuals
## 2963   0031118203 4.089 1997     1703      3     2.825
## 3237   0031383380 3.875 1997     1705      4     2.611
## 3339   0031185845 4.068 1997     1702      5     2.804
## 4580   0031153727 4.130 1997     1708      3     2.866
```

```
## 6810   0032251894 4.293 1998     1712      2     3.010
## 10328  4243148480 4.090 1999     1702      5     2.867
## 13973  0035478854 4.469 2001     1702      2     3.192
## 14183  0035455653 4.472 2001     1708      3     3.195
## 20611  0141607824 4.104 2003     1702      4     2.829
## 22218 10944266504 4.134 2004     1705      3     2.830
## 25594 13944252629 4.211 2005     1705      3     2.869
## 27067 33750369245 3.834 2006     1704      2     2.515
## 28243 33646023117 3.828 2006     1702      4     2.509
## 28967 31144478351 3.964 2006     1704      2     2.645
## 40880 77950537175 3.862 2010     1712      3     2.731
## 43126 79953048649 3.731 2011     1702      5     2.590
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.3788 -0.4970  0.0224  0.5004  3.1952
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.26057    0.02434   51.78  < 2e-16 ***
## FirstAuthorFemale1  0.03669    0.01576    2.33  0.01991 *
## Year1997            0.00333    0.03397    0.10  0.92195
## Year1998            0.02253    0.03355    0.67  0.50179
## Year1999           -0.03737    0.03427   -1.09  0.27555
## Year2000           -0.03546    0.03213   -1.10  0.26968
## Year2001            0.01622    0.03314    0.49  0.62458
## Year2002            0.04651    0.03299    1.41  0.15858
## Year2003            0.01405    0.03388    0.41  0.67830
## Year2004            0.04370    0.03414    1.28  0.20053
## Year2005            0.08156    0.03304    2.47  0.01358 *
## Year2006            0.05866    0.03268    1.79  0.07270 .
## Year2007           -0.06422    0.03117   -2.06  0.03938 *
## Year2008           -0.09221    0.03197   -2.88  0.00393 **
## Year2009           -0.09608    0.03114   -3.09  0.00204 **
## Year2010           -0.12929    0.03200   -4.04  5.3e-05 ***
## Year2011           -0.11954    0.03299   -3.62  0.00029 ***
## Year2012           -0.08296    0.03288   -2.52  0.01163 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.746
## Multiple R-squared:  0.00718,    Adjusted R-squared:  0.00634
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
```

```
##  1717 weights are ~= 1. The remaining 18216 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0268  0.8680  0.9500  0.9090  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          5.02e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                    psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##                   GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.002  1           1.001
## Year             1.002 16           1.000
```

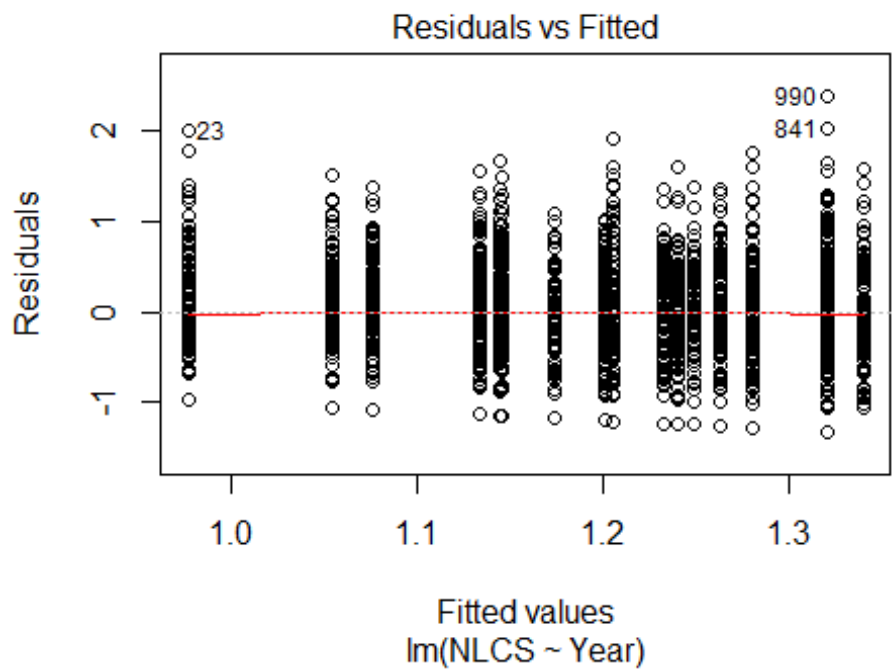



```
## [1] "List of  16 outliers with residuals above  2.5"
##          ScopusId  NLCS Year OneField Fields residuals
## 2963   0031118203 4.089 1997     1703      3     2.825
```

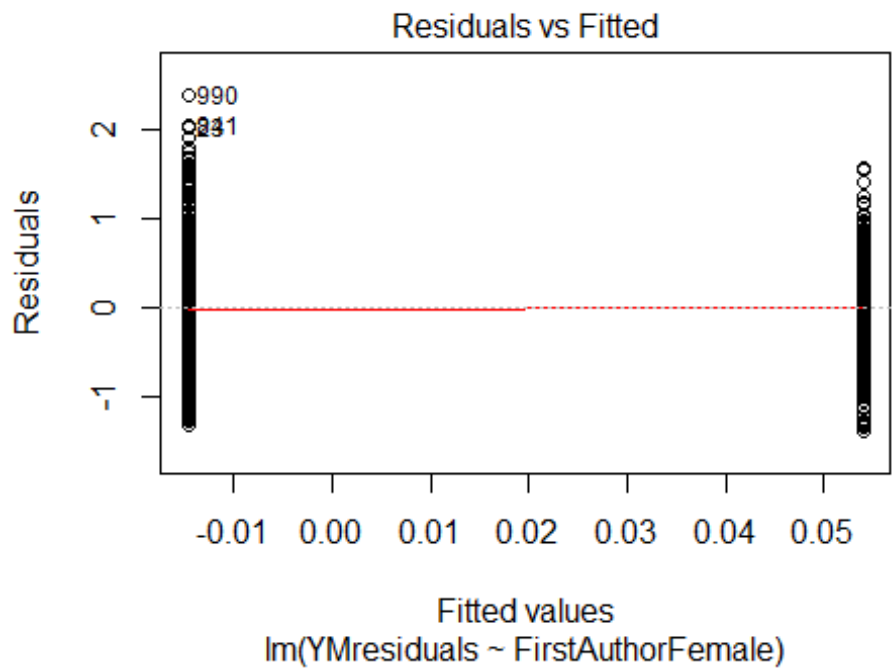
```
## 3237   0031383380 3.875 1997     1705      4     2.611
## 3339   0031185845 4.068 1997     1702      5     2.804
## 4580   0031153727 4.130 1997     1708      3     2.866
## 6810   0032251894 4.293 1998     1712      2     3.010
## 10328  4243148480 4.090 1999     1702      5     2.867
## 13973  0035478854 4.469 2001     1702      2     3.192
## 14183  0035455653 4.472 2001     1708      3     3.195
## 20611  0141607824 4.104 2003     1702      4     2.829
## 22218 10944266504 4.134 2004     1705      3     2.830
## 25594 13944252629 4.211 2005     1705      3     2.869
## 27067 33750369245 3.834 2006     1704      2     2.515
## 28243 33646023117 3.828 2006     1702      4     2.509
## 28967 31144478351 3.964 2006     1704      2     2.645
## 40880 77950537175 3.862 2010     1712      3     2.731
## 43126 79953048649 3.731 2011     1702      5     2.590
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.3570 -0.4973  0.0222  0.5007  3.1915
##
## Coefficients:
##                   Estimate Std. Error t value Pr(>|t|)
## (Intercept)        1.26325    0.02434   51.91  < 2e-16 ***
## LastAuthorFemale1  0.01173    0.01598    0.73  0.46277
## Year1997           0.00265    0.03398    0.08  0.93791
## Year1998           0.02306    0.03356    0.69  0.49200
## Year1999          -0.03746    0.03427   -1.09  0.27443
## Year2000          -0.03537    0.03214   -1.10  0.27110
## Year2001           0.01725    0.03316    0.52  0.60285
## Year2002           0.04784    0.03300    1.45  0.14717
## Year2003           0.01495    0.03391    0.44  0.65933
## Year2004           0.04475    0.03417    1.31  0.19038
## Year2005           0.08200    0.03306    2.48  0.01313 *
## Year2006           0.05951    0.03270    1.82  0.06882 .
## Year2007          -0.06339    0.03120   -2.03  0.04218 *
## Year2008          -0.09095    0.03198   -2.84  0.00447 **
## Year2009          -0.09549    0.03117   -3.06  0.00219 **
## Year2010          -0.12913    0.03202   -4.03  5.5e-05 ***
## Year2011          -0.11835    0.03300   -3.59  0.00034 ***
## Year2012          -0.08201    0.03289   -2.49  0.01266 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.746
## Multiple R-squared:  0.00695,    Adjusted R-squared:  0.0061
```

```
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  1731 weights are ~= 1. The remaining 18202 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0277  0.8680  0.9500  0.9090  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          5.02e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Sample size for the above analysis:  19933"
## [1] ""
## [1] ""
## [1] "####################################"
## [1] "Analysis of AJSC 1800"
## [1] "####################################"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  145  157  133  132  145  146  136  121  107  110  151  159  210  201  223
## 2011 2012
##  221  235
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  119  122  103  108  108  114  109  101   77   90  120  121  164  156  171
## 2011 2012
##  167  181
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  112  111   94   95   99   97   97   85   71   80  111  111  142  143  153
## 2011 2012
##  143  164
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
##  Bartlett test of homogeneity of variances
##
```

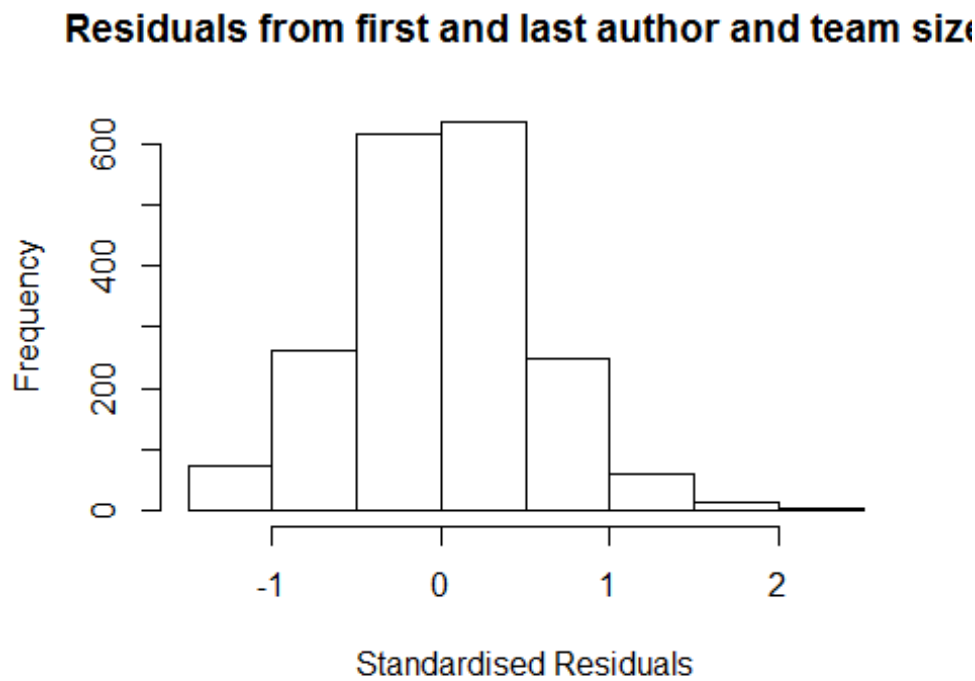
```
## data:  NLCS by Year
## Bartlett's K-squared = 45, df = 16, p-value = 1e-04
```




```
##
##  Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 1.2, df = 1, p-value = 0.3
```

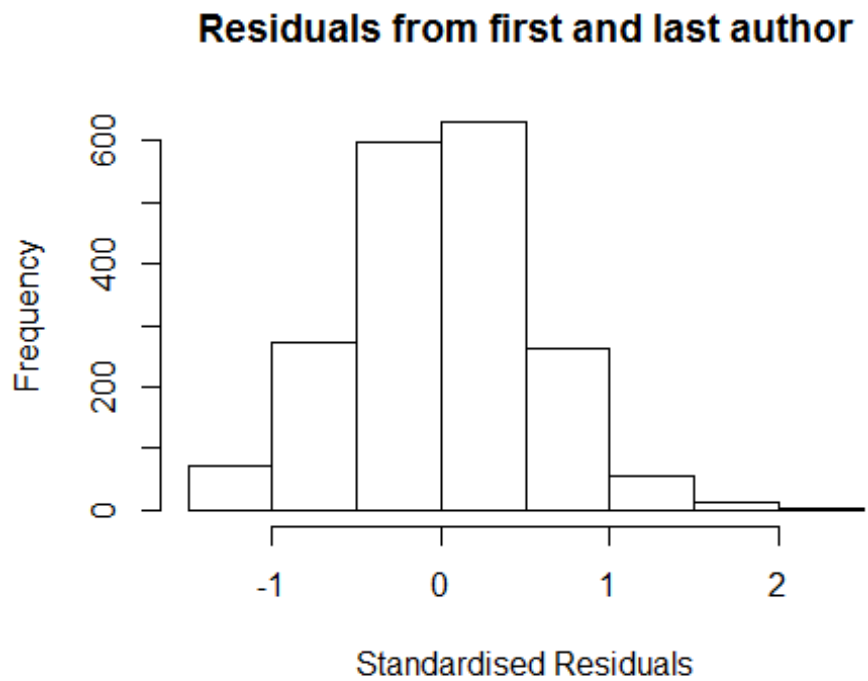



```
## [1] "Female first author team size 2018 geometric mean: 2.46637897403433"
## [1] "Male first author team size 2018 geometric mean: 2.21646335078082"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 2900, p-value = 0.4
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 2.19652533625302"
## [1] "Male last author team size 2018 geometric mean: 2.31074674489461"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 2200, p-value = 0.5
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.192  1           1.092
## LastAuthorFemale  1.175  1           1.084
## UniqueAuthors     1.200  4           1.023
## Year              1.283 16           1.008
```



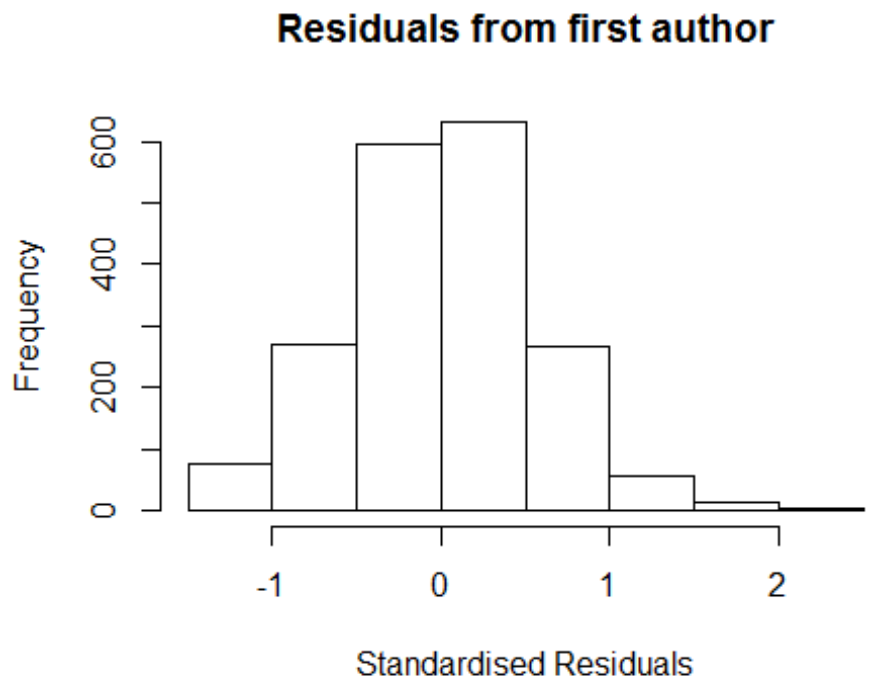

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
UniqueAuthors +
##     Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##     k.max = 1000))
##  \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.46337 -0.34359  0.00636  0.35099  2.33925
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)          0.8408     0.0656   12.82  < 2e-16 ***
## FirstAuthorFemale1   0.0656     0.0321    2.04  0.04128 *
## LastAuthorFemale1    0.0486     0.0324    1.50  0.13410
## UniqueAuthors2       0.0870     0.0326    2.67  0.00772 **
## UniqueAuthors3       0.2036     0.0366    5.56    3e-08 ***
## UniqueAuthors4       0.1403     0.0508    2.76  0.00580 **
## UniqueAuthors5       0.2142     0.0609    3.52  0.00044 ***
## Year1997             0.0890     0.0842    1.06  0.29051
## Year1998             0.2143     0.0955    2.25  0.02488 *
## Year1999             0.3005     0.0861    3.49  0.00050 ***
```

```
## Year2000             0.3367     0.0867    3.88  0.00011 ***
## Year2001             0.3533     0.0951    3.71  0.00021 ***
## Year2002             0.3807     0.0910    4.18    3e-05 ***
## Year2003             0.3465     0.0889    3.90    1e-04 ***
## Year2004             0.2103     0.0886    2.37  0.01771 *
## Year2005             0.3120     0.0864    3.61  0.00031 ***
## Year2006             0.1984     0.0855    2.32  0.02037 *
## Year2007             0.2569     0.0771    3.33  0.00088 ***
## Year2008             0.2992     0.0784    3.82  0.00014 ***
## Year2009             0.1758     0.0769    2.29  0.02227 *
## Year2010             0.2322     0.0749    3.10  0.00195 **
## Year2011             0.1818     0.0801    2.27  0.02336 *
## Year2012             0.1190     0.0749    1.59  0.11227
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.526
## Multiple R-squared:  0.0577, Adjusted R-squared:  0.0467
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
##  157 weights are ~= 1. The remaining 1751 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0096  0.8640  0.9520  0.8990  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          5.24e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.185  1           1.089
## LastAuthorFemale  1.171  1           1.082
## Year              1.082 16           1.002
```



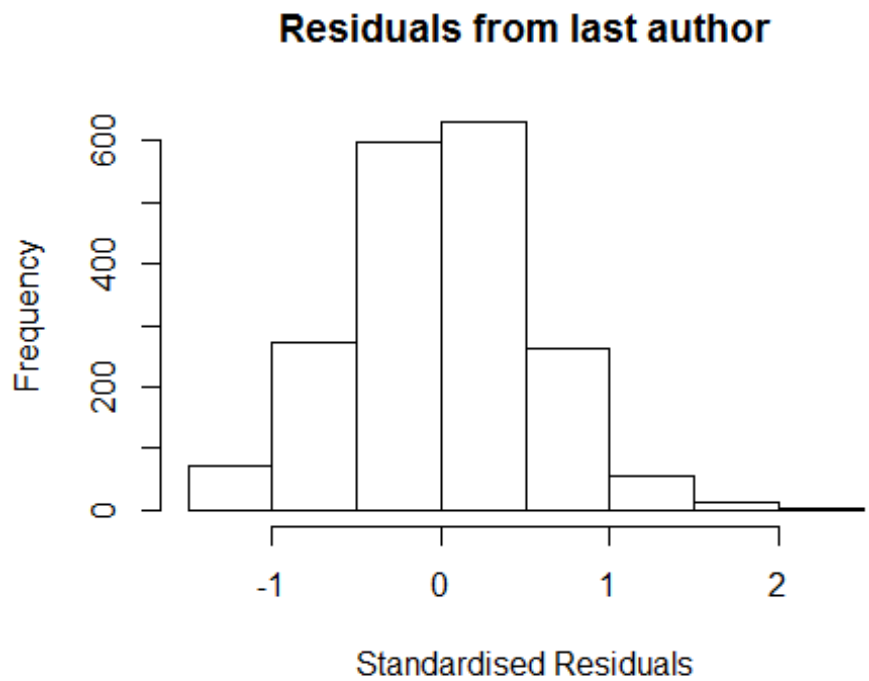

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.40227 -0.34338  0.00384  0.34581  2.41415
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)           0.8938     0.0651   13.73  < 2e-16 ***
## FirstAuthorFemale1    0.0727     0.0325    2.24  0.02537 *
## LastAuthorFemale1     0.0557     0.0329    1.69  0.09102 .
## Year1997              0.1004     0.0838    1.20  0.23125
## Year1998              0.2242     0.0949    2.36  0.01825 *
## Year1999              0.3159     0.0860    3.67  0.00025 ***
## Year2000              0.3522     0.0875    4.02  5.9e-05 ***
## Year2001              0.3801     0.0969    3.92  9.1e-05 ***
## Year2002              0.3930     0.0914    4.30  1.8e-05 ***
## Year2003              0.3927     0.0893    4.40  1.2e-05 ***
## Year2004              0.2519     0.0892    2.82  0.00480 **
## Year2005              0.3521     0.0870    4.05  5.4e-05 ***
```

```
## Year2006             0.2344     0.0862    2.72  0.00659 **
## Year2007             0.3021     0.0769    3.93  8.8e-05 ***
## Year2008             0.3469     0.0777    4.46  8.6e-06 ***
## Year2009             0.2205     0.0768    2.87  0.00414 **
## Year2010             0.2835     0.0745    3.81  0.00015 ***
## Year2011             0.2308     0.0802    2.88  0.00404 **
## Year2012             0.1767     0.0744    2.38  0.01759 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.529
## Multiple R-squared:  0.0389, Adjusted R-squared:  0.0298
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
##  175 weights are ~= 1. The remaining 1733 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0026  0.8590  0.9510  0.8980  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          5.24e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.046  1           1.023
## Year              1.046 16           1.001
```




```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.37654 -0.34170  0.00891  0.34559  2.40539
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)           0.9004     0.0647   13.92  < 2e-16 ***
## FirstAuthorFemale1    0.0925     0.0304    3.04  0.00242 **
## Year1997              0.1040     0.0835    1.24  0.21335
## Year1998              0.2234     0.0949    2.35  0.01866 *
## Year1999              0.3170     0.0859    3.69  0.00023 ***
## Year2000              0.3521     0.0875    4.02  5.9e-05 ***
## Year2001              0.3837     0.0963    3.99  7.0e-05 ***
## Year2002              0.3952     0.0912    4.33  1.6e-05 ***
## Year2003              0.3905     0.0892    4.38  1.3e-05 ***
## Year2004              0.2530     0.0886    2.85  0.00436 **
## Year2005              0.3525     0.0870    4.05  5.3e-05 ***
## Year2006              0.2346     0.0860    2.73  0.00641 **
```

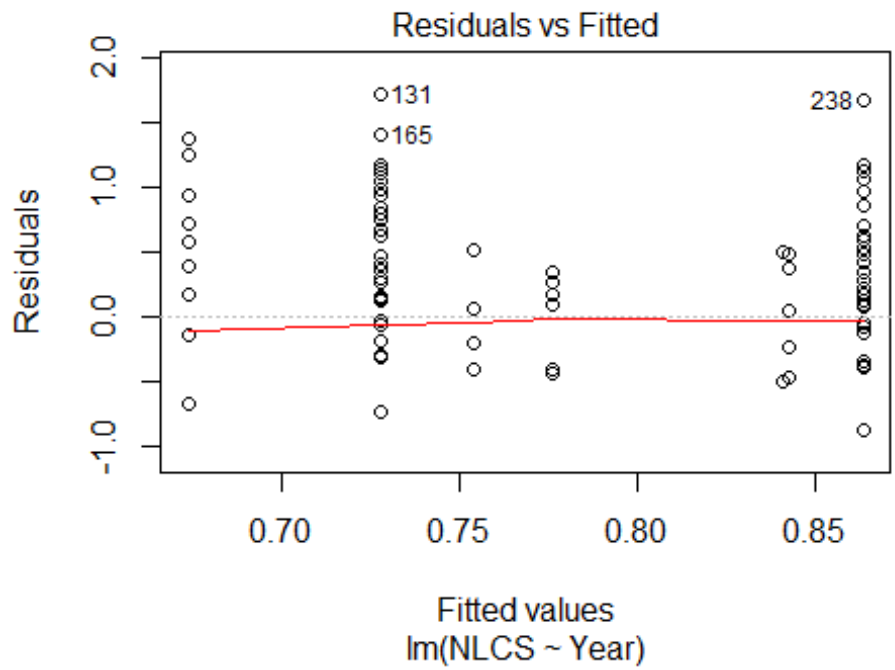
```
## Year2007             0.3074     0.0766    4.01  6.3e-05 ***
## Year2008             0.3474     0.0773    4.49  7.5e-06 ***
## Year2009             0.2221     0.0764    2.91  0.00368 **
## Year2010             0.2819     0.0744    3.79  0.00015 ***
## Year2011             0.2302     0.0802    2.87  0.00413 **
## Year2012             0.1792     0.0742    2.41  0.01585 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.529
## Multiple R-squared:  0.0376, Adjusted R-squared:  0.0289
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
##  173 weights are ~= 1. The remaining 1735 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0033  0.8590  0.9520  0.8980  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          5.24e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##                  GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.03  1           1.015
## Year             1.03 16           1.001
```




```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.36833 -0.34365  0.00396  0.35006  2.40353
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)          0.9087     0.0642   14.15  < 2e-16 ***
## LastAuthorFemale1    0.0817     0.0309    2.64  0.00827 **
## Year1997             0.0936     0.0836    1.12  0.26324
## Year1998             0.2150     0.0940    2.29  0.02233 *
## Year1999             0.3082     0.0860    3.58  0.00035 ***
## Year2000             0.3478     0.0875    3.98  7.3e-05 ***
## Year2001             0.3779     0.0969    3.90  1.0e-04 ***
## Year2002             0.3888     0.0914    4.25  2.2e-05 ***
## Year2003             0.3868     0.0890    4.34  1.5e-05 ***
## Year2004             0.2468     0.0894    2.76  0.00586 **
## Year2005             0.3463     0.0866    4.00  6.6e-05 ***
## Year2006             0.2379     0.0870    2.74  0.00629 **
```

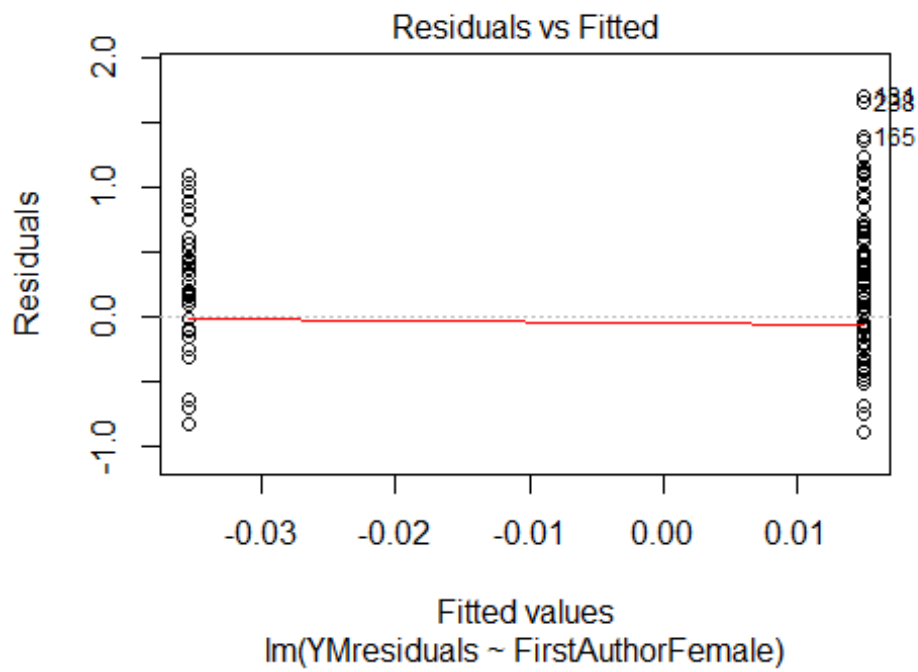
```
## Year2007            0.2940     0.0766    3.84  0.00013 ***
## Year2008            0.3405     0.0779    4.37  1.3e-05 ***
## Year2009            0.2124     0.0767    2.77  0.00570 **
## Year2010            0.2803     0.0745    3.76  0.00017 ***
## Year2011            0.2294     0.0803    2.86  0.00431 **
## Year2012            0.1732     0.0744    2.33  0.02010 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.533
## Multiple R-squared:  0.036,  Adjusted R-squared:  0.0273
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
##  178 weights are ~= 1. The remaining 1730 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0054  0.8610  0.9510  0.8990  0.9840  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          5.24e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Sample size for the above analysis:  1908"
## [1] ""
## [1] ""
## [1] "####################################"
## [1] "Analysis of AJSC 1801"
## [1] "####################################"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 2005 2007 2008 2009 2010 2011 2012
##    4    9    6   10   37  103   85
##
## 2005 2007 2008 2009 2010 2011 2012
##    4    6    2    9   24   89   72
##
## 2005 2007 2008 2009 2010 2011 2012
##    3    5    2    8   19   81   55
```

```
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
##  Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 6.6, df = 6, p-value = 0.4
```

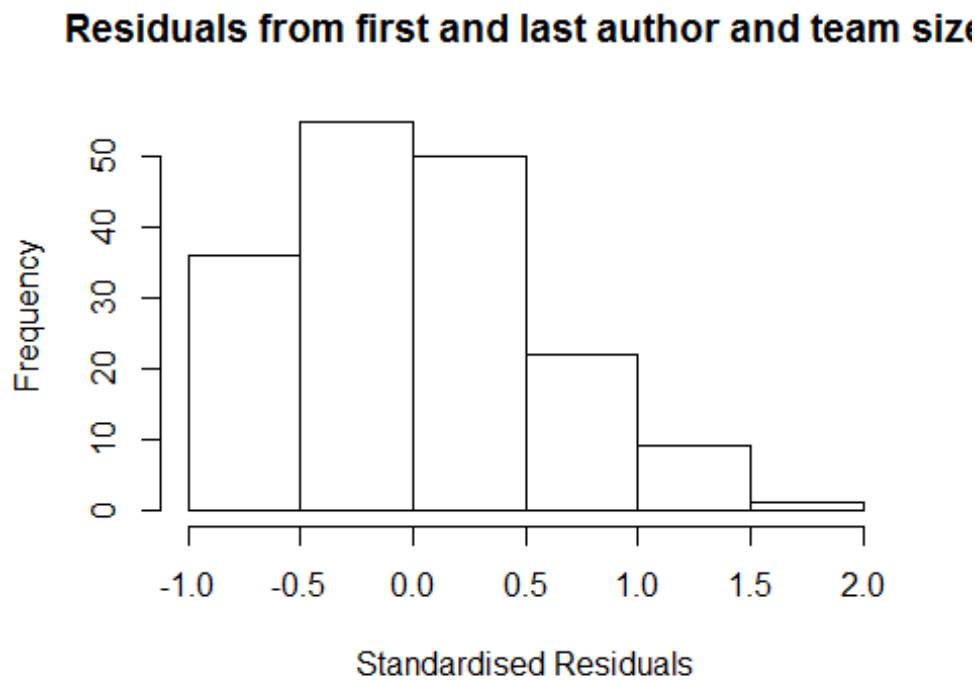



```
##
##  Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 1.7, df = 1, p-value = 0.2

## [1] "Female first author team size 2018 geometric mean: 2.35215804504935"
## [1] "Male first author team size 2018 geometric mean: 1.94328331572615"

## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties

##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 32, p-value = 0.4
## alternative hypothesis: true location shift is not equal to 0
##
```

```
## [1] "Female last author team size 2018 geometric mean: 1.8612097182042"
## [1] "Male last author team size 2018 geometric mean: 2.15301268319815"

## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties
```

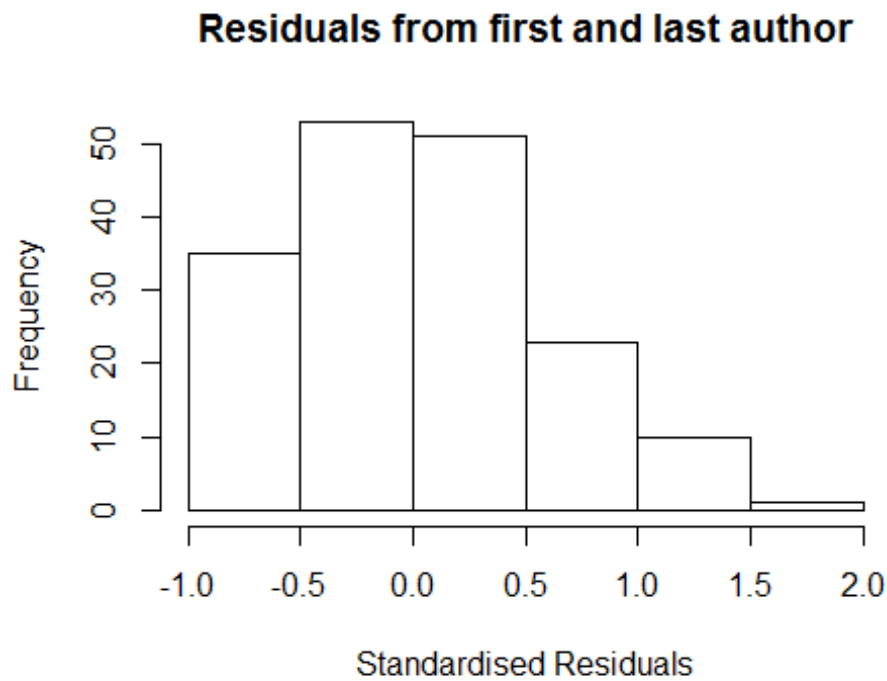



```
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 18, p-value = 0.6
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.887  1           1.374
## LastAuthorFemale  1.231  1           1.109
## UniqueAuthors     2.828  4           1.139
## Year              2.093  6           1.063
```



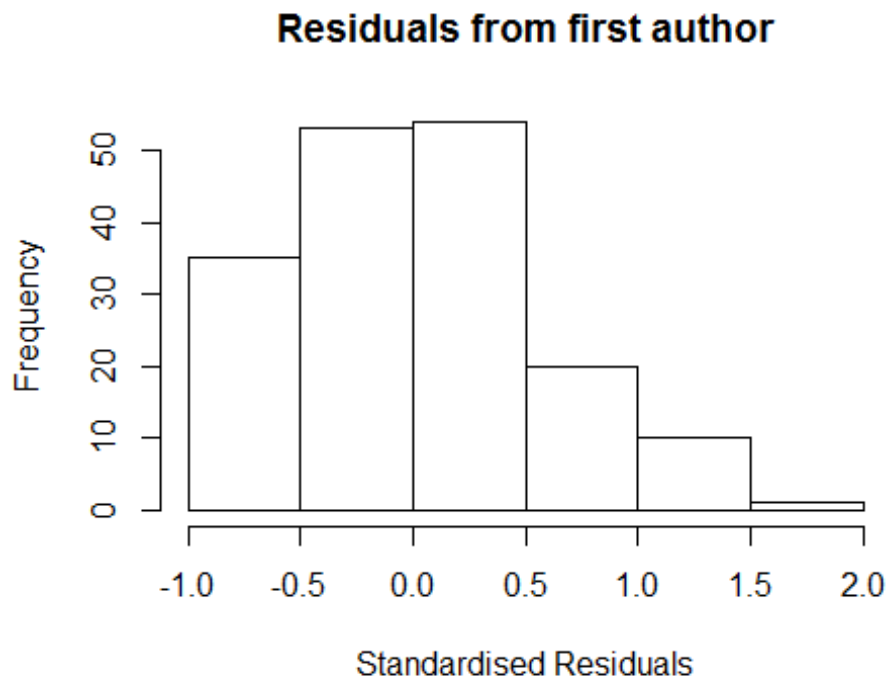

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
UniqueAuthors +
##     Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##     k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -0.9707 -0.3176 -0.0109  0.3564  1.8299
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         0.64617    0.30095    2.15    0.033 *
## FirstAuthorFemale1 -0.05090    0.11823   -0.43    0.667
## LastAuthorFemale1   0.17402    0.10569    1.65    0.102
## UniqueAuthors2      0.05108    0.12276    0.42    0.678
## UniqueAuthors3      0.06229    0.13018    0.48    0.633
## UniqueAuthors4      0.05906    0.20724    0.29    0.776
## UniqueAuthors5     -0.03229    0.17414   -0.19    0.853
## Year2007            0.02014    0.34810    0.06    0.954
## Year2008            0.14375    0.51508    0.28    0.781
## Year2009            0.00161    0.32104    0.01    0.996
```

```
## Year2010            0.07649    0.34387    0.22    0.824
## Year2011            0.00421    0.30892    0.01    0.989
## Year2012            0.09140    0.31613    0.29    0.773
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.589
## Multiple R-squared:  0.0226, Adjusted R-squared:  -0.0507
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
##  17 weights are ~= 1. The remaining 156 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.313   0.875   0.952   0.918   0.987   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          5.78e-04          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.201  1           1.096
## LastAuthorFemale  1.129  1           1.062
## Year              1.338  6           1.025
```



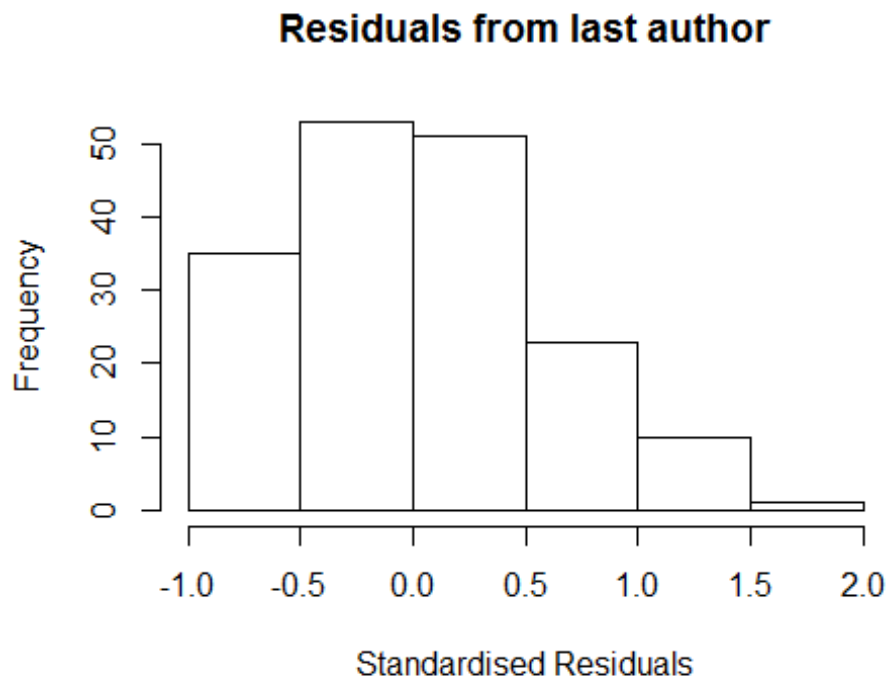

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -0.9496 -0.2976 -0.0216  0.3510  1.7750
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)           0.6601     0.2887    2.29    0.024 *
## FirstAuthorFemale1   -0.0571     0.0950   -0.60    0.548
## LastAuthorFemale1     0.1810     0.1019    1.78    0.078 .
## Year2007              0.0433     0.3215    0.13    0.893
## Year2008              0.1809     0.5010    0.36    0.718
## Year2009              0.0273     0.3078    0.09    0.929
## Year2010              0.0728     0.3314    0.22    0.826
## Year2011              0.0130     0.2971    0.04    0.965
## Year2012              0.1085     0.2983    0.36    0.716
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
```

```
## Robust residual standard error: 0.581
## Multiple R-squared:  0.0192, Adjusted R-squared:  -0.0286
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  14 weights are ~= 1. The remaining 159 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.330   0.881   0.954   0.918   0.986   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          5.78e-04          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.209  1           1.100
## Year              1.209  6           1.016
```




```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.78160 -0.31060 -0.00928  0.34653  1.73972
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)           0.7208     0.2414    2.99   0.0033 **
## FirstAuthorFemale1   -0.0308     0.0959   -0.32   0.7483
## Year2007              0.0204     0.2747    0.07   0.9408
## Year2008              0.1202     0.4722    0.25   0.7993
## Year2009              0.0312     0.2698    0.12   0.9080
## Year2010              0.0235     0.2935    0.08   0.9362
## Year2011             -0.0125     0.2533   -0.05   0.9608
## Year2012              0.0608     0.2538    0.24   0.8109
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.598
```

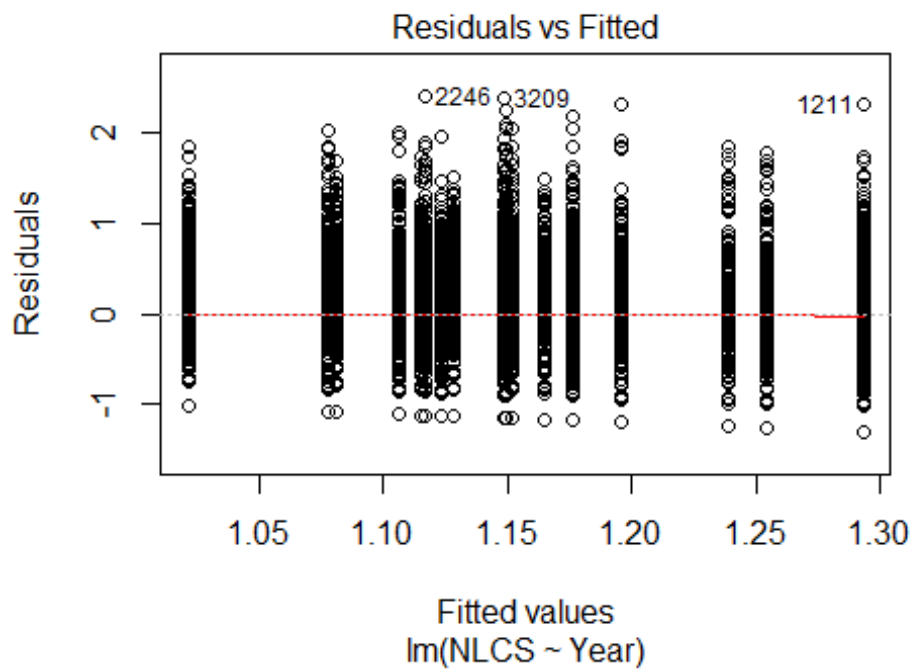
```
## Multiple R-squared:  0.00452,    Adjusted R-squared:  -0.0377
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  9 weights are ~= 1. The remaining 164 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.377   0.876   0.967   0.923   0.988   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          5.78e-04          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##                   GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.118  1           1.057
## Year             1.118  6           1.009
```



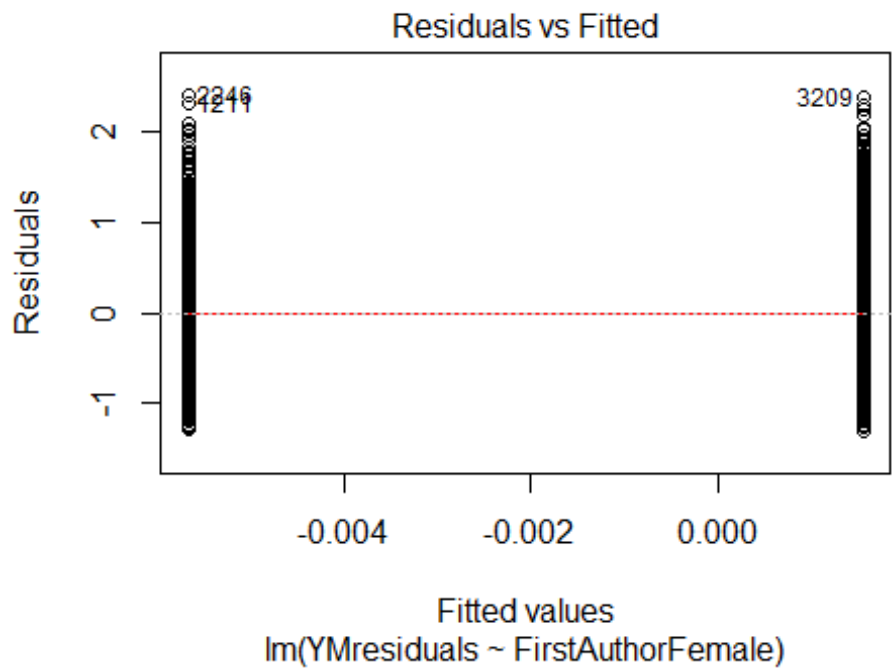

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -0.9235 -0.2835 -0.0075  0.3685  1.7925
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)          0.6639     0.2857    2.32    0.021 *
## LastAuthorFemale1    0.1690     0.1027    1.65    0.102
## Year2007             0.0419     0.3181    0.13    0.895
## Year2008             0.1771     0.4993    0.35    0.723
## Year2009             0.0134     0.3029    0.04    0.965
## Year2010             0.0398     0.3181    0.13    0.901
## Year2011            -0.0084     0.2916   -0.03    0.977
## Year2012             0.0906     0.2937    0.31    0.758
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.58
```

```
## Multiple R-squared:  0.0174, Adjusted R-squared:  -0.0243
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  15 weights are ~= 1. The remaining 158 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.320   0.887   0.954   0.917   0.986   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          5.78e-04          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Sample size for the above analysis:  173"
## [1] ""
## [1] ""
## [1] "####################################"
## [1] "Analysis of AJSC 1802"
## [1] "####################################"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  414  355  431  377  370  408  384  351  338  403  565  596  625  508  484
## 2011 2012
##  444  440
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  284  224  281  250  265  279  285  271  253  268  395  390  426  359  334
## 2011 2012
##  313  267
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  239  184  246  213  233  254  243  223  210  234  320  319  355  308  278
## 2011 2012
##  263  206
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
##  Bartlett test of homogeneity of variances
```

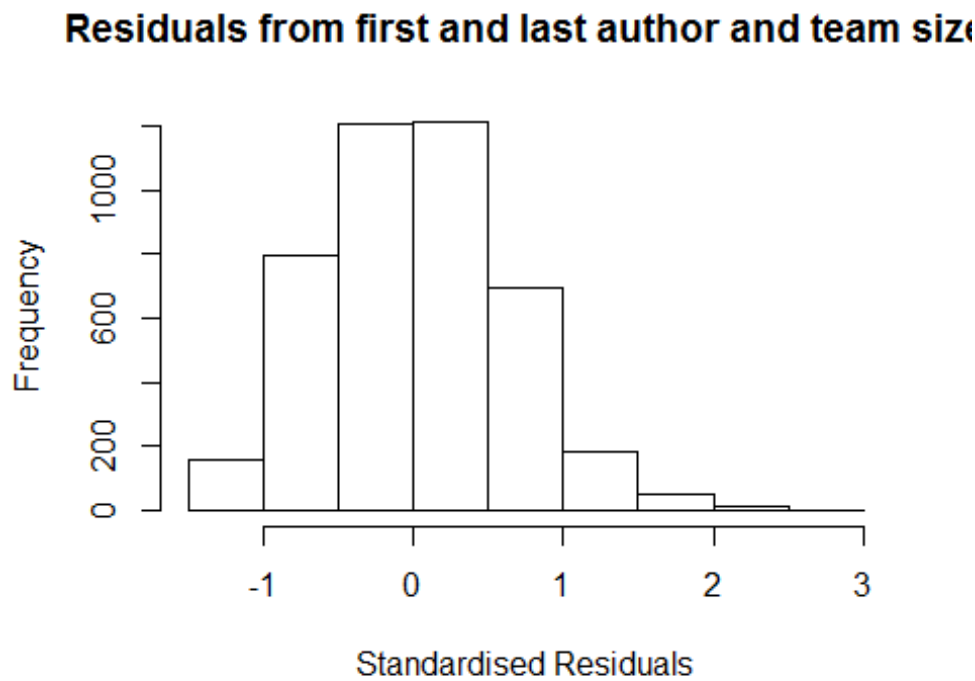
```
##
## data:  NLCS by Year
## Bartlett's K-squared = 34, df = 16, p-value = 0.005
```




```
##
##  Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 30, df = 1, p-value = 5e-08
```

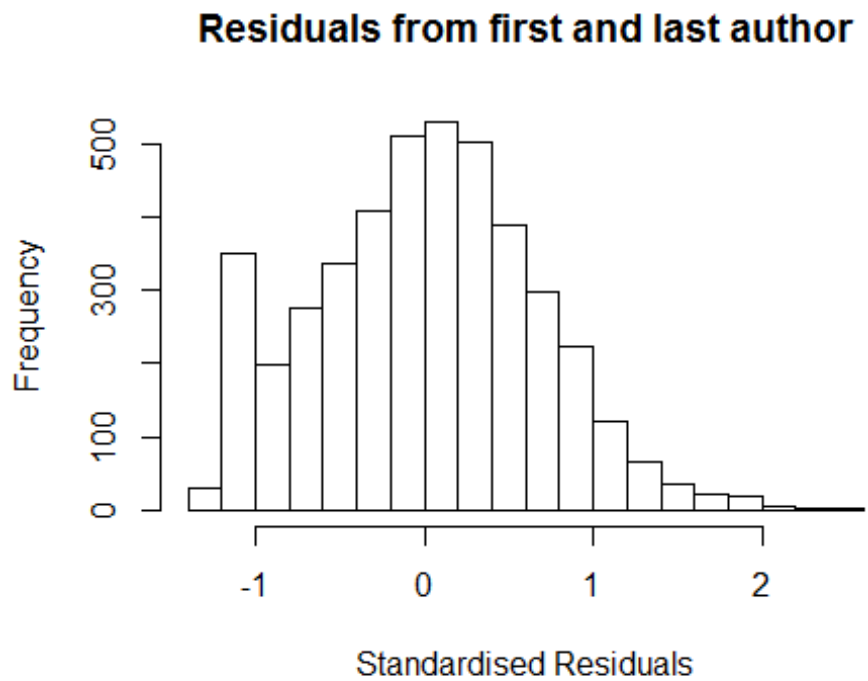



```
## [1] "Female first author team size 2018 geometric mean: 2.2442903186398"
## [1] "Male first author team size 2018 geometric mean: 2.26844676611441"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 3700, p-value = 0.6
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 2.25735369092427"
## [1] "Male last author team size 2018 geometric mean: 2.26361389643411"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 3500, p-value = 0.9
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.167  1           1.080
## LastAuthorFemale  1.172  1           1.083
## UniqueAuthors     1.070  4           1.009
## Year              1.105 16           1.003
```



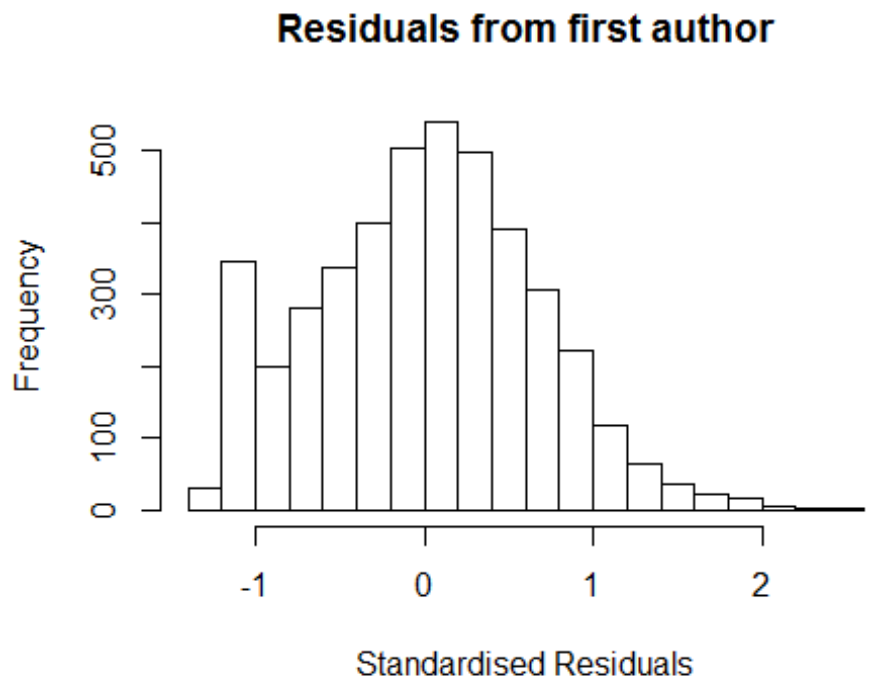

```
## [1] "List of  2 outliers with residuals above  2.5"
##        ScopusId  NLCS Year OneField Fields residuals
## 2160 0034559541 3.488 2000     1404      5     2.578
## 2987 0036015968 3.391 2002     1705      4     2.532
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
UniqueAuthors +
##     Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##     k.max = 1000))
##  \--> method = "MM"
## Residuals:
##       Min        1Q    Median        3Q       Max
## -1.449017 -0.436582 -0.000837  0.433286  2.577925
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         0.93588    0.05209   17.97  < 2e-16 ***
## FirstAuthorFemale1 -0.00535    0.02781   -0.19  0.84736
## LastAuthorFemale1  -0.04040    0.02928   -1.38  0.16778
## UniqueAuthors2      0.38069    0.02406   15.82  < 2e-16 ***
## UniqueAuthors3      0.45699    0.02727   16.76  < 2e-16 ***
## UniqueAuthors4      0.51059    0.04008   12.74  < 2e-16 ***
## UniqueAuthors5      0.33264    0.05726    5.81  6.7e-09 ***
## Year1997            0.02960    0.06903    0.43  0.66807
## Year1998            0.06150    0.06465    0.95  0.34156
```

```
## Year1999            0.01741    0.06448    0.27  0.78711
## Year2000           -0.02581    0.06504   -0.40  0.69151
## Year2001           -0.06513    0.06530   -1.00  0.31867
## Year2002           -0.07691    0.06836   -1.13  0.26062
## Year2003           -0.11257    0.06783   -1.66  0.09707 .
## Year2004           -0.07919    0.06854   -1.16  0.24800
## Year2005           -0.10338    0.06750   -1.53  0.12574
## Year2006           -0.18210    0.06208   -2.93  0.00337 **
## Year2007           -0.13135    0.06076   -2.16  0.03068 *
## Year2008           -0.20437    0.05933   -3.44  0.00058 ***
## Year2009           -0.11639    0.06179   -1.88  0.05970 .
## Year2010           -0.18025    0.06392   -2.82  0.00482 **
## Year2011           -0.09071    0.06567   -1.38  0.16726
## Year2012           -0.07883    0.06727   -1.17  0.24131
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.645
## Multiple R-squared:  0.106,  Adjusted R-squared:  0.102
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
##  347 weights are ~= 1. The remaining 3981 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0744  0.8730  0.9510  0.9120  0.9860  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          2.31e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.167  1           1.080
## LastAuthorFemale  1.167  1           1.080
## Year              1.044 16           1.001
```



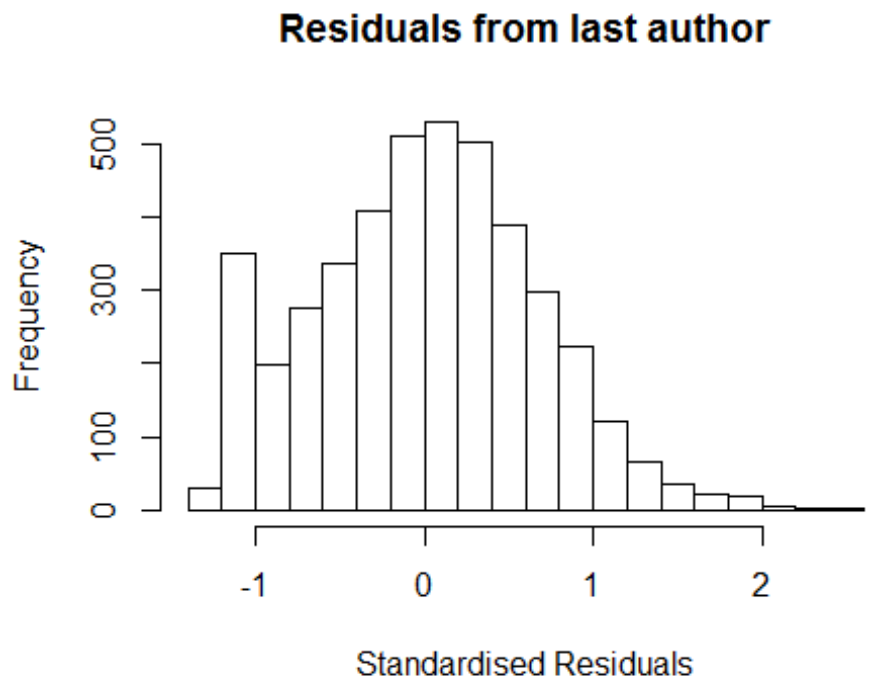

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.2877 -0.4549  0.0182  0.4497  2.4623
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)          1.17643    0.04819   24.41  < 2e-16 ***
## FirstAuthorFemale1   0.02287    0.02947    0.78  0.43779
## LastAuthorFemale1   -0.06190    0.03143   -1.97  0.04900 *
## Year1997             0.04837    0.06754    0.72  0.47394
## Year1998             0.08839    0.06293    1.40  0.16026
## Year1999             0.03930    0.06349    0.62  0.53598
## Year2000             0.00355    0.06435    0.06  0.95603
## Year2001            -0.07168    0.06641   -1.08  0.28049
## Year2002            -0.07886    0.06971   -1.13  0.25798
## Year2003            -0.11175    0.07044   -1.59  0.11269
## Year2004            -0.07803    0.07124   -1.10  0.27344
## Year2005            -0.08586    0.06690   -1.28  0.19940
```

```
## Year2006           -0.16935    0.06268   -2.70  0.00692 **
## Year2007           -0.09380    0.05950   -1.58  0.11496
## Year2008           -0.20016    0.05930   -3.38  0.00074 ***
## Year2009           -0.05774    0.06075   -0.95  0.34192
## Year2010           -0.11671    0.06391   -1.83  0.06789 .
## Year2011           -0.03666    0.06534   -0.56  0.57476
## Year2012           -0.03272    0.06623   -0.49  0.62129
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.672
## Multiple R-squared:  0.014,  Adjusted R-squared:  0.00992
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  393 weights are ~= 1. The remaining 3935 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.151   0.866   0.949   0.910   0.985   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          2.31e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.025  1           1.012
## Year              1.025 16           1.001
```




```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.2601 -0.4641  0.0196  0.4510  2.4704
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.16995    0.04833   24.21  < 2e-16 ***
## FirstAuthorFemale1 -0.00197    0.02802   -0.07  0.94406
## Year1997            0.04891    0.06792    0.72  0.47151
## Year1998            0.09015    0.06311    1.43  0.15323
## Year1999            0.04243    0.06360    0.67  0.50472
## Year2000            0.00683    0.06443    0.11  0.91557
## Year2001           -0.07107    0.06654   -1.07  0.28556
## Year2002           -0.07704    0.06996   -1.10  0.27086
## Year2003           -0.11336    0.07082   -1.60  0.10950
## Year2004           -0.07592    0.07140   -1.06  0.28768
## Year2005           -0.08780    0.06702   -1.31  0.19024
## Year2006           -0.16867    0.06277   -2.69  0.00723 **
```

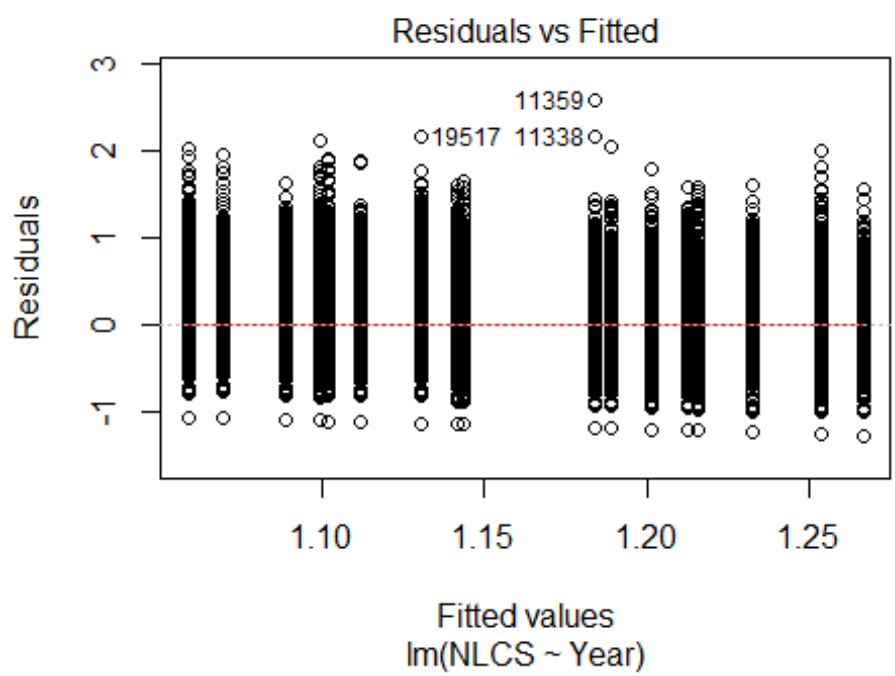
```
## Year2007           -0.09416    0.05968   -1.58  0.11469
## Year2008           -0.20103    0.05947   -3.38  0.00073 ***
## Year2009           -0.05499    0.06092   -0.90  0.36669
## Year2010           -0.11888    0.06413   -1.85  0.06383 .
## Year2011           -0.03771    0.06543   -0.58  0.56439
## Year2012           -0.03237    0.06646   -0.49  0.62624
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.672
## Multiple R-squared:  0.013,  Adjusted R-squared:  0.00911
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  398 weights are ~= 1. The remaining 3930 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.148   0.868   0.949   0.910   0.985   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          2.31e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##                   GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.024  1           1.012
## Year             1.024 16           1.001
```



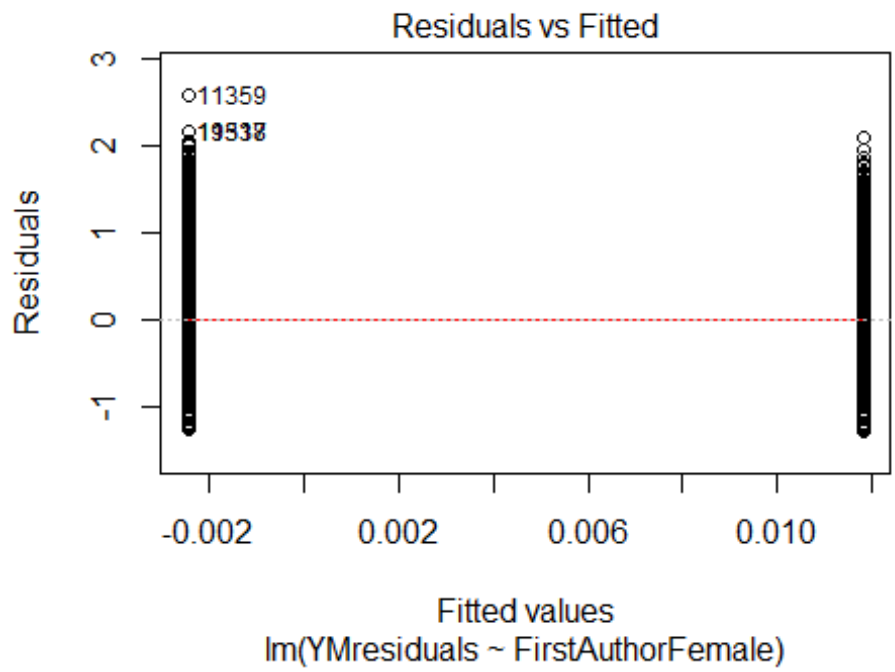

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.2675 -0.4567  0.0185  0.4518  2.4593
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.17858    0.04813   24.49  < 2e-16 ***
## LastAuthorFemale1 -0.05157    0.02939   -1.75  0.07939 .
## Year1997            0.04947    0.06743    0.73  0.46319
## Year1998            0.08889    0.06289    1.41  0.15760
## Year1999            0.03913    0.06343    0.62  0.53731
## Year2000            0.00403    0.06430    0.06  0.95006
## Year2001           -0.07058    0.06627   -1.07  0.28691
## Year2002           -0.07786    0.06961   -1.12  0.26342
## Year2003           -0.11087    0.07039   -1.57  0.11535
## Year2004           -0.07734    0.07122   -1.09  0.27757
## Year2005           -0.08530    0.06680   -1.28  0.20166
## Year2006           -0.16791    0.06249   -2.69  0.00724 **
```

```
## Year2007          -0.09301    0.05939   -1.57  0.11741
## Year2008          -0.19988    0.05923   -3.37  0.00075 ***
## Year2009          -0.05592    0.06053   -0.92  0.35566
## Year2010          -0.11557    0.06377   -1.81  0.07002 .
## Year2011          -0.03662    0.06525   -0.56  0.57465
## Year2012          -0.03133    0.06612   -0.47  0.63564
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.673
## Multiple R-squared:  0.0139, Adjusted R-squared:  0.00998
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  390 weights are ~= 1. The remaining 3938 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.153   0.867   0.949   0.910   0.985   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          2.31e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Sample size for the above analysis:  4328"
## [1] ""
## [1] ""
## [1] "####################################"
## [1] "Analysis of AJSC 1803"
## [1] "####################################"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 1109 1064  992  922  854  946  930  819  758  865 1014 1113 1226 1194 1157
## 2011 2012
## 1209 1119
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  737  655  629  611  596  585  640  562  527  559  649  731  813  795  774
## 2011 2012
```

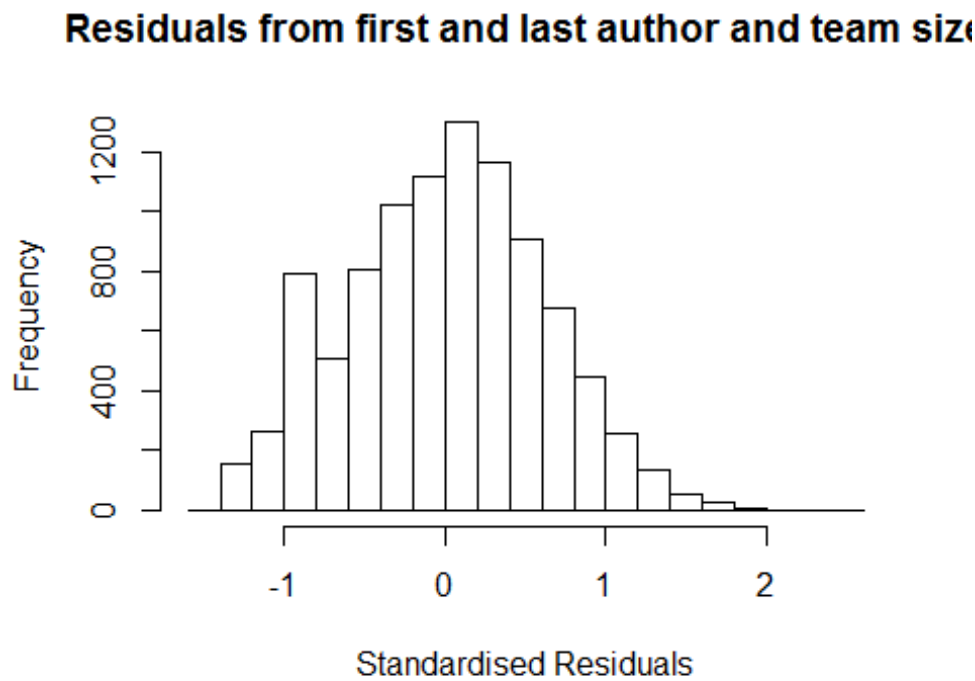
```
##  831  756
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  651  586  561  536  518  507  551  473  441  450  527  598  663  656  621
## 2011 2012
##  680  612
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
##  Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 130, df = 16, p-value <2e-16
```




```
##
##  Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 5.3, df = 1, p-value = 0.02
```

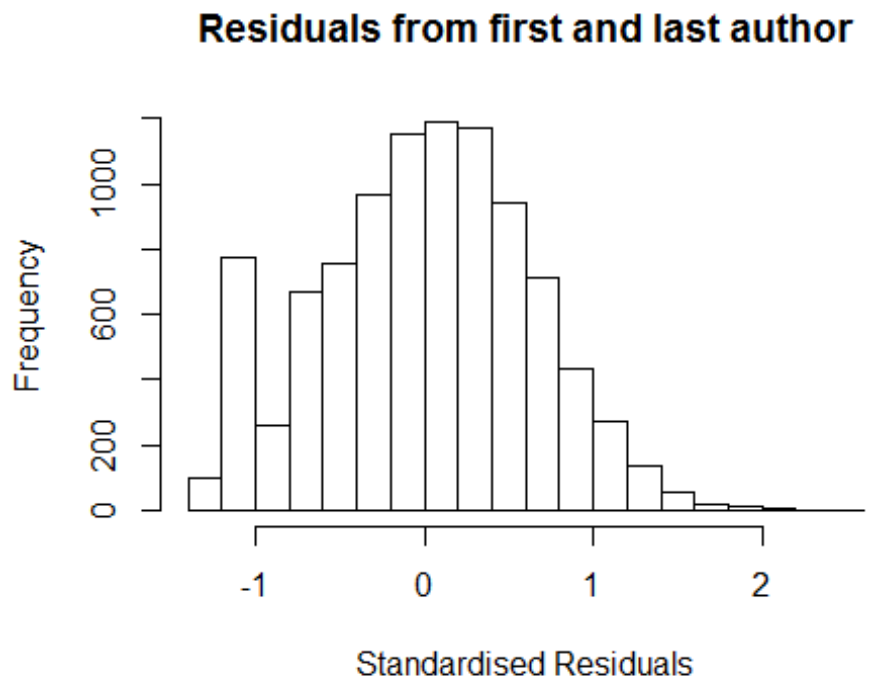



```
## [1] "Female first author team size 2018 geometric mean: 2.13676814967353"
## [1] "Male first author team size 2018 geometric mean: 2.16266140567915"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 27000, p-value = 0.9
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 2.05416011657984"
## [1] "Male last author team size 2018 geometric mean: 2.18483775764533"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 24000, p-value = 0.3
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.133  1           1.064
## LastAuthorFemale  1.133  1           1.065
## UniqueAuthors     1.065  4           1.008
## Year              1.076 16           1.002
```



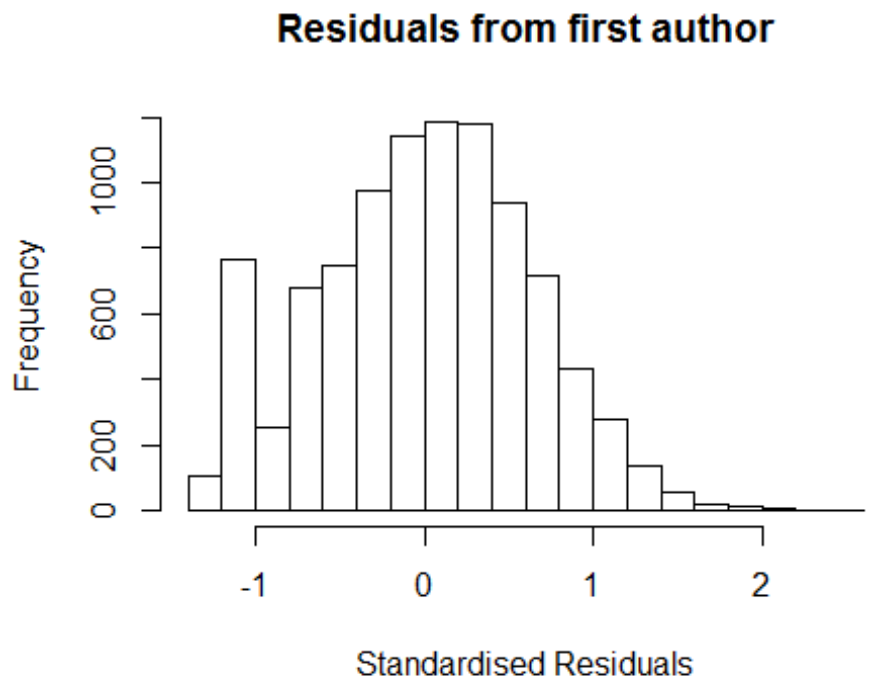

```
## [1] "List of  1 outliers with residuals above  2.5"
##          ScopusId NLCS Year OneField Fields residuals
## 11359 29144523061 3.76 2006     1700      7      2.53
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
UniqueAuthors +
##     Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##     k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.4447 -0.4265  0.0214  0.4194  2.5296
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         0.92654    0.03028   30.59  < 2e-16 ***
## FirstAuthorFemale1  0.00982    0.01879    0.52   0.6013
## LastAuthorFemale1  -0.02003    0.01922   -1.04   0.2974
## UniqueAuthors2      0.28203    0.01597   17.66  < 2e-16 ***
## UniqueAuthors3      0.35382    0.01812   19.52  < 2e-16 ***
## UniqueAuthors4      0.36110    0.02774   13.02  < 2e-16 ***
## UniqueAuthors5      0.24466    0.04151    5.89  3.9e-09 ***
## Year1997           -0.07676    0.04199   -1.83   0.0676 .
## Year1998           -0.03307    0.04249   -0.78   0.4364
## Year1999            0.03542    0.03986    0.89   0.3742
```

```
## Year2000            0.06204    0.03934    1.58   0.1149
## Year2001            0.11189    0.04237    2.64   0.0083 **
## Year2002            0.05699    0.03860    1.48   0.1398
## Year2003            0.08350    0.03963    2.11   0.0351 *
## Year2004            0.16437    0.03968    4.14  3.5e-05 ***
## Year2005            0.10422    0.03909    2.67   0.0077 **
## Year2006            0.02188    0.03964    0.55   0.5810
## Year2007            0.03362    0.03767    0.89   0.3722
## Year2008           -0.00573    0.03793   -0.15   0.8798
## Year2009           -0.07966    0.03938   -2.02   0.0431 *
## Year2010           -0.06267    0.03890   -1.61   0.1072
## Year2011           -0.11930    0.03813   -3.13   0.0018 **
## Year2012           -0.07456    0.04067   -1.83   0.0668 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.629
## Multiple R-squared:  0.0674, Adjusted R-squared:  0.0653
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  806 weights are ~= 1. The remaining 8825 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0689  0.8620  0.9500  0.9120  0.9860  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.04e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.114  1           1.056
## LastAuthorFemale  1.114  1           1.056
## Year              1.016 16           1.001
```



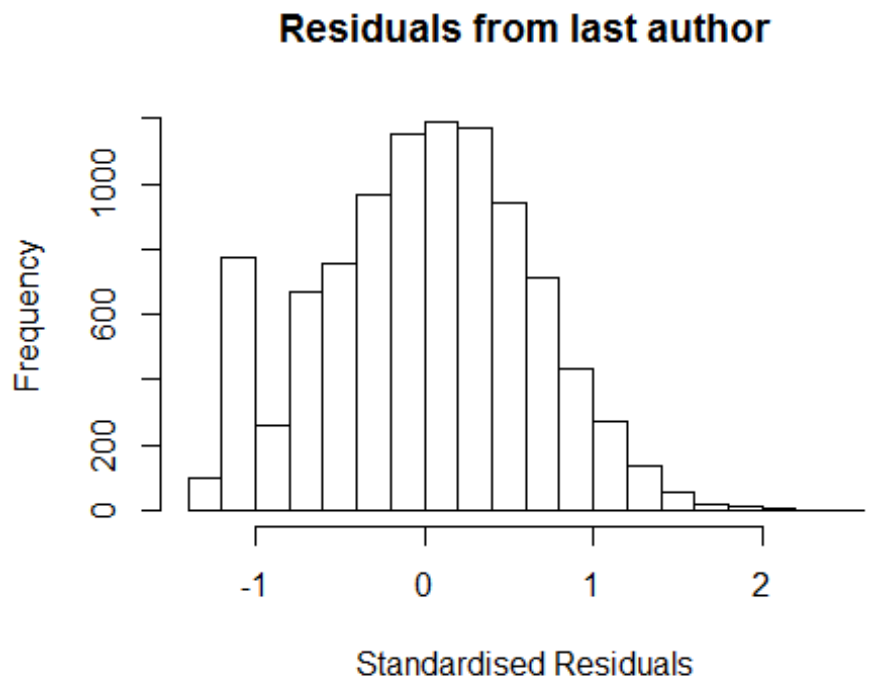

```
## [1] "List of  1 outliers with residuals above  2.5"
##          ScopusId NLCS Year OneField Fields residuals
## 11359 29144523061 3.76 2006     1700      7     2.594
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.3016 -0.4381  0.0226  0.4329  2.5936
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)          1.07609    0.02985   36.05  < 2e-16 ***
## FirstAuthorFemale1   0.02598    0.01932    1.34  0.17882
## LastAuthorFemale1   -0.01203    0.01972   -0.61  0.54180
## Year1997            -0.05593    0.04308   -1.30  0.19419
## Year1998            -0.01067    0.04383   -0.24  0.80773
## Year1999             0.06578    0.04031    1.63  0.10277
## Year2000             0.10392    0.03951    2.63  0.00854 **
## Year2001             0.15803    0.04350    3.63  0.00028 ***
## Year2002             0.10838    0.04005    2.71  0.00683 **
## Year2003             0.12545    0.04024    3.12  0.00183 **
## Year2004             0.21156    0.03996    5.29  1.2e-07 ***
## Year2005             0.15854    0.03980    3.98  6.8e-05 ***
```

```
## Year2006            0.09030    0.04023    2.24  0.02482 *
## Year2007            0.11259    0.03809    2.96  0.00312 **
## Year2008            0.05494    0.03860    1.42  0.15464
## Year2009           -0.01121    0.04025   -0.28  0.78059
## Year2010            0.00532    0.04007    0.13  0.89438
## Year2011           -0.04536    0.03907   -1.16  0.24565
## Year2012           -0.00325    0.04138   -0.08  0.93748
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.644
## Multiple R-squared:  0.0137, Adjusted R-squared:  0.0118
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  769 weights are ~= 1. The remaining 8862 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0679  0.8740  0.9500  0.9120  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.04e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.009  1           1.004
## Year              1.009 16           1.000
```




```
## [1] "List of  1 outliers with residuals above  2.5"
##          ScopusId NLCS Year OneField Fields residuals
## 11359 29144523061 3.76 2006     1700      7     2.594
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.3078 -0.4385  0.0237  0.4330  2.5948
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)          1.07499    0.02979   36.08  < 2e-16 ***
## FirstAuthorFemale1   0.02147    0.01847    1.16  0.24497
## Year1997            -0.05611    0.04309   -1.30  0.19286
## Year1998            -0.01066    0.04385   -0.24  0.80793
## Year1999             0.06573    0.04034    1.63  0.10322
## Year2000             0.10410    0.03951    2.63  0.00843 **
## Year2001             0.15832    0.04350    3.64  0.00027 ***
## Year2002             0.10817    0.04006    2.70  0.00694 **
## Year2003             0.12535    0.04024    3.12  0.00184 **
## Year2004             0.21135    0.03996    5.29  1.3e-07 ***
## Year2005             0.15836    0.03979    3.98  6.9e-05 ***
## Year2006             0.09023    0.04023    2.24  0.02495 *
```

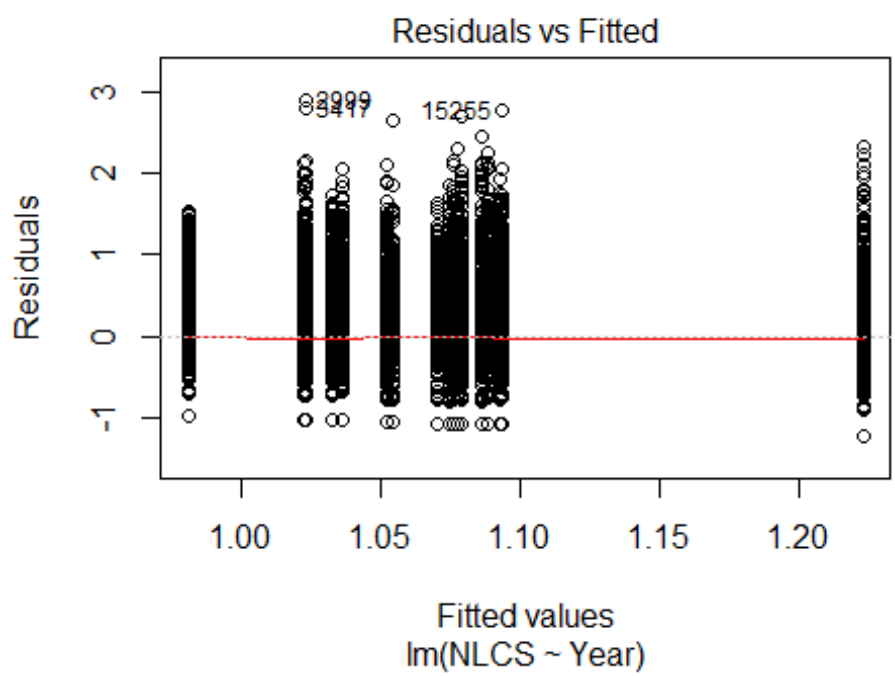
```
## Year2007            0.11226    0.03808    2.95  0.00320 **
## Year2008            0.05498    0.03860    1.42  0.15440
## Year2009           -0.01125    0.04025   -0.28  0.77979
## Year2010            0.00500    0.04007    0.12  0.90073
## Year2011           -0.04553    0.03907   -1.17  0.24390
## Year2012           -0.00371    0.04136   -0.09  0.92855
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.644
## Multiple R-squared:  0.0136, Adjusted R-squared:  0.0119
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  767 weights are ~= 1. The remaining 8864 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0675  0.8740  0.9500  0.9120  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.04e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##                   GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.009  1           1.004
## Year             1.009 16           1.000
```



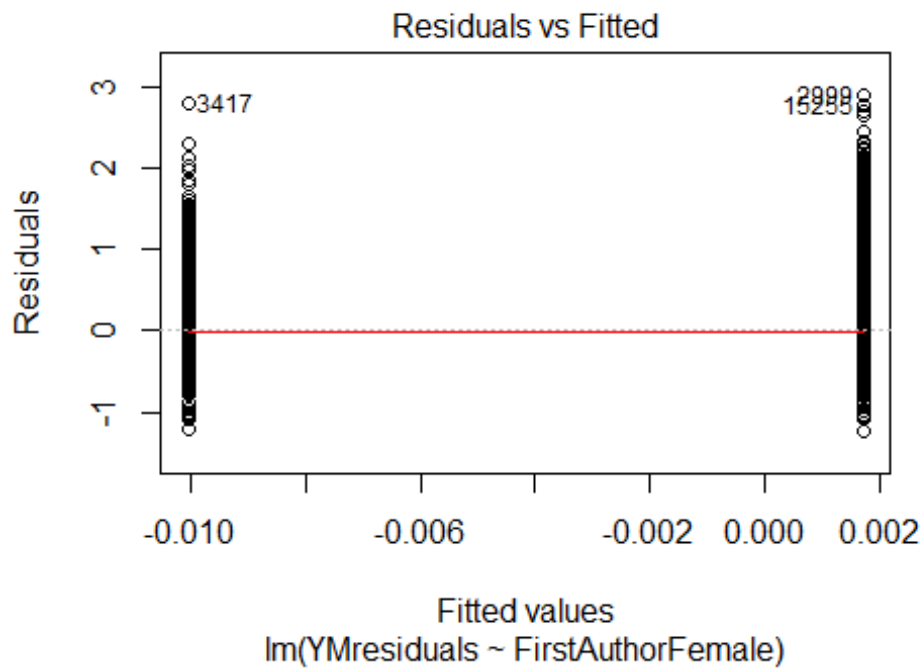

```
## [1] "List of  1 outliers with residuals above  2.5"
##          ScopusId NLCS Year OneField Fields residuals
## 11359 29144523061 3.76 2006     1700      7     2.594
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##    Min     1Q Median     3Q    Max
## -1.290 -0.437  0.025  0.431  2.590
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.07846    0.02981   36.18  < 2e-16 ***
## LastAuthorFemale1 -0.00204    0.01875   -0.11  0.91336
## Year1997          -0.05591    0.04309   -1.30  0.19444
## Year1998          -0.01027    0.04386   -0.23  0.81478
## Year1999           0.06548    0.04033    1.62  0.10452
## Year2000           0.10429    0.03953    2.64  0.00835 **
## Year2001           0.15872    0.04353    3.65  0.00027 ***
## Year2002           0.10798    0.04008    2.69  0.00707 **
## Year2003           0.12577    0.04027    3.12  0.00179 **
## Year2004           0.21170    0.03999    5.29  1.2e-07 ***
## Year2005           0.15855    0.03981    3.98  6.9e-05 ***
## Year2006           0.09190    0.04022    2.29  0.02233 *
```

```
## Year2007           0.11324    0.03810    2.97  0.00296 **
## Year2008           0.05542    0.03864    1.43  0.15149
## Year2009          -0.01048    0.04028   -0.26  0.79470
## Year2010           0.00591    0.04011    0.15  0.88278
## Year2011          -0.04453    0.03909   -1.14  0.25458
## Year2012          -0.00294    0.04140   -0.07  0.94330
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.644
## Multiple R-squared:  0.0135, Adjusted R-squared:  0.0117
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  772 weights are ~= 1. The remaining 8859 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0694  0.8740  0.9510  0.9120  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.04e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Sample size for the above analysis:  9631"
## [1] ""
## [1] ""
## [1] "####################################"
## [1] "Analysis of AJSC 1804"
## [1] "####################################"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  776  843  785  852  820  707  766  699  693  813  827  950 1088 1161 1086
## 2011 2012
## 1132 1069
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  507  544  517  575  552  430  499  444  427  515  488  558  685  674  642
## 2011 2012
```

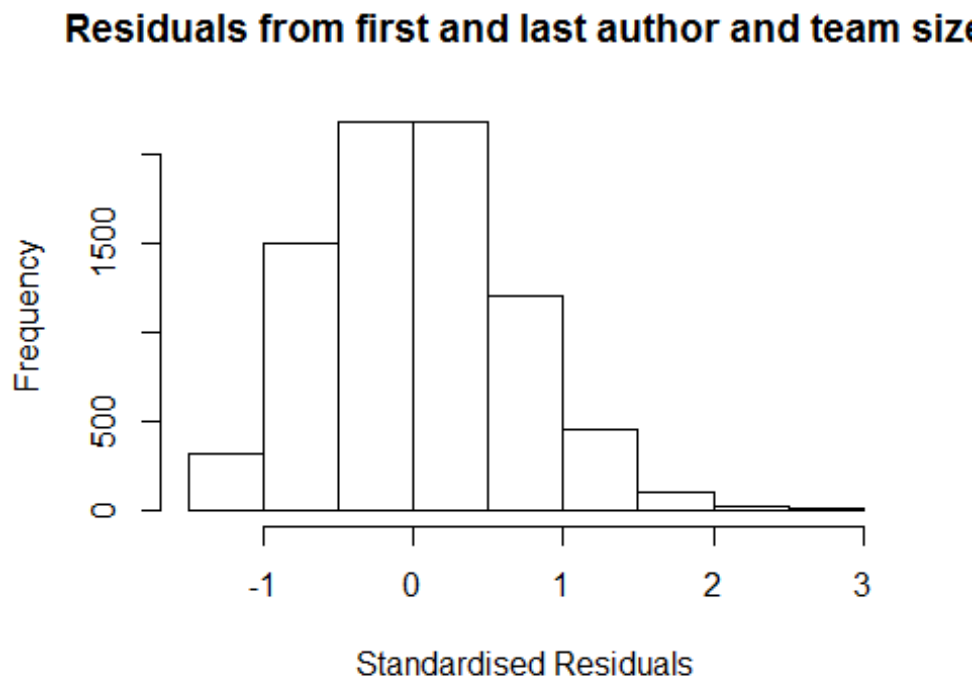
```
##  669  630
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  451  469  459  513  490  356  423  390  363  438  417  494  573  565  510
## 2011 2012
##  540  533
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
##  Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 89, df = 16, p-value = 5e-12
```




```
##
##  Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 0.77, df = 1, p-value = 0.4
```



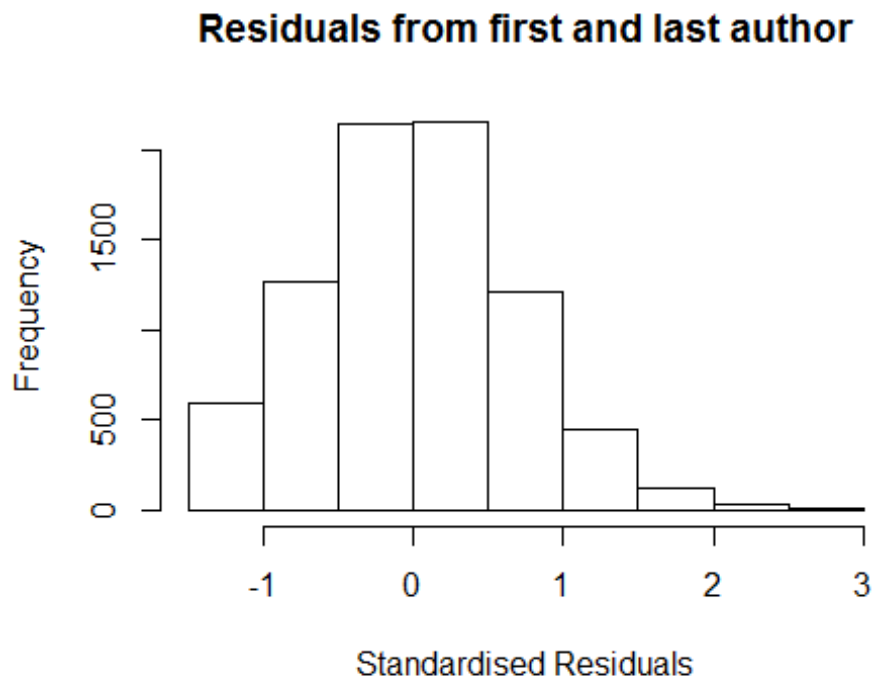

```
## [1] "Female first author team size 2018 geometric mean: 2.72155754955865"
## [1] "Male first author team size 2018 geometric mean: 2.01333541374583"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 33000, p-value = 4e-06
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 2.52697522126878"
## [1] "Male last author team size 2018 geometric mean: 2.07378697991921"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 28000, p-value = 0.002
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##                     GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.287  1           1.135
## LastAuthorFemale  1.274  1           1.129
## UniqueAuthors     1.070  4           1.008
## Year              1.069 16           1.002
```




```
## [1] "List of  6 outliers with residuals above  2.5"
##          ScopusId  NLCS Year OneField Fields residuals
## 2780   0032377357 3.789 1998     1804      2     2.649
## 2999   1542532754 3.921 1999     1804      2     2.821
## 3417   0001259111 3.825 1999     1804      3     2.560
## 4915   0035470889 3.537 2001     1804      2     2.631
## 13792 58449116301 3.389 2009     1804      4     2.509
## 15255 77950537175 3.862 2010     1712      3     2.636
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
UniqueAuthors +
##     Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##     k.max = 1000))
##  \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.40143 -0.44281 -0.00315  0.44736  2.82141
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)          1.0741     0.0353   30.39  < 2e-16 ***
## FirstAuthorFemale1  -0.0449     0.0233   -1.93  0.05374 .
## LastAuthorFemale1   -0.0447     0.0243   -1.84  0.06623 .
## UniqueAuthors2       0.2513     0.0169   14.88  < 2e-16 ***
## UniqueAuthors3       0.3274     0.0228   14.35  < 2e-16 ***
```

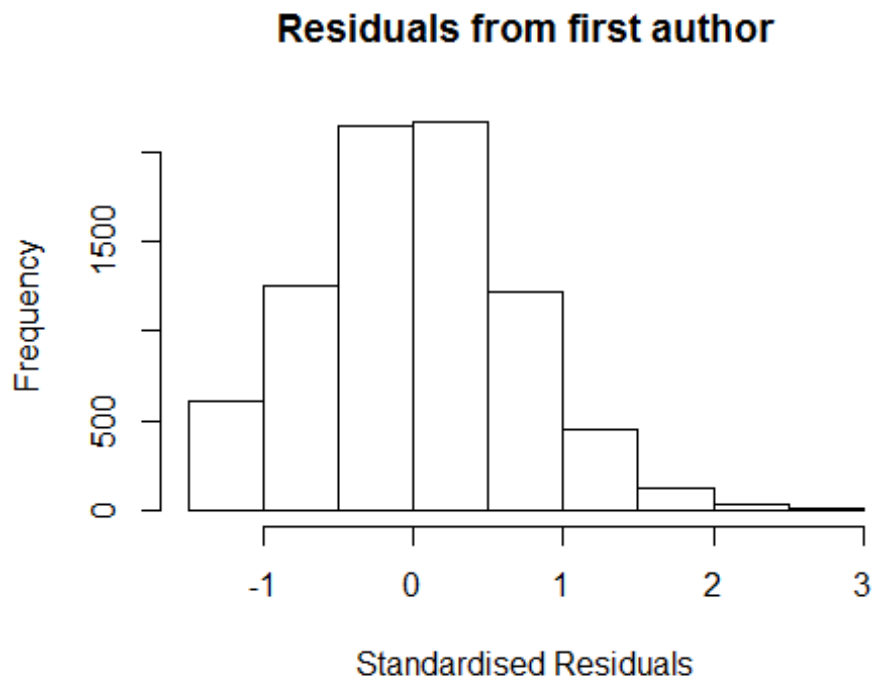
```
## UniqueAuthors4       0.4615     0.0418   11.05  < 2e-16 ***
## UniqueAuthors5       0.4446     0.0486    9.16  < 2e-16 ***
## Year1997            -0.1515     0.0483   -3.14  0.00171 **
## Year1998            -0.1852     0.0489   -3.79  0.00015 ***
## Year1999            -0.2258     0.0462   -4.89  1.0e-06 ***
## Year2000            -0.1693     0.0462   -3.66  0.00025 ***
## Year2001            -0.1678     0.0501   -3.35  0.00082 ***
## Year2002            -0.1475     0.0466   -3.16  0.00156 **
## Year2003            -0.1725     0.0458   -3.77  0.00017 ***
## Year2004            -0.1568     0.0483   -3.25  0.00116 **
## Year2005            -0.2579     0.0457   -5.64  1.8e-08 ***
## Year2006            -0.1986     0.0455   -4.36  1.3e-05 ***
## Year2007            -0.2421     0.0440   -5.50  3.9e-08 ***
## Year2008            -0.1962     0.0431   -4.55  5.5e-06 ***
## Year2009            -0.1938     0.0436   -4.44  9.0e-06 ***
## Year2010            -0.1755     0.0459   -3.83  0.00013 ***
## Year2011            -0.1949     0.0457   -4.27  2.0e-05 ***
## Year2012            -0.2622     0.0480   -5.46  4.9e-08 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.66
## Multiple R-squared:  0.0568, Adjusted R-squared:  0.0542
## Convergence in 10 IRWLS iterations
##
## Robustness weights:
##  654 weights are ~= 1. The remaining 7330 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.028   0.867   0.950   0.912   0.986   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.25e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.244  1           1.115
```

```
## LastAuthorFemale  1.239  1           1.113
## Year              1.021 16           1.001
```



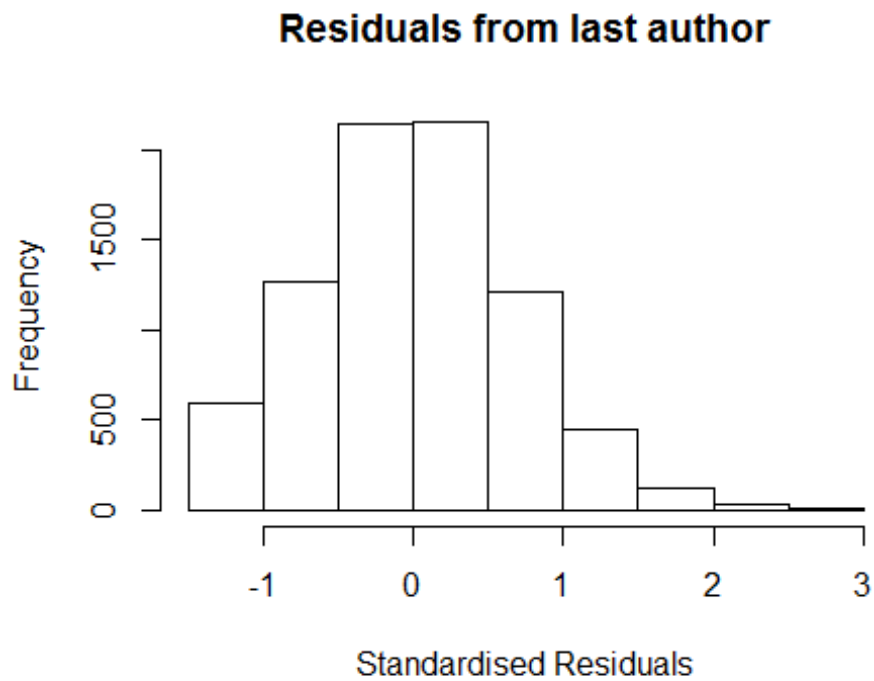

```
## [1] "List of  5 outliers with residuals above  2.5"
##          ScopusId  NLCS Year OneField Fields residuals
## 2780   0032377357 3.789 1998     1804      2     2.773
## 2999   1542532754 3.921 1999     1804      2     2.958
## 3417   0001259111 3.825 1999     1804      3     2.865
## 7148   0142121516 3.710 2003     1804      2     2.662
## 15255 77950537175 3.862 2010     1712      3     2.794
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.18690 -0.46657 -0.00262  0.45292  2.95782
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.18690    0.03603   32.95  < 2e-16 ***
## FirstAuthorFemale1 -0.00331    0.02363   -0.14  0.88861
## LastAuthorFemale1  -0.02892    0.02468   -1.17  0.24142
## Year1997           -0.13697    0.05008   -2.74  0.00625 **
## Year1998           -0.17064    0.05080   -3.36  0.00079 ***
```

```
## Year1999           -0.22372    0.04719   -4.74  2.2e-06 ***
## Year2000           -0.16807    0.04771   -3.52  0.00043 ***
## Year2001           -0.13882    0.05145   -2.70  0.00699 **
## Year2002           -0.12251    0.04786   -2.56  0.01048 *
## Year2003           -0.13933    0.04691   -2.97  0.00298 **
## Year2004           -0.11634    0.04948   -2.35  0.01873 *
## Year2005           -0.22290    0.04751   -4.69  2.8e-06 ***
## Year2006           -0.13298    0.04671   -2.85  0.00442 **
## Year2007           -0.19445    0.04552   -4.27  2.0e-05 ***
## Year2008           -0.15166    0.04458   -3.40  0.00067 ***
## Year2009           -0.14576    0.04482   -3.25  0.00115 **
## Year2010           -0.11937    0.04708   -2.54  0.01124 *
## Year2011           -0.12877    0.04724   -2.73  0.00642 **
## Year2012           -0.19280    0.04886   -3.95  8.0e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.681
## Multiple R-squared:  0.0056, Adjusted R-squared:  0.00335
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  667 weights are ~= 1. The remaining 7317 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0194  0.8720  0.9510  0.9120  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.25e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.011  1           1.006
## Year              1.011 16           1.000
```




```
## [1] "List of  5 outliers with residuals above  2.5"
##          ScopusId  NLCS Year OneField Fields residuals
## 2780   0032377357 3.789 1998     1804      2     2.773
## 2999   1542532754 3.921 1999     1804      2     2.958
## 3417   0001259111 3.825 1999     1804      3     2.865
## 7148   0142121516 3.710 2003     1804      2     2.662
## 15255 77950537175 3.862 2010     1712      3     2.794
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.18513 -0.46561 -0.00239  0.45338  2.96003
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)           1.1851     0.0360   32.95  < 2e-16 ***
## FirstAuthorFemale1   -0.0154     0.0214   -0.72  0.47221
## Year1997             -0.1371     0.0501   -2.74  0.00622 **
## Year1998             -0.1708     0.0508   -3.36  0.00078 ***
## Year1999             -0.2242     0.0472   -4.75  2.1e-06 ***
## Year2000             -0.1684     0.0477   -3.53  0.00042 ***
## Year2001             -0.1387     0.0515   -2.69  0.00707 **
## Year2002             -0.1228     0.0478   -2.57  0.01030 *
```

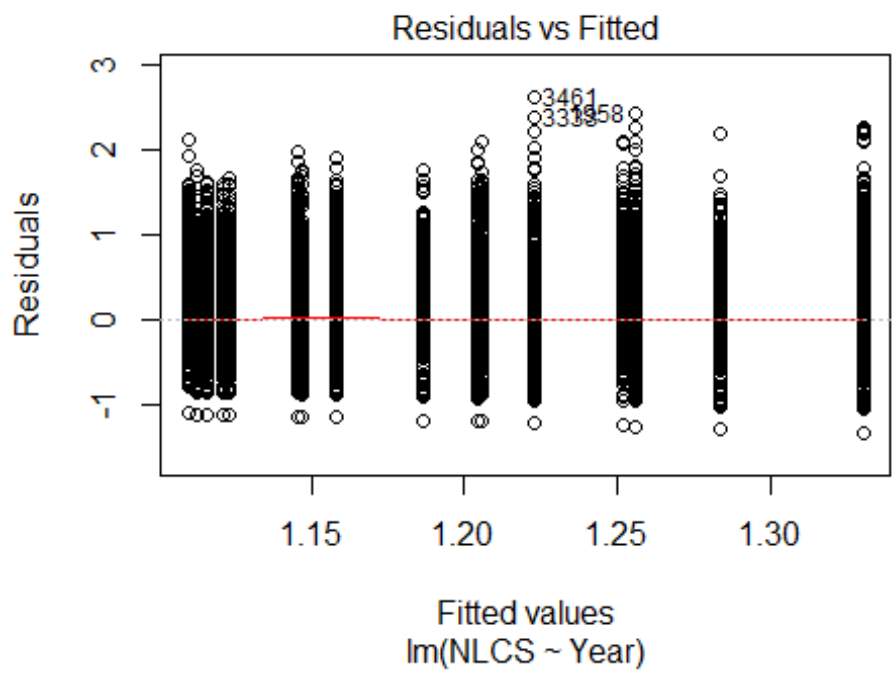
```
## Year2003            -0.1401     0.0469   -2.99  0.00280 **
## Year2004            -0.1168     0.0495   -2.36  0.01818 *
## Year2005            -0.2228     0.0475   -4.69  2.7e-06 ***
## Year2006            -0.1337     0.0467   -2.86  0.00421 **
## Year2007            -0.1957     0.0455   -4.30  1.7e-05 ***
## Year2008            -0.1517     0.0446   -3.40  0.00067 ***
## Year2009            -0.1460     0.0448   -3.26  0.00113 **
## Year2010            -0.1201     0.0471   -2.55  0.01075 *
## Year2011            -0.1295     0.0472   -2.74  0.00611 **
## Year2012            -0.1922     0.0488   -3.93  8.4e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.681
## Multiple R-squared:  0.00542,    Adjusted R-squared:  0.0033
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  671 weights are ~= 1. The remaining 7313 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0192  0.8720  0.9510  0.9120  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.25e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                    psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##                   GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.007  1           1.004
## Year             1.007 16           1.000
```



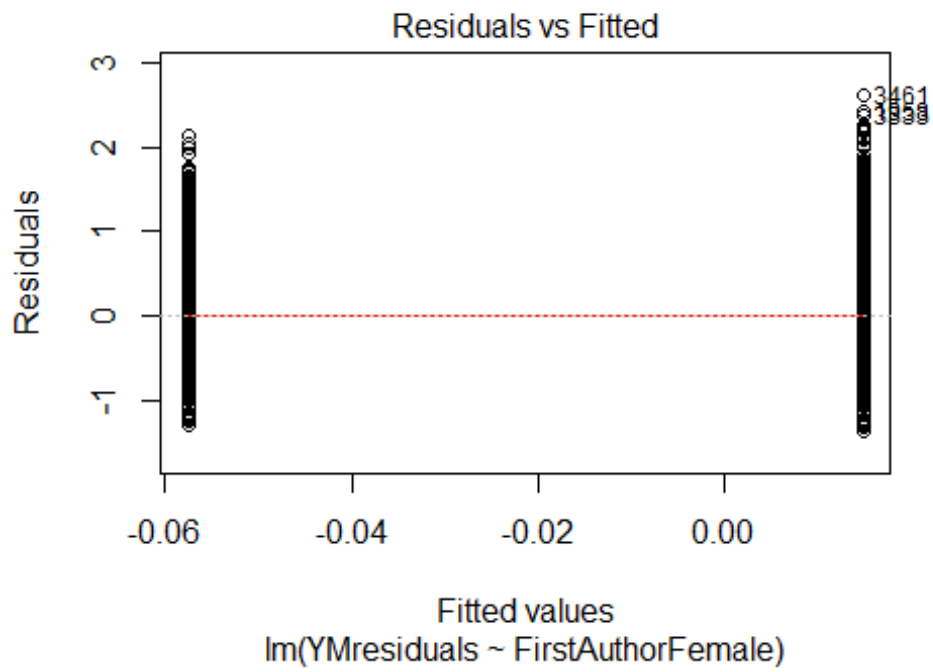

```
## [1] "List of  5 outliers with residuals above  2.5"
##          ScopusId  NLCS Year OneField Fields residuals
## 2780   0032377357 3.789 1998     1804      2     2.773
## 2999   1542532754 3.921 1999     1804      2     2.958
## 3417   0001259111 3.825 1999     1804      3     2.865
## 7148   0142121516 3.710 2003     1804      2     2.662
## 15255 77950537175 3.862 2010     1712      3     2.794
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.18664 -0.46618 -0.00381  0.45309  2.95804
##
## Coefficients:
##                   Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.1866     0.0360   33.00  < 2e-16 ***
## LastAuthorFemale1  -0.0304     0.0223   -1.37  0.17193
## Year1997           -0.1370     0.0501   -2.74  0.00624 **
## Year1998           -0.1707     0.0508   -3.36  0.00078 ***
## Year1999           -0.2237     0.0472   -4.74  2.2e-06 ***
## Year2000           -0.1680     0.0477   -3.52  0.00043 ***
## Year2001           -0.1389     0.0515   -2.70  0.00697 **
## Year2002           -0.1226     0.0479   -2.56  0.01045 *
```

```
## Year2003           -0.1395     0.0469   -2.97  0.00296 **
## Year2004           -0.1164     0.0495   -2.35  0.01870 *
## Year2005           -0.2229     0.0475   -4.69  2.8e-06 ***
## Year2006           -0.1330     0.0467   -2.85  0.00442 **
## Year2007           -0.1944     0.0455   -4.27  2.0e-05 ***
## Year2008           -0.1518     0.0446   -3.41  0.00066 ***
## Year2009           -0.1459     0.0448   -3.25  0.00114 **
## Year2010           -0.1195     0.0471   -2.54  0.01118 *
## Year2011           -0.1288     0.0472   -2.73  0.00642 **
## Year2012           -0.1931     0.0488   -3.96  7.7e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.68
## Multiple R-squared:  0.0056, Adjusted R-squared:  0.00348
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  673 weights are ~= 1. The remaining 7311 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0193  0.8720  0.9510  0.9120  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.25e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Sample size for the above analysis:  7984"
## [1] ""
## [1] ""
## [1] "##################################"
## [1] "Analysis of AJSC 1900"
## [1] "##################################"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 1395 1409 1163  936  991 1113  978 1065  980  983  964  980 1127 1135 1046
## 2011 2012
## 1168 1114
```

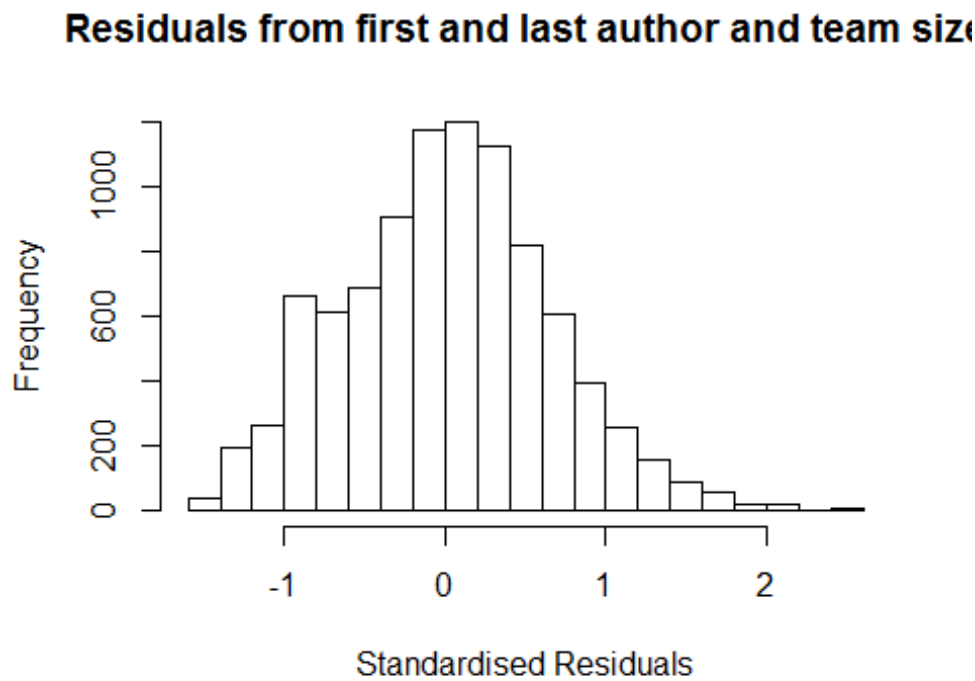
```
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  774  756  661  529  422  459  589  634  560  596  623  619  718  676  653
## 2011 2012
##  742  708
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  695  687  574  464  366  406  516  550  482  493  510  499  609  593  564
## 2011 2012
##  647  602
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
##  Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 100, df = 16, p-value = 2e-14
```




```
##
##  Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 0.3, df = 1, p-value = 0.6
```

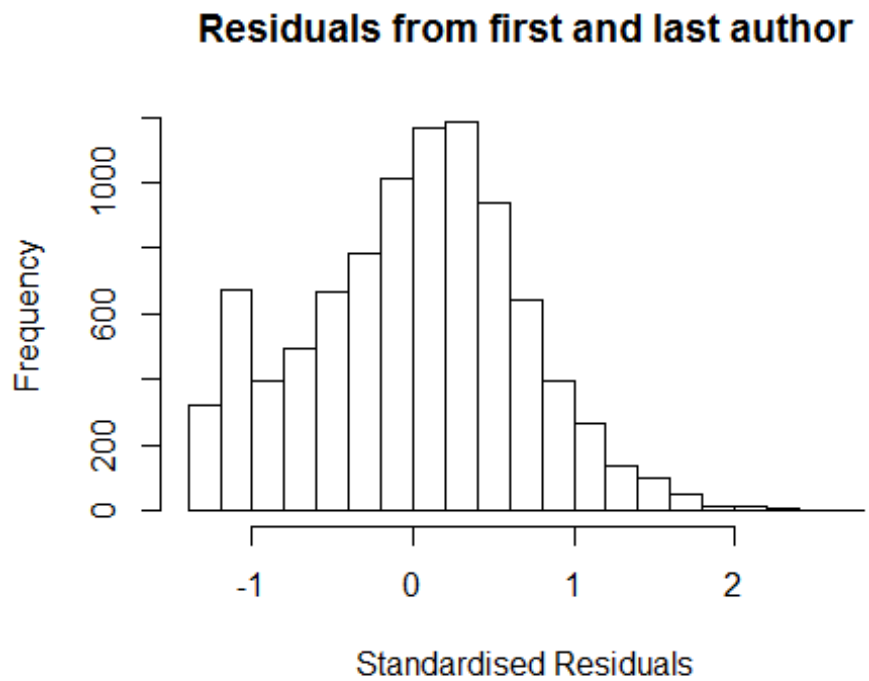



```
## [1] "Female first author team size 2018 geometric mean: 3.1351991562037"
## [1] "Male first author team size 2018 geometric mean: 2.92426245170664"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 51000, p-value = 0.2
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 2.80773874784484"
## [1] "Male last author team size 2018 geometric mean: 3.03863449701992"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 38000, p-value = 0.3
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.115  1           1.056
## LastAuthorFemale  1.117  1           1.057
## UniqueAuthors     1.054  4           1.007
## Year              1.064 16           1.002
```



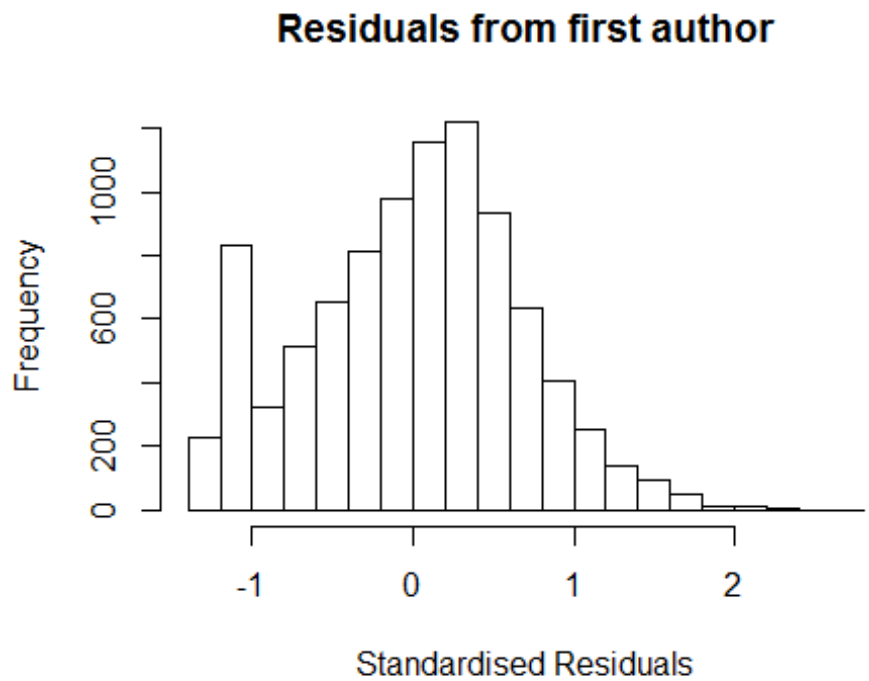

```
## [1] "List of  2 outliers with residuals above  2.5"
##         ScopusId  NLCS Year OneField Fields residuals
## 1243  0030443129 3.584 1996     1900      2     2.507
## 3461 33747921774 3.844 1998     1900      2     2.561
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
UniqueAuthors +
##     Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##     k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.5891 -0.4419  0.0162  0.4189  2.5606
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)          1.0769     0.0303   35.57  < 2e-16 ***
## FirstAuthorFemale1  -0.0485     0.0177   -2.73  0.00627 **
## LastAuthorFemale1   -0.0715     0.0194   -3.68  0.00023 ***
## UniqueAuthors2       0.3504     0.0196   17.85  < 2e-16 ***
## UniqueAuthors3       0.4243     0.0207   20.46  < 2e-16 ***
## UniqueAuthors4       0.5091     0.0239   21.29  < 2e-16 ***
## UniqueAuthors5       0.5837     0.0224   26.02  < 2e-16 ***
## Year1997            -0.0952     0.0379   -2.51  0.01213 *
## Year1998            -0.1439     0.0403   -3.57  0.00036 ***
```

```
## Year1999            -0.1526     0.0408   -3.74  0.00019 ***
## Year2000            -0.1685     0.0442   -3.82  0.00014 ***
## Year2001            -0.0840     0.0423   -1.99  0.04702 *
## Year2002            -0.2687     0.0381   -7.05  1.9e-12 ***
## Year2003            -0.2816     0.0379   -7.44  1.1e-13 ***
## Year2004            -0.2751     0.0377   -7.29  3.3e-13 ***
## Year2005            -0.2421     0.0381   -6.35  2.2e-10 ***
## Year2006            -0.2694     0.0387   -6.97  3.5e-12 ***
## Year2007            -0.2720     0.0401   -6.79  1.2e-11 ***
## Year2008            -0.2601     0.0406   -6.41  1.5e-10 ***
## Year2009            -0.2251     0.0385   -5.85  5.1e-09 ***
## Year2010            -0.1581     0.0393   -4.03  5.7e-05 ***
## Year2011            -0.2078     0.0387   -5.37  7.9e-08 ***
## Year2012            -0.3009     0.0401   -7.50  6.9e-14 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.638
## Multiple R-squared:  0.11,   Adjusted R-squared:  0.108
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
##  764 weights are ~= 1. The remaining 8493 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0706  0.8610  0.9510  0.9070  0.9870  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.08e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.090  1           1.044
## LastAuthorFemale  1.089  1           1.044
## Year              1.017 16           1.001
```



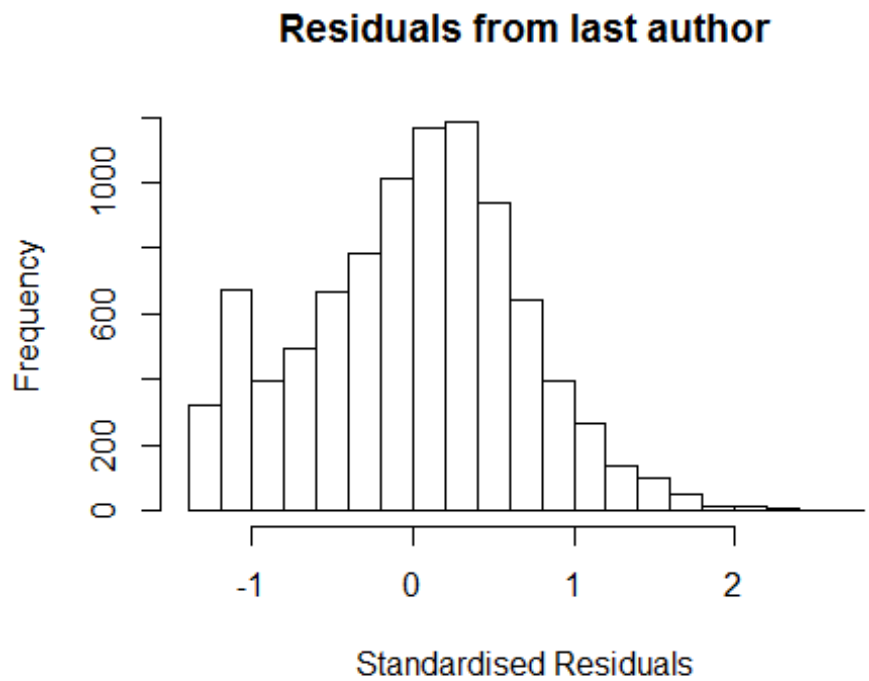

```
## [1] "List of  1 outliers with residuals above  2.5"
##         ScopusId  NLCS Year OneField Fields residuals
## 3461 33747921774 3.844 1998     1900      2     2.624
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.3373 -0.4669  0.0454  0.4450  2.6238
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)           1.3373     0.0275   48.64  < 2e-16 ***
## FirstAuthorFemale1   -0.0417     0.0185   -2.25  0.02436 *
## LastAuthorFemale1    -0.0955     0.0204   -4.68  2.9e-06 ***
## Year1997             -0.0890     0.0384   -2.32  0.02054 *
## Year1998             -0.1171     0.0413   -2.84  0.00459 **
## Year1999             -0.1365     0.0415   -3.29  0.00101 **
## Year2000             -0.1681     0.0451   -3.72  0.00020 ***
## Year2001             -0.0385     0.0428   -0.90  0.36891
## Year2002             -0.2241     0.0397   -5.64  1.7e-08 ***
## Year2003             -0.2154     0.0387   -5.56  2.8e-08 ***
## Year2004             -0.2143     0.0392   -5.47  4.7e-08 ***
## Year2005             -0.1776     0.0391   -4.54  5.7e-06 ***
```

```
## Year2006            -0.1911     0.0403   -4.74  2.2e-06 ***
## Year2007            -0.1823     0.0416   -4.38  1.2e-05 ***
## Year2008            -0.2048     0.0426   -4.81  1.5e-06 ***
## Year2009            -0.1550     0.0409   -3.79  0.00015 ***
## Year2010            -0.0850     0.0405   -2.10  0.03606 *
## Year2011            -0.1274     0.0401   -3.18  0.00150 **
## Year2012            -0.2253     0.0431   -5.22  1.8e-07 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.662
## Multiple R-squared:  0.0144, Adjusted R-squared:  0.0124
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  793 weights are ~= 1. The remaining 8464 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0805  0.8560  0.9500  0.9040  0.9840  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.08e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.009  1           1.004
## Year              1.009 16           1.000
```




```
## [1] "List of  1 outliers with residuals above  2.5"
##         ScopusId  NLCS Year OneField Fields residuals
## 3461 33747921774 3.844 1998     1900      2     2.624
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.3289 -0.4667  0.0479  0.4445  2.6324
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)           1.3289     0.0276   48.18  < 2e-16 ***
## FirstAuthorFemale1   -0.0731     0.0182   -4.02  5.8e-05 ***
## Year1997             -0.0895     0.0385   -2.33  0.01995 *
## Year1998             -0.1173     0.0414   -2.83  0.00460 **
## Year1999             -0.1415     0.0415   -3.41  0.00065 ***
## Year2000             -0.1675     0.0454   -3.69  0.00022 ***
## Year2001             -0.0385     0.0430   -0.90  0.37041
## Year2002             -0.2273     0.0398   -5.71  1.2e-08 ***
## Year2003             -0.2197     0.0387   -5.67  1.5e-08 ***
## Year2004             -0.2170     0.0394   -5.51  3.8e-08 ***
## Year2005             -0.1808     0.0391   -4.62  3.9e-06 ***
## Year2006             -0.1942     0.0404   -4.81  1.5e-06 ***
```

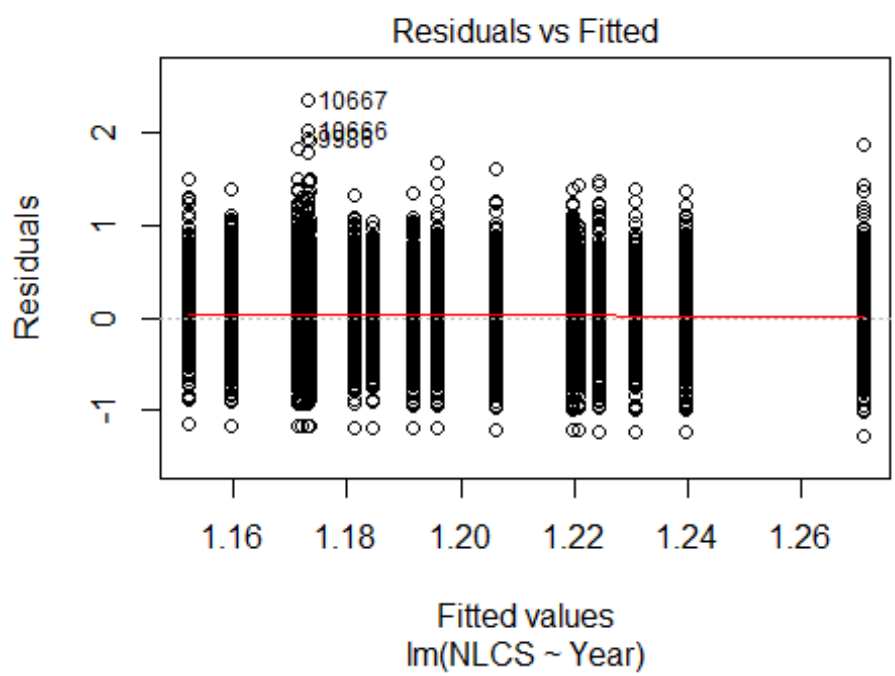
```
## Year2007            -0.1867     0.0418   -4.46  8.2e-06 ***
## Year2008            -0.2062     0.0429   -4.81  1.5e-06 ***
## Year2009            -0.1563     0.0411   -3.80  0.00014 ***
## Year2010            -0.0895     0.0405   -2.21  0.02730 *
## Year2011            -0.1303     0.0402   -3.24  0.00121 **
## Year2012            -0.2275     0.0434   -5.25  1.6e-07 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.662
## Multiple R-squared:  0.0118, Adjusted R-squared:  0.00997
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  789 weights are ~= 1. The remaining 8468 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0781  0.8580  0.9500  0.9040  0.9840  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.08e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##                   GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.008  1           1.004
## Year             1.008 16           1.000
```



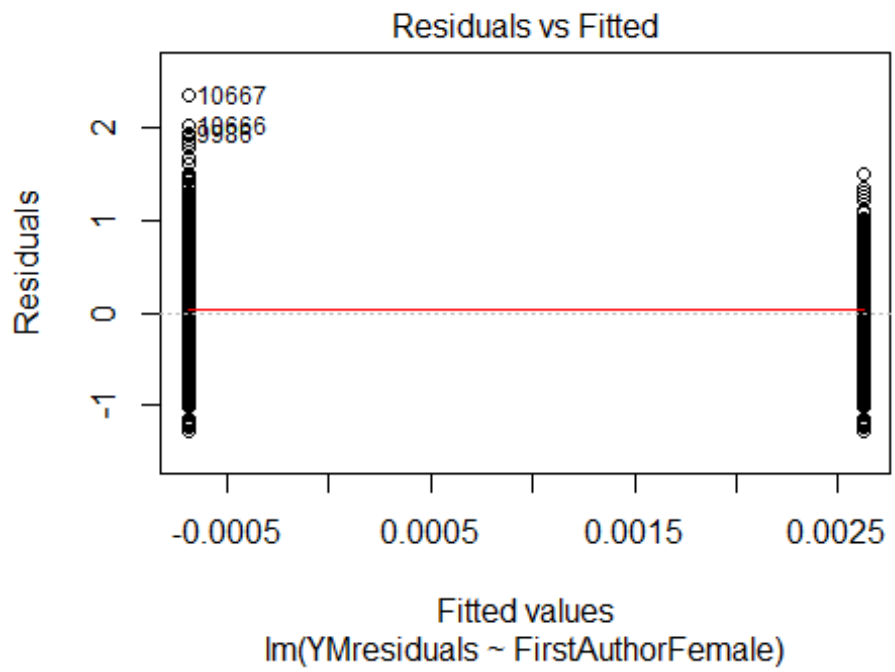

```
## [1] "List of  1 outliers with residuals above  2.5"
##         ScopusId  NLCS Year OneField Fields residuals
## 3461 33747921774 3.844 1998     1900      2     2.624
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##    Min     1Q Median     3Q    Max
## -1.332 -0.470  0.043  0.444  2.629
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.3319     0.0274   48.70  < 2e-16 ***
## LastAuthorFemale1  -0.1110     0.0198   -5.60  2.2e-08 ***
## Year1997           -0.0890     0.0384   -2.32  0.02054 *
## Year1998           -0.1170     0.0413   -2.83  0.00461 **
## Year1999           -0.1356     0.0415   -3.27  0.00109 **
## Year2000           -0.1673     0.0452   -3.70  0.00021 ***
## Year2001           -0.0382     0.0428   -0.89  0.37266
## Year2002           -0.2234     0.0397   -5.63  1.9e-08 ***
## Year2003           -0.2151     0.0387   -5.55  2.9e-08 ***
## Year2004           -0.2153     0.0392   -5.49  4.0e-08 ***
## Year2005           -0.1796     0.0391   -4.60  4.3e-06 ***
## Year2006           -0.1919     0.0403   -4.76  1.9e-06 ***
```

```
## Year2007           -0.1844     0.0417   -4.43  9.8e-06 ***
## Year2008           -0.2060     0.0425   -4.84  1.3e-06 ***
## Year2009           -0.1564     0.0408   -3.83  0.00013 ***
## Year2010           -0.0855     0.0405   -2.11  0.03490 *
## Year2011           -0.1295     0.0401   -3.23  0.00124 **
## Year2012           -0.2267     0.0431   -5.26  1.5e-07 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.662
## Multiple R-squared:  0.0138, Adjusted R-squared:  0.012
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  784 weights are ~= 1. The remaining 8473 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0793  0.8570  0.9500  0.9040  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.08e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Sample size for the above analysis:  9257"
## [1] ""
## [1] ""
## [1] "####################################"
## [1] "Analysis of AJSC 1901"
## [1] "####################################"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 1260 1467 1464 1351 1168 1435 1380 1551 1371 1353 1557 1637 1489 1154 1239
## 2011 2012
## 1195 1305
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  737  784  803  704  573  675  822  855  783  786  909  986  845  644  654
## 2011 2012
```

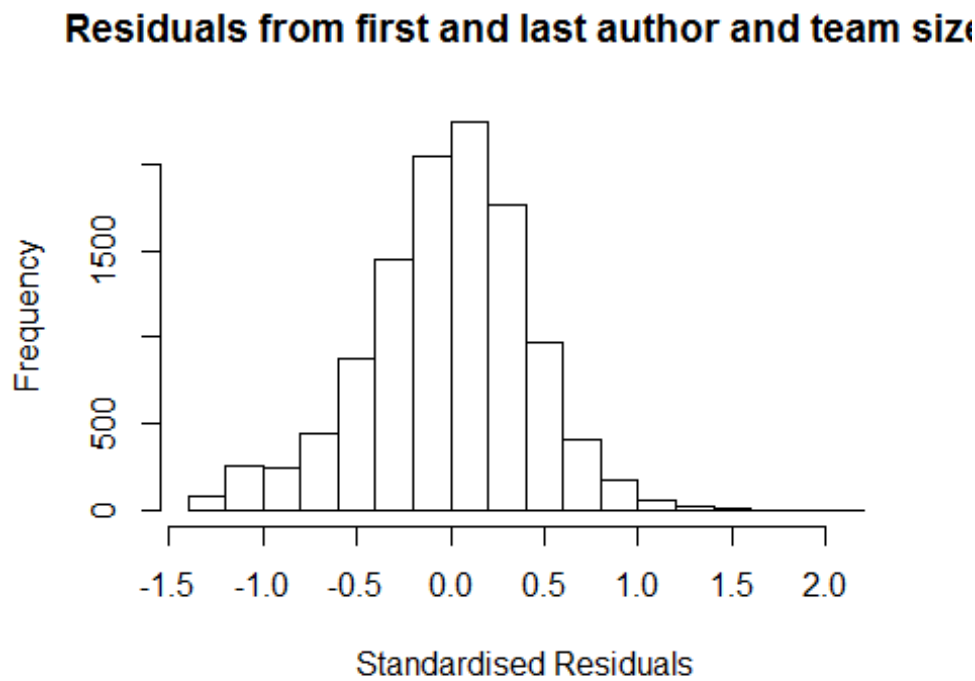
```
##  650  704
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  654  713  680  619  498  571  708  728  674  654  763  851  698  557  566
## 2011 2012
##  555  585
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
##  Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 110, df = 16, p-value = 5e-16
```




```
##
##  Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 9.1, df = 1, p-value = 0.003
```

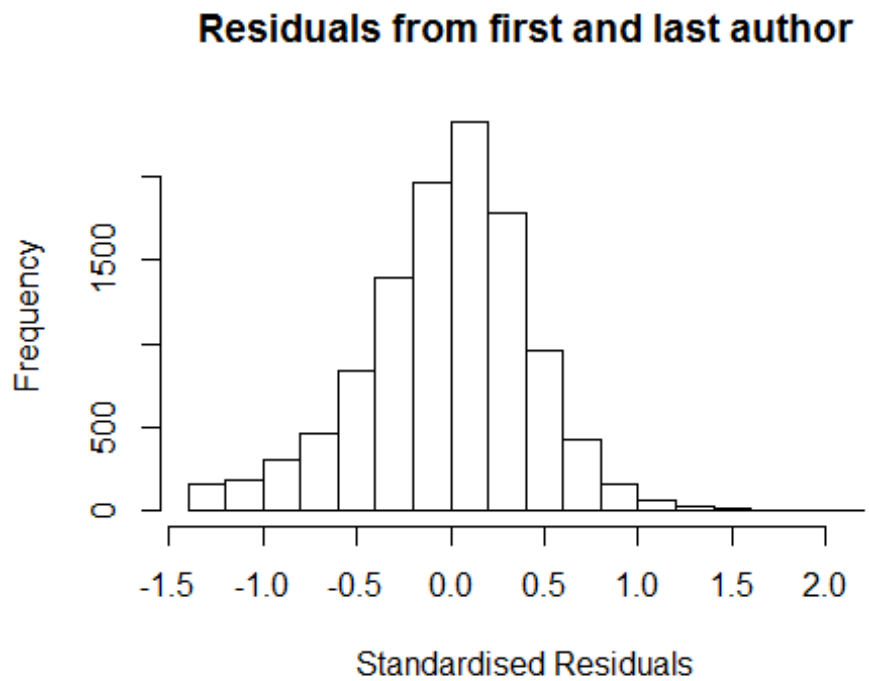



```
## [1] "Female first author team size 2018 geometric mean: 3.27470549467803"
## [1] "Male first author team size 2018 geometric mean: 2.99772249843164"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 20000, p-value = 0.4
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 2.93827558996241"
## [1] "Male last author team size 2018 geometric mean: 3.11042743499362"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 14000, p-value = 0.6
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.031  1           1.015
## LastAuthorFemale  1.019  1           1.009
## UniqueAuthors     1.098  4           1.012
## Year              1.099 16           1.003
```



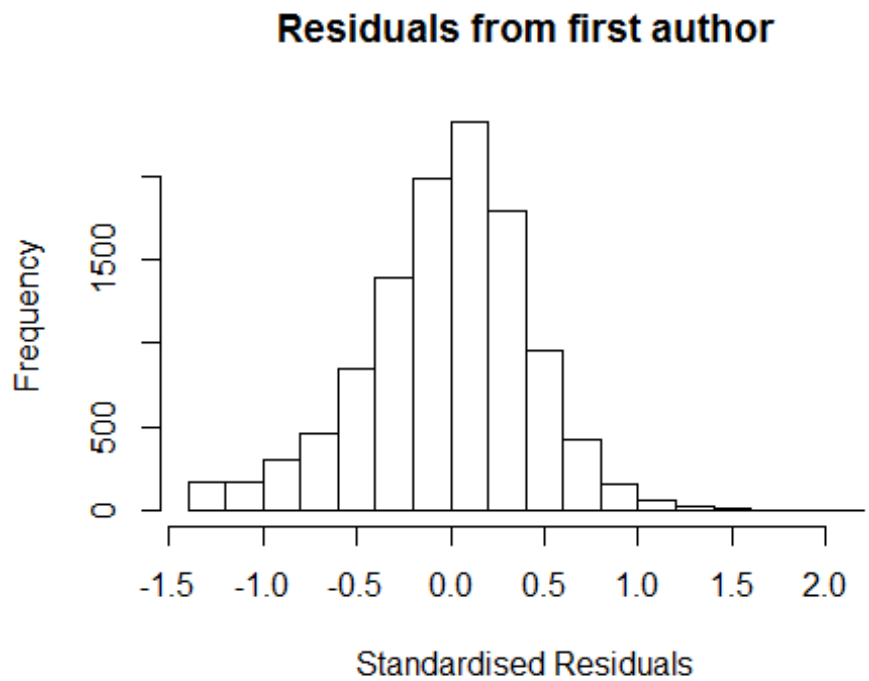

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
UniqueAuthors +
##     Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##     k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.3700 -0.2692  0.0126  0.2681  2.0886
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.14182    0.02013   56.72  < 2e-16 ***
## FirstAuthorFemale1 -0.00176    0.00967   -0.18  0.85583
## LastAuthorFemale1  -0.02684    0.01150   -2.33  0.01967 *
## UniqueAuthors2      0.17395    0.01415   12.29  < 2e-16 ***
## UniqueAuthors3      0.18928    0.01468   12.90  < 2e-16 ***
## UniqueAuthors4      0.22141    0.01595   13.88  < 2e-16 ***
## UniqueAuthors5      0.26290    0.01523   17.26  < 2e-16 ***
## Year1997           -0.03469    0.02490   -1.39  0.16358
## Year1998           -0.02777    0.02396   -1.16  0.24658
## Year1999           -0.04732    0.02458   -1.92  0.05427 .
```

```
## Year2000           -0.05227    0.02444   -2.14  0.03249 *
## Year2001           -0.08530    0.02519   -3.39  0.00071 ***
## Year2002           -0.12339    0.02583   -4.78  1.8e-06 ***
## Year2003           -0.10761    0.02279   -4.72  2.4e-06 ***
## Year2004           -0.13968    0.02332   -5.99  2.2e-09 ***
## Year2005           -0.08435    0.02412   -3.50  0.00047 ***
## Year2006           -0.13416    0.02357   -5.69  1.3e-08 ***
## Year2007           -0.12949    0.02308   -5.61  2.1e-08 ***
## Year2008           -0.10687    0.02326   -4.59  4.4e-06 ***
## Year2009           -0.08128    0.02346   -3.46  0.00053 ***
## Year2010           -0.10356    0.02479   -4.18  3.0e-05 ***
## Year2011           -0.15017    0.02434   -6.17  7.1e-10 ***
## Year2012           -0.13162    0.02518   -5.23  1.8e-07 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.395
## Multiple R-squared:  0.0469, Adjusted R-squared:  0.045
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
##  4 observations c(3871,4212,4220,10745)
##   are outliers with |weight| = 0 ( < 9e-06);
##  926 weights are ~= 1. The remaining 10144 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0012  0.8630  0.9500  0.8910  0.9860  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          9.03e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.021  1           1.011
## LastAuthorFemale  1.017  1           1.008
## Year              1.019 16           1.001
```



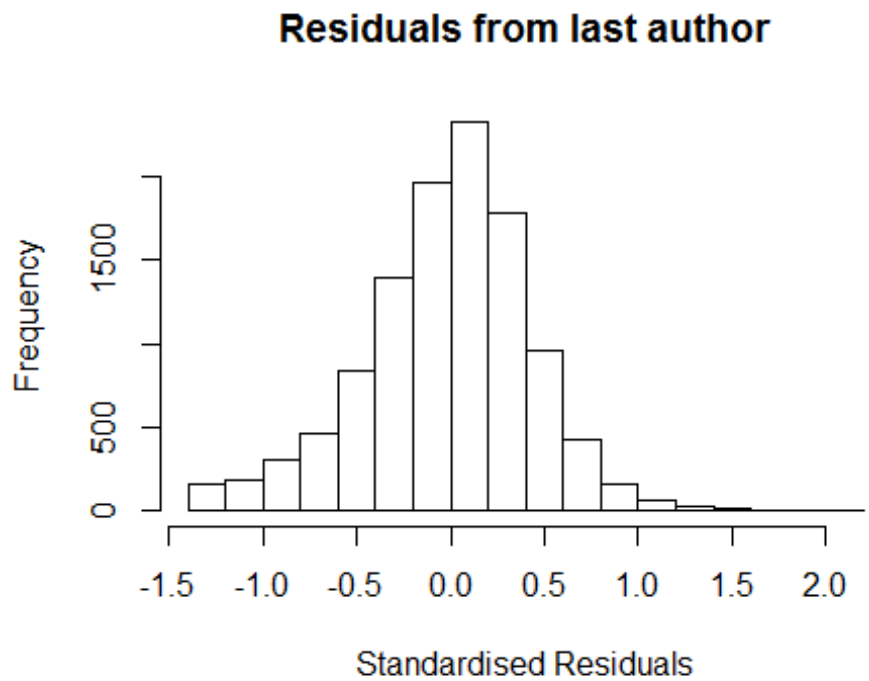

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.2792 -0.2731  0.0178  0.2690  2.0161
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)          1.27918    0.01720   74.38  < 2e-16 ***
## FirstAuthorFemale1  0.01194    0.00971    1.23  0.21916
## LastAuthorFemale1  -0.02556    0.01162   -2.20  0.02780 *
## Year1997           -0.04040    0.02486   -1.62  0.10420
## Year1998           -0.02097    0.02383   -0.88  0.37897
## Year1999           -0.04059    0.02462   -1.65  0.09925 .
## Year2000           -0.03405    0.02448   -1.39  0.16427
## Year2001           -0.05344    0.02503   -2.14  0.03277 *
## Year2002           -0.09228    0.02562   -3.60  0.00032 ***
## Year2003           -0.07417    0.02262   -3.28  0.00104 **
## Year2004           -0.10031    0.02326   -4.31  1.6e-05 ***
## Year2005           -0.05624    0.02412   -2.33  0.01974 *
```

```
## Year2006           -0.10168    0.02350   -4.33  1.5e-05 ***
## Year2007           -0.09569    0.02318   -4.13  3.7e-05 ***
## Year2008           -0.07065    0.02308   -3.06  0.00221 **
## Year2009           -0.03522    0.02315   -1.52  0.12825
## Year2010           -0.06892    0.02480   -2.78  0.00547 **
## Year2011           -0.10768    0.02448   -4.40  1.1e-05 ***
## Year2012           -0.08488    0.02506   -3.39  0.00071 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.397
## Multiple R-squared:  0.0059, Adjusted R-squared:  0.00428
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  4 observations c(283,3871,4212,10745)
##   are outliers with |weight| = 0 ( < 9e-06);
##  936 weights are ~= 1. The remaining 10134 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0029  0.8600  0.9500  0.8880  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          9.03e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.011  1           1.006
## Year              1.011 16           1.000
```




```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##    Min     1Q Median     3Q    Max
## -1.285 -0.271  0.018  0.268  2.019
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)          1.27659    0.01713   74.54  < 2e-16 ***
## FirstAuthorFemale1  0.00806    0.00972    0.83  0.40693
## Year1997            -0.04033    0.02486   -1.62  0.10484
## Year1998            -0.02086    0.02383   -0.88  0.38133
## Year1999            -0.04091    0.02461   -1.66  0.09650 .
## Year2000            -0.03447    0.02448   -1.41  0.15922
## Year2001            -0.05335    0.02502   -2.13  0.03299 *
## Year2002            -0.09255    0.02565   -3.61  0.00031 ***
## Year2003            -0.07476    0.02263   -3.30  0.00096 ***
## Year2004            -0.10071    0.02328   -4.33  1.5e-05 ***
## Year2005            -0.05651    0.02411   -2.34  0.01909 *
## Year2006            -0.10174    0.02348   -4.33  1.5e-05 ***
```

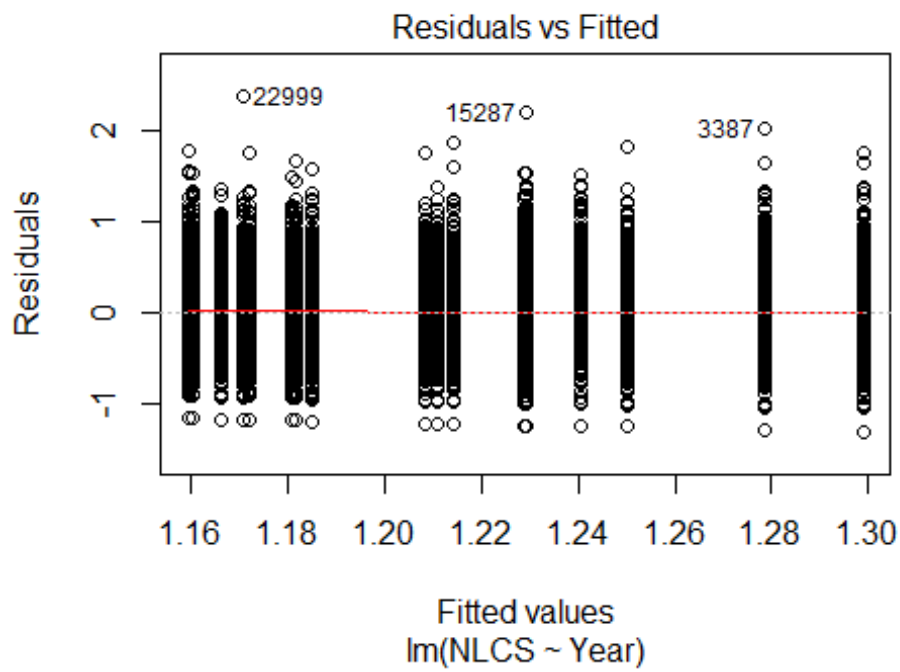
```
## Year2007           -0.09560    0.02318   -4.12  3.8e-05 ***
## Year2008           -0.07108    0.02306   -3.08  0.00205 **
## Year2009           -0.03599    0.02317   -1.55  0.12034
## Year2010           -0.06942    0.02479   -2.80  0.00511 **
## Year2011           -0.10906    0.02448   -4.46  8.5e-06 ***
## Year2012           -0.08618    0.02506   -3.44  0.00059 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.397
## Multiple R-squared:  0.00543,    Adjusted R-squared:  0.0039
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  4 observations c(283,3871,4212,10745)
##   are outliers with |weight| = 0 ( < 9e-06);
##  907 weights are ~= 1. The remaining 10163 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0026  0.8610  0.9500  0.8880  0.9860  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          9.03e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##                   GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.007  1           1.003
## Year             1.007 16           1.000
```



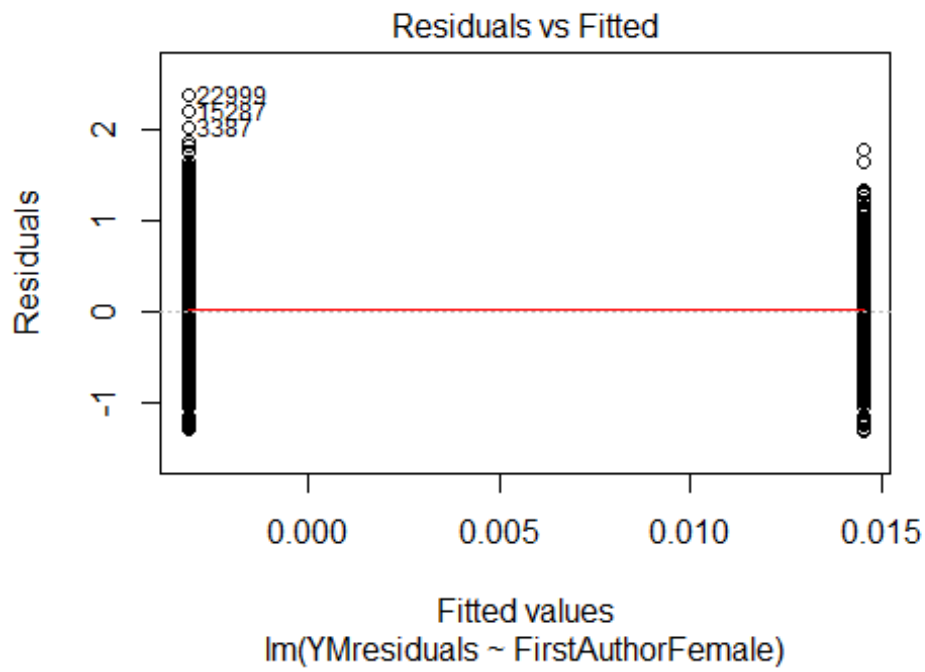

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.2805 -0.2728  0.0174  0.2695  2.0137
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.2805     0.0172   74.60  < 2e-16 ***
## LastAuthorFemale1  -0.0231     0.0116   -1.99  0.04607 *
## Year1997           -0.0399     0.0249   -1.60  0.10860
## Year1998           -0.0205     0.0238   -0.86  0.38934
## Year1999           -0.0400     0.0246   -1.63  0.10382
## Year2000           -0.0331     0.0245   -1.35  0.17547
## Year2001           -0.0528     0.0250   -2.11  0.03498 *
## Year2002           -0.0912     0.0256   -3.56  0.00037 ***
## Year2003           -0.0735     0.0226   -3.25  0.00115 **
## Year2004           -0.0991     0.0232   -4.27  2.0e-05 ***
## Year2005           -0.0553     0.0241   -2.30  0.02174 *
## Year2006           -0.1007     0.0235   -4.29  1.8e-05 ***
```

```
## Year2007           -0.0944     0.0231   -4.08  4.6e-05 ***
## Year2008           -0.0693     0.0230   -3.01  0.00264 **
## Year2009           -0.0343     0.0231   -1.48  0.13858
## Year2010           -0.0679     0.0248   -2.74  0.00618 **
## Year2011           -0.1063     0.0244   -4.35  1.4e-05 ***
## Year2012           -0.0834     0.0250   -3.33  0.00087 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.397
## Multiple R-squared:  0.00578,    Adjusted R-squared:  0.00425
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  4 observations c(283,3871,4212,10745)
##   are outliers with |weight| = 0 ( < 9e-06);
##  933 weights are ~= 1. The remaining 10137 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0032  0.8610  0.9500  0.8880  0.9860  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          9.03e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                  psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Sample size for the above analysis:  11074"
## [1] ""
## [1] ""
## [1] "##################################"
## [1] "Analysis of AJSC 1902"
## [1] "##################################"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 1783 1964 1945 1959 2005 2155 2064 2212 2082 2044 2072 2402 2445 2168 2231
## 2011 2012
## 2246 2434
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
```

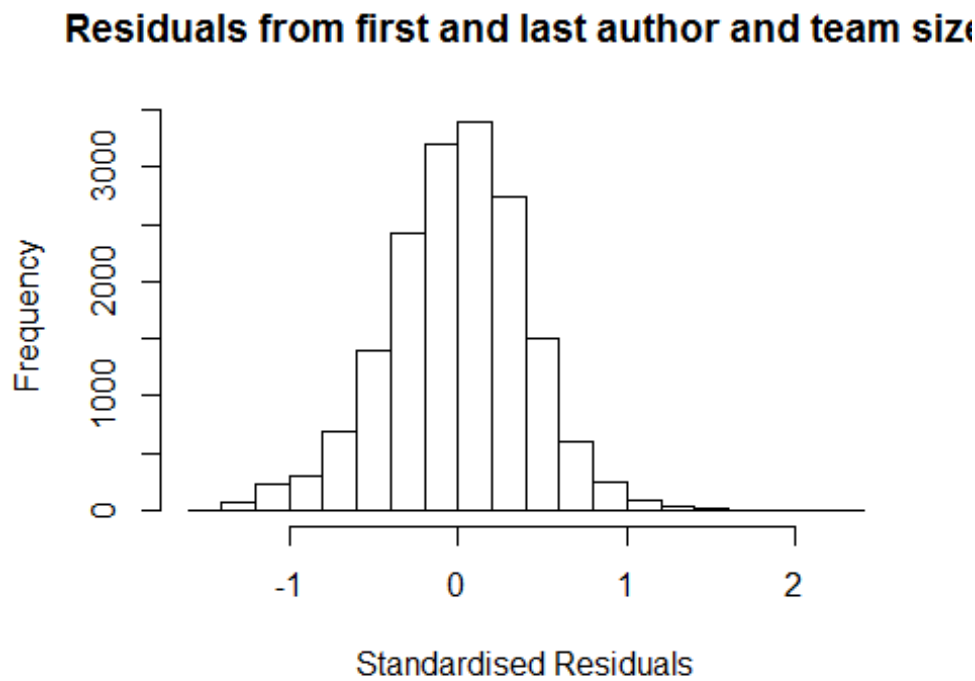
```
## 1056 1124 1101 1082  940  792 1241 1266 1216 1268 1250 1434 1393 1225 1195
## 2011 2012
## 1229 1355
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  920  999  952  931  810  665 1071 1054 1013 1026  980 1193 1128 1045  983
## 2011 2012
## 1027 1128
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
##  Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 110, df = 16, p-value = 4e-16
```




```
##
##  Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 35, df = 1, p-value = 3e-09
```

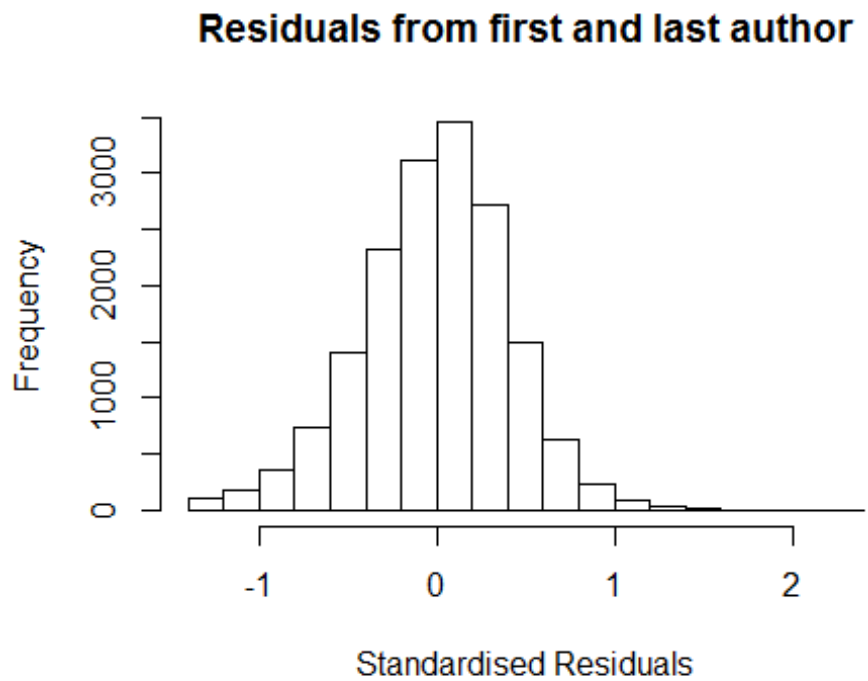



```
## [1] "Female first author team size 2018 geometric mean: 3.77586423193766"
## [1] "Male first author team size 2018 geometric mean: 3.22307147927268"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 180000, p-value = 3e-04
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 3.36290065635469"
## [1] "Male last author team size 2018 geometric mean: 3.39465189872483"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 120000, p-value = 0.9
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.039  1           1.019
## LastAuthorFemale  1.030  1           1.015
## UniqueAuthors     1.066  4           1.008
## Year              1.075 16           1.002
```



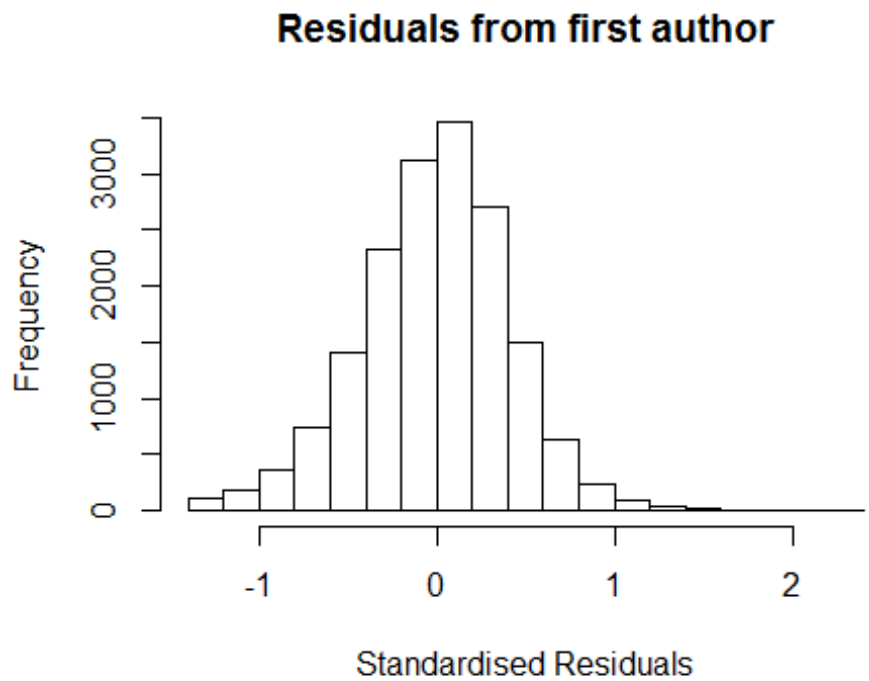

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
UniqueAuthors +
##     Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##     k.max = 1000))
##  \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.45325 -0.26356  0.00936  0.26277  2.37438
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.18660    0.01669   71.08  < 2e-16 ***
## FirstAuthorFemale1  0.01117    0.00793    1.41  0.15925
## LastAuthorFemale1  -0.00712    0.00925   -0.77  0.44126
## UniqueAuthors2      0.13043    0.01028   12.69  < 2e-16 ***
## UniqueAuthors3      0.16510    0.01088   15.18  < 2e-16 ***
## UniqueAuthors4      0.18574    0.01212   15.32  < 2e-16 ***
## UniqueAuthors5      0.26666    0.01129   23.62  < 2e-16 ***
## Year1997           -0.02991    0.02074   -1.44  0.14926
## Year1998           -0.04900    0.02062   -2.38  0.01748 *
## Year1999           -0.07457    0.02069   -3.60  0.00031 ***
```

```
## Year2000           -0.08560    0.02101   -4.07  4.6e-05 ***
## Year2001           -0.10298    0.02248   -4.58  4.6e-06 ***
## Year2002           -0.09396    0.01996   -4.71  2.5e-06 ***
## Year2003           -0.11447    0.01925   -5.95  2.8e-09 ***
## Year2004           -0.12897    0.01911   -6.75  1.6e-11 ***
## Year2005           -0.14741    0.02004   -7.36  2.0e-13 ***
## Year2006           -0.18009    0.01979   -9.10  < 2e-16 ***
## Year2007           -0.15681    0.01884   -8.32  < 2e-16 ***
## Year2008           -0.18327    0.01937   -9.46  < 2e-16 ***
## Year2009           -0.15911    0.01943   -8.19  2.8e-16 ***
## Year2010           -0.15763    0.02001   -7.88  3.5e-15 ***
## Year2011           -0.18341    0.01952   -9.40  < 2e-16 ***
## Year2012           -0.17096    0.01930   -8.86  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.39
## Multiple R-squared:  0.0493, Adjusted R-squared:  0.0481
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  3 observations c(4693,6081,9183)
##   are outliers with |weight| <= 4.7e-06 ( < 5.9e-06);
##  1415 weights are ~= 1. The remaining 15507 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0005  0.8710  0.9510  0.8990  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          5.91e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.031  1           1.015
## LastAuthorFemale  1.026  1           1.013
## Year              1.018 16           1.001
```



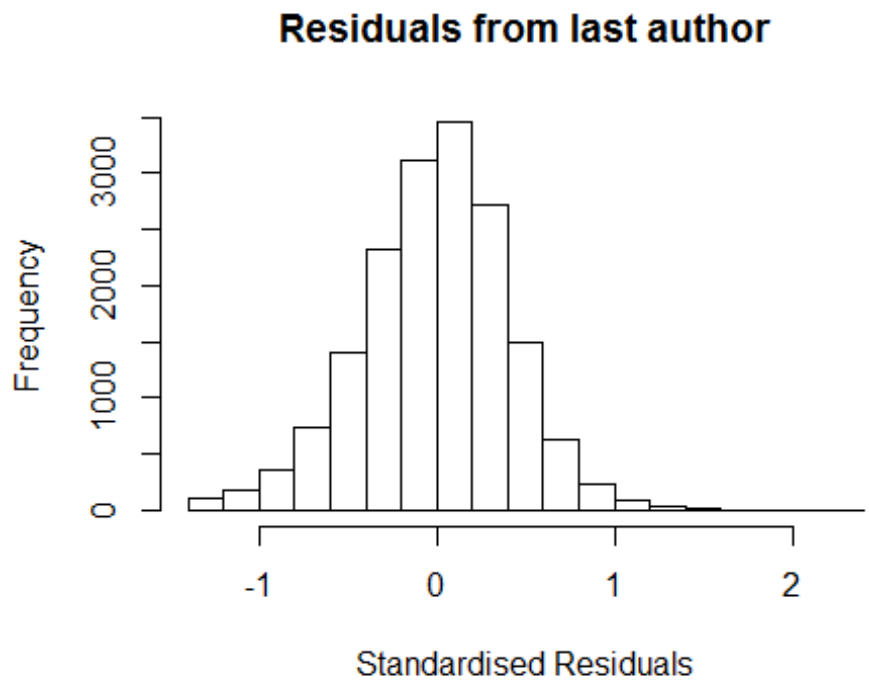

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.3027 -0.2660  0.0135  0.2666  2.3602
##
## Coefficients:
##                      Estimate Std. Error t value Pr(>|t|)
## (Intercept)          1.302712   0.015120   86.16  < 2e-16 ***
## FirstAuthorFemale1   0.022734   0.008025    2.83  0.00462 **
## LastAuthorFemale1   -0.000915   0.009427   -0.10  0.92270
## Year1997            -0.032723   0.020787   -1.57  0.11546
## Year1998            -0.037988   0.020743   -1.83  0.06707 .
## Year1999            -0.069698   0.020821   -3.35  0.00082 ***
## Year2000            -0.070431   0.021192   -3.32  0.00089 ***
## Year2001            -0.073425   0.022714   -3.23  0.00123 **
## Year2002            -0.065357   0.019866   -3.29  0.00100 **
## Year2003            -0.084238   0.019246   -4.38  1.2e-05 ***
## Year2004            -0.096674   0.019165   -5.04  4.6e-07 ***
## Year2005            -0.118870   0.020155   -5.90  3.8e-09 ***
```

```
## Year2006           -0.141691   0.019902   -7.12  1.1e-12 ***
## Year2007           -0.119462   0.018834   -6.34  2.3e-10 ***
## Year2008           -0.144515   0.019479   -7.42  1.2e-13 ***
## Year2009           -0.123683   0.019457   -6.36  2.1e-10 ***
## Year2010           -0.116037   0.020046   -5.79  7.2e-09 ***
## Year2011           -0.141962   0.019607   -7.24  4.7e-13 ***
## Year2012           -0.129627   0.019437   -6.67  2.7e-11 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.395
## Multiple R-squared:  0.0104, Adjusted R-squared:  0.00933
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  2 observations c(6081,9183) are outliers with |weight| = 0 ( < 5.9e-06);
##  1343 weights are ~= 1. The remaining 15580 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0001  0.8710  0.9500  0.8980  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          5.91e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.011  1           1.006
## Year              1.011 16           1.000
```




```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.3026 -0.2659  0.0134  0.2667  2.3602
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)          1.30264    0.01509   86.30  < 2e-16 ***
## FirstAuthorFemale1  0.02260    0.00795    2.84  0.00450 **
## Year1997           -0.03273    0.02079   -1.57  0.11539
## Year1998           -0.03800    0.02074   -1.83  0.06699 .
## Year1999           -0.06971    0.02082   -3.35  0.00082 ***
## Year2000           -0.07045    0.02119   -3.32  0.00089 ***
## Year2001           -0.07344    0.02271   -3.23  0.00122 **
## Year2002           -0.06536    0.01987   -3.29  0.00100 **
## Year2003           -0.08427    0.01925   -4.38  1.2e-05 ***
## Year2004           -0.09670    0.01916   -5.05  4.6e-07 ***
## Year2005           -0.11887    0.02015   -5.90  3.7e-09 ***
## Year2006           -0.14168    0.01990   -7.12  1.1e-12 ***
```

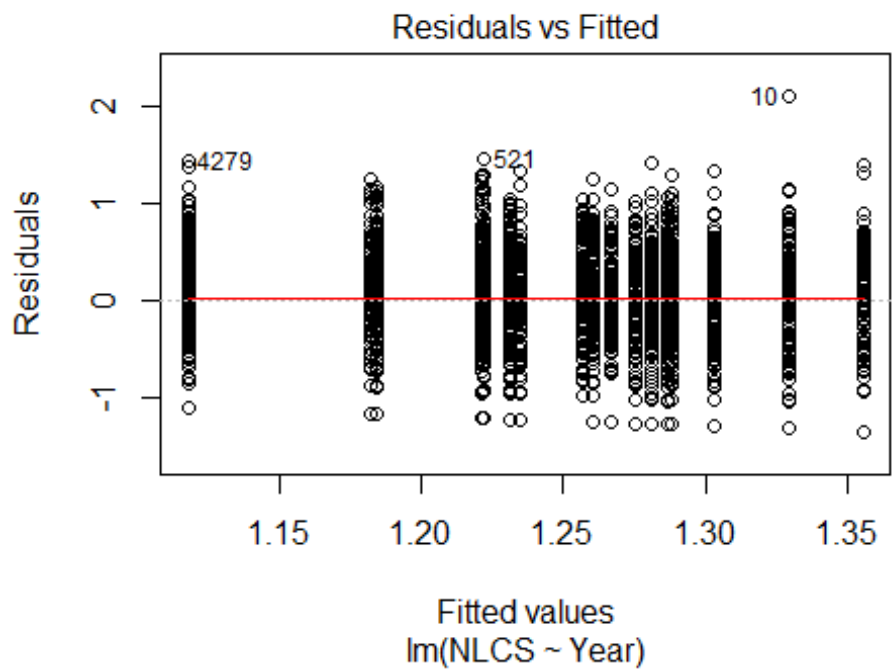
```
## Year2007           -0.11947    0.01883   -6.34  2.3e-10 ***
## Year2008           -0.14455    0.01948   -7.42  1.2e-13 ***
## Year2009           -0.12371    0.01946   -6.36  2.1e-10 ***
## Year2010           -0.11606    0.02005   -5.79  7.2e-09 ***
## Year2011           -0.14200    0.01961   -7.24  4.6e-13 ***
## Year2012           -0.12970    0.01942   -6.68  2.5e-11 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.395
## Multiple R-squared:  0.0104, Adjusted R-squared:  0.00939
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  2 observations c(6081,9183) are outliers with |weight| = 0 ( < 5.9e-06);
##  1337 weights are ~= 1. The remaining 15586 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0001  0.8710  0.9500  0.8980  0.9860  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          5.91e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##                   GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.007  1           1.003
## Year             1.007 16           1.000
```



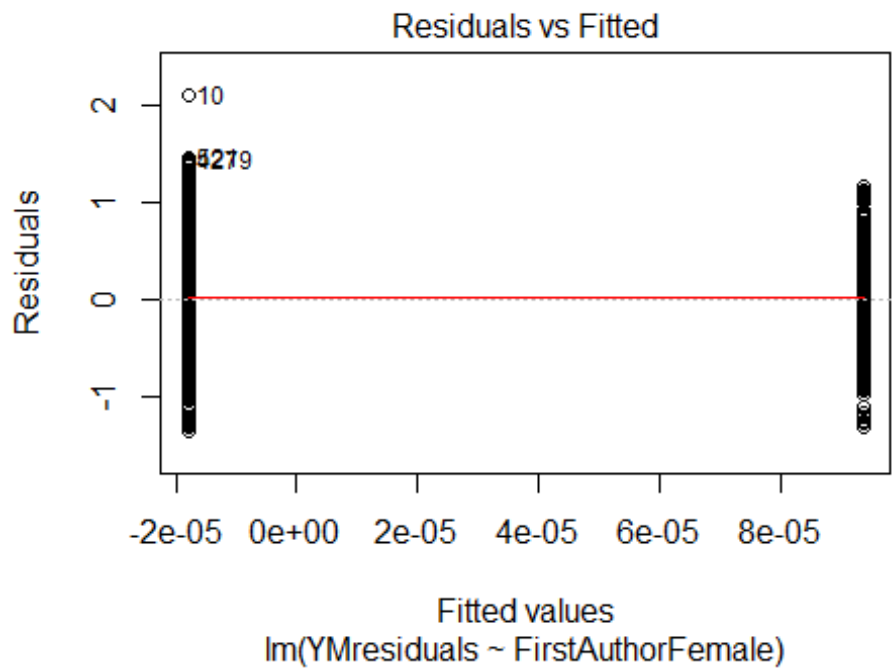

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.3052 -0.2654  0.0142  0.2674  2.3567
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.30517    0.01511   86.38  < 2e-16 ***
## LastAuthorFemale1   0.00335    0.00932    0.36  0.71881
## Year1997           -0.03239    0.02081   -1.56  0.11958
## Year1998           -0.03793    0.02076   -1.83  0.06770 .
## Year1999           -0.06937    0.02081   -3.33  0.00086 ***
## Year2000           -0.06973    0.02119   -3.29  0.00100 **
## Year2001           -0.07248    0.02274   -3.19  0.00143 **
## Year2002           -0.06462    0.01987   -3.25  0.00115 **
## Year2003           -0.08302    0.01925   -4.31  1.6e-05 ***
## Year2004           -0.09549    0.01917   -4.98  6.4e-07 ***
## Year2005           -0.11790    0.02015   -5.85  5.0e-09 ***
## Year2006           -0.14003    0.01991   -7.03  2.1e-12 ***
```

```
## Year2007          -0.11734    0.01882   -6.24  4.6e-10 ***
## Year2008          -0.14274    0.01947   -7.33  2.4e-13 ***
## Year2009          -0.12217    0.01945   -6.28  3.4e-10 ***
## Year2010          -0.11433    0.02005   -5.70  1.2e-08 ***
## Year2011          -0.13956    0.01959   -7.13  1.1e-12 ***
## Year2012          -0.12773    0.01943   -6.57  5.1e-11 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.395
## Multiple R-squared:  0.00992,    Adjusted R-squared:  0.00893
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  2 observations c(6081,9183) are outliers with |weight| = 0 ( < 5.9e-06);
##  1357 weights are ~= 1. The remaining 15566 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0001  0.8710  0.9510  0.8980  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          5.91e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Sample size for the above analysis:  16925"
## [1] ""
## [1] ""
## [1] "####################################"
## [1] "Analysis of AJSC 1903"
## [1] "####################################"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  184  236  192  173  183  211  199  222  178  213  235  191  288  271  214
## 2011 2012
##  277  321
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  121  160  128  112  118  103  137  145  120  144  159  116  172  162  131
```

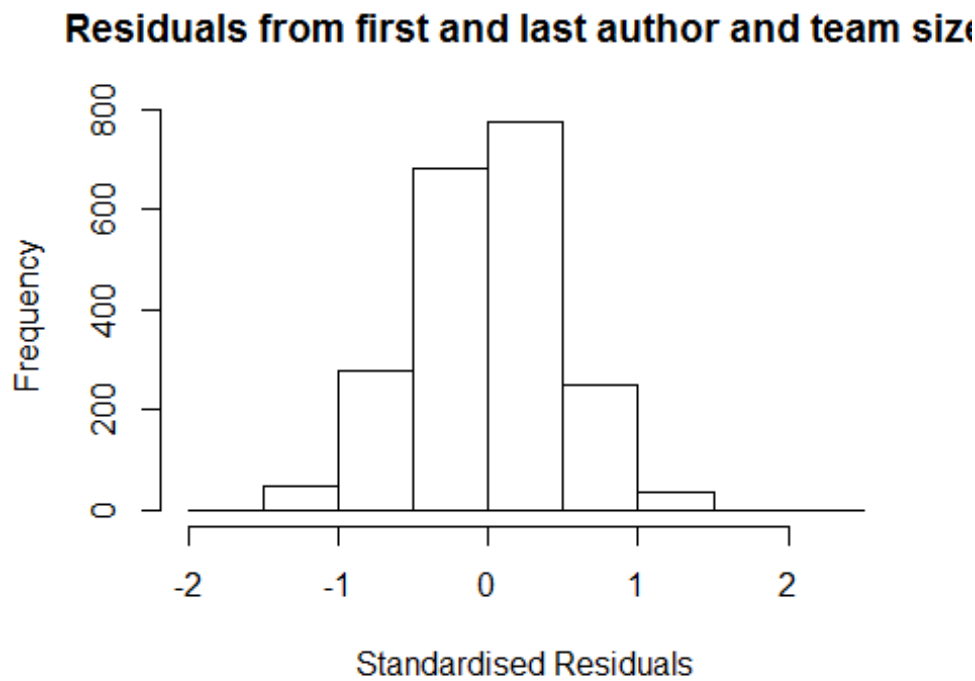
```
## 2011 2012
##  179  199
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  107  140  113  101  108   79  127  129  102  122  127  100  150  143  111
## 2011 2012
##  150  162
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
##  Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 34, df = 16, p-value = 0.006
```




```
##
##  Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 1, df = 1, p-value = 0.3
```

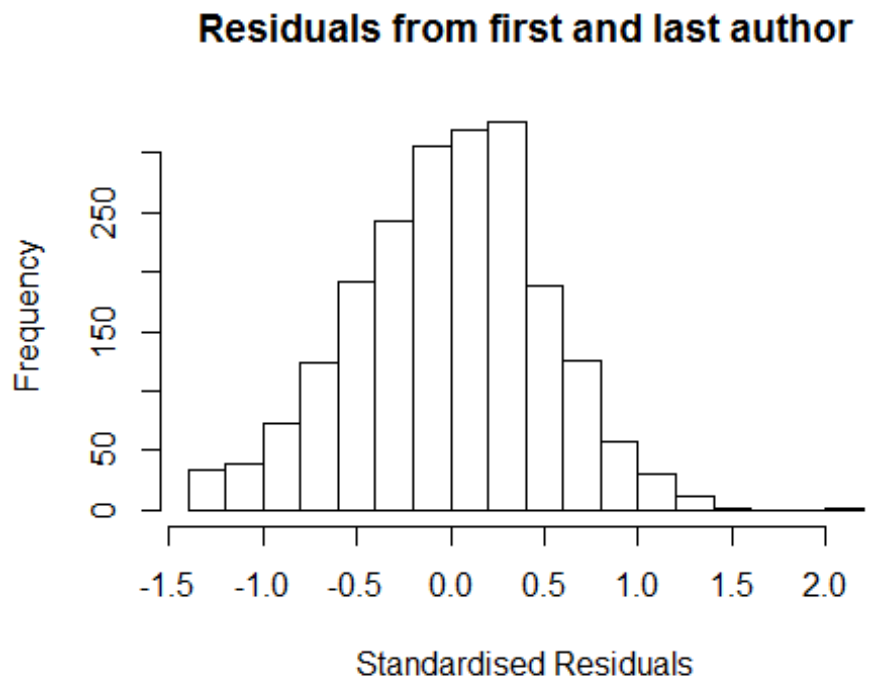



```
## [1] "Female first author team size 2018 geometric mean: 3.49163463928205"
## [1] "Male first author team size 2018 geometric mean: 3.31576971064282"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 2300, p-value = 0.7
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 3.52085085942577"
## [1] "Male last author team size 2018 geometric mean: 3.31167060828195"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 2200, p-value = 0.8
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##                     GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.084  1           1.041
## LastAuthorFemale  1.077  1           1.038
## UniqueAuthors     1.248  4           1.028
## Year              1.277 16           1.008
```



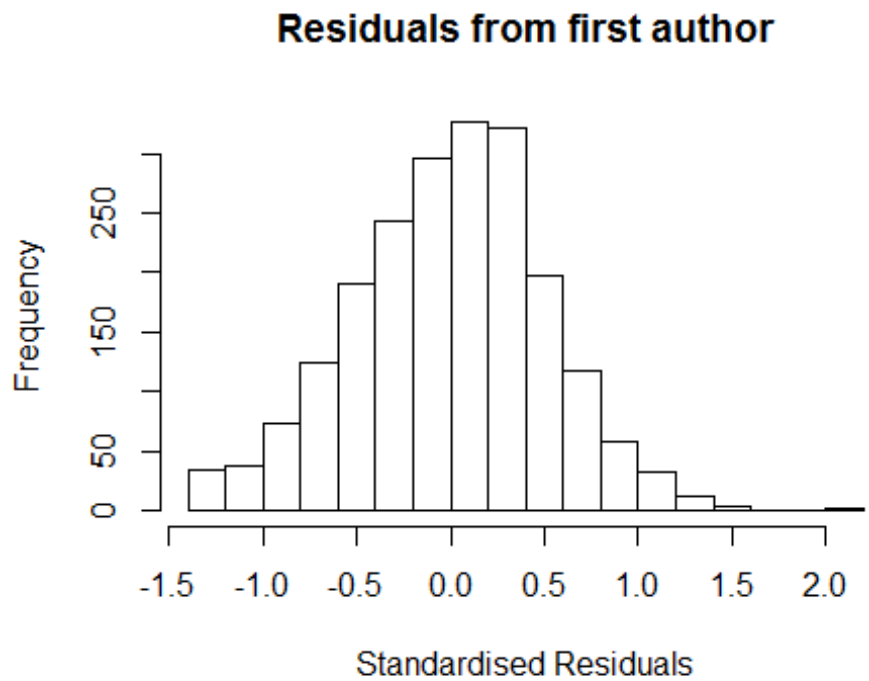

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
UniqueAuthors +
##     Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##     k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.5156 -0.3226  0.0176  0.3239  2.2703
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)          1.1597     0.0657   17.65  < 2e-16 ***
## FirstAuthorFemale1  -0.0375     0.0305   -1.23  0.21802
## LastAuthorFemale1    0.0128     0.0312    0.41  0.68141
## UniqueAuthors2       0.2367     0.0334    7.09  1.8e-12 ***
## UniqueAuthors3       0.2415     0.0346    6.97  4.1e-12 ***
## UniqueAuthors4       0.3766     0.0405    9.30  < 2e-16 ***
## UniqueAuthors5       0.5239     0.0373   14.04  < 2e-16 ***
## Year1997            -0.0771     0.0741   -1.04  0.29846
## Year1998            -0.1509     0.0829   -1.82  0.06888 .
## Year1999            -0.1680     0.0812   -2.07  0.03874 *
```

```
## Year2000            -0.1169     0.0763   -1.53  0.12574
## Year2001            -0.1042     0.0850   -1.23  0.22002
## Year2002            -0.0891     0.0741   -1.20  0.22953
## Year2003            -0.1783     0.0780   -2.28  0.02244 *
## Year2004            -0.1722     0.0761   -2.26  0.02371 *
## Year2005            -0.1760     0.0742   -2.37  0.01782 *
## Year2006            -0.1772     0.0755   -2.35  0.01897 *
## Year2007            -0.1588     0.0773   -2.06  0.03996 *
## Year2008            -0.1243     0.0715   -1.74  0.08226 .
## Year2009            -0.1644     0.0754   -2.18  0.02939 *
## Year2010            -0.2361     0.0800   -2.95  0.00320 **
## Year2011            -0.2554     0.0764   -3.34  0.00084 ***
## Year2012            -0.3394     0.0760   -4.47  8.3e-06 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.477
## Multiple R-squared:  0.112,  Adjusted R-squared:  0.102
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
##  observation 4 is an outlier with |weight| = 0 ( < 4.8e-05);
##  172 weights are ~= 1. The remaining 1898 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.292   0.873   0.950   0.903   0.986   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          4.83e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.044  1           1.022
## LastAuthorFemale  1.025  1           1.012
## Year              1.060 16           1.002
```



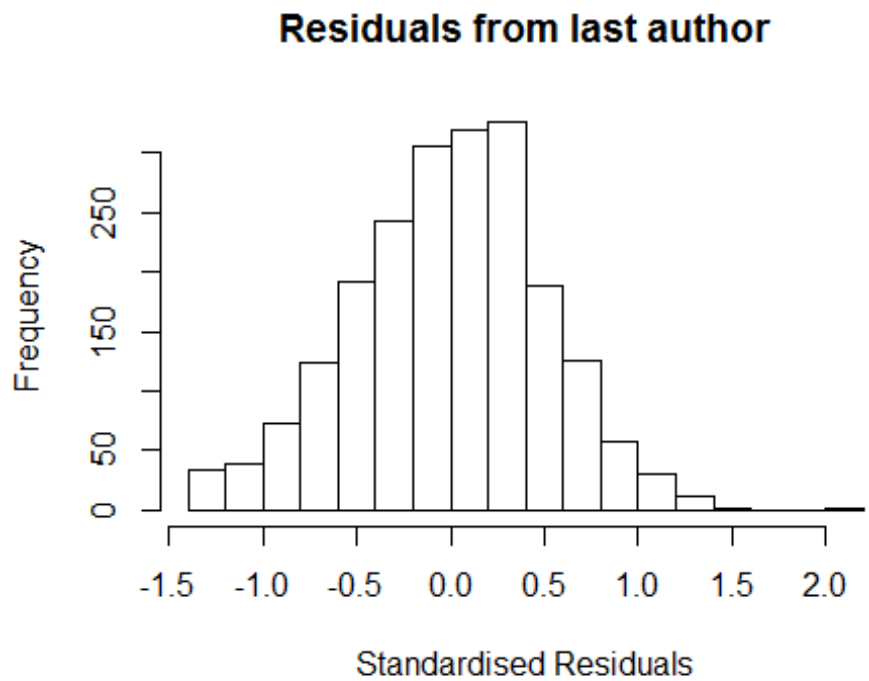

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.3891 -0.3478  0.0185  0.3366  2.1010
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)          1.32902    0.06174   21.53   <2e-16 ***
## FirstAuthorFemale1   0.00241    0.03193    0.08   0.9398
## LastAuthorFemale1    0.05770    0.03250    1.78   0.0760 .
## Year1997            -0.05560    0.07324   -0.76   0.4479
## Year1998            -0.13151    0.08348   -1.58   0.1153
## Year1999            -0.13791    0.08396   -1.64   0.1007
## Year2000            -0.08834    0.07611   -1.16   0.2459
## Year2001            -0.00513    0.09138   -0.06   0.9553
## Year2002            -0.03896    0.07558   -0.52   0.6063
## Year2003            -0.06392    0.07846   -0.81   0.4153
## Year2004            -0.08770    0.07671   -1.14   0.2531
## Year2005            -0.09728    0.07584   -1.28   0.1997
```

```
## Year2006           -0.09205    0.07462   -1.23   0.2175
## Year2007           -0.08282    0.07789   -1.06   0.2878
## Year2008           -0.05316    0.07279   -0.73   0.4652
## Year2009           -0.09705    0.07726   -1.26   0.2092
## Year2010           -0.16931    0.08180   -2.07   0.0386 *
## Year2011           -0.14482    0.07775   -1.86   0.0626 .
## Year2012           -0.23778    0.07744   -3.07   0.0022 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.503
## Multiple R-squared:  0.0144, Adjusted R-squared:  0.00576
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
##  177 weights are ~= 1. The remaining 1894 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0417  0.8710  0.9500  0.9030  0.9840  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          4.83e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.043  1           1.021
## Year              1.043 16           1.001
```




```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.3460 -0.3486  0.0215  0.3343  2.0951
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)          1.33493    0.06182   21.59   <2e-16 ***
## FirstAuthorFemale1  0.01104    0.03184    0.35   0.7289
## Year1997           -0.05677    0.07336   -0.77   0.4391
## Year1998           -0.13164    0.08361   -1.57   0.1155
## Year1999           -0.13598    0.08394   -1.62   0.1054
## Year2000           -0.08815    0.07607   -1.16   0.2467
## Year2001           -0.00347    0.09172   -0.04   0.9699
## Year2002           -0.03848    0.07558   -0.51   0.6107
## Year2003           -0.06286    0.07848   -0.80   0.4232
## Year2004           -0.08363    0.07682   -1.09   0.2764
## Year2005           -0.09475    0.07605   -1.25   0.2129
## Year2006           -0.09083    0.07484   -1.21   0.2250
```

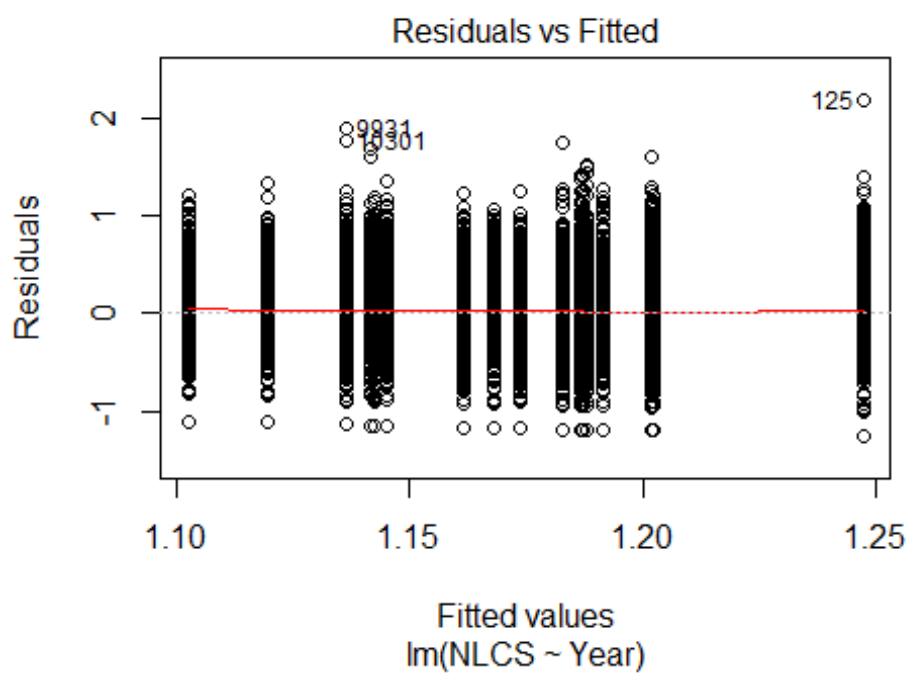
```
## Year2007           -0.07798    0.07793   -1.00   0.3172
## Year2008           -0.05079    0.07286   -0.70   0.4858
## Year2009           -0.09320    0.07726   -1.21   0.2278
## Year2010           -0.16845    0.08171   -2.06   0.0394 *
## Year2011           -0.14245    0.07769   -1.83   0.0669 .
## Year2012           -0.23529    0.07745   -3.04   0.0024 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.504
## Multiple R-squared:  0.0128, Adjusted R-squared:  0.00464
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
##  196 weights are ~= 1. The remaining 1875 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0452  0.8690  0.9500  0.9030  0.9840  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          4.83e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##                   GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.022  1           1.011
## Year             1.022 16           1.001
```



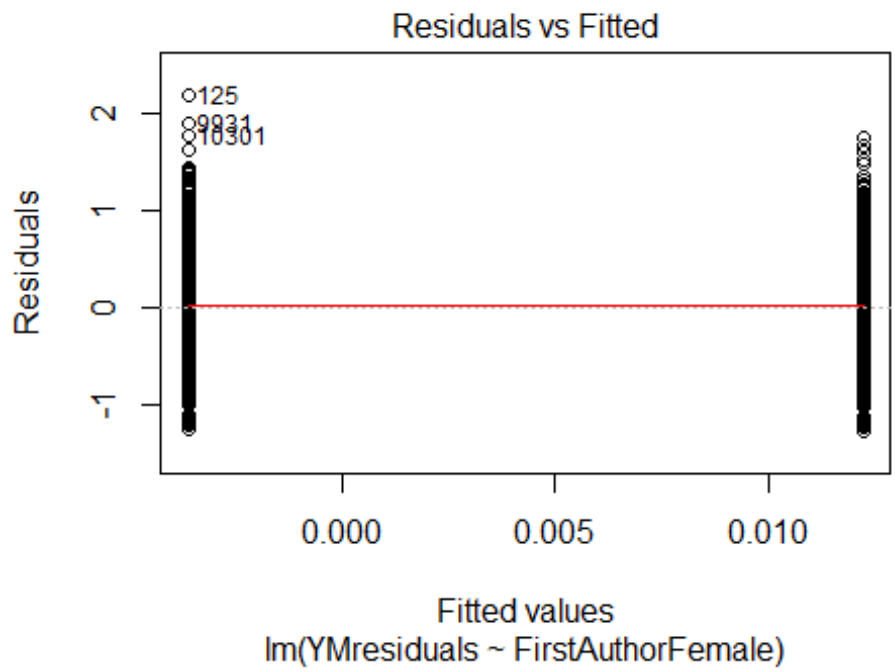

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.3872 -0.3479  0.0199  0.3376  2.1008
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.32915    0.06172   21.54   <2e-16 ***
## LastAuthorFemale1   0.05803    0.03255    1.78   0.0748 .
## Year1997           -0.05558    0.07323   -0.76   0.4479
## Year1998           -0.13135    0.08340   -1.57   0.1154
## Year1999           -0.13779    0.08390   -1.64   0.1007
## Year2000           -0.08823    0.07609   -1.16   0.2464
## Year2001           -0.00493    0.09125   -0.05   0.9570
## Year2002           -0.03875    0.07553   -0.51   0.6080
## Year2003           -0.06383    0.07840   -0.81   0.4157
## Year2004           -0.08755    0.07667   -1.14   0.2537
## Year2005           -0.09706    0.07577   -1.28   0.2004
## Year2006           -0.09199    0.07461   -1.23   0.2178
```

```
## Year2007          -0.08253    0.07770   -1.06   0.2883
## Year2008          -0.05283    0.07259   -0.73   0.4668
## Year2009          -0.09688    0.07718   -1.26   0.2096
## Year2010          -0.16899    0.08172   -2.07   0.0388 *
## Year2011          -0.14450    0.07748   -1.87   0.0623 .
## Year2012          -0.23749    0.07728   -3.07   0.0021 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.503
## Multiple R-squared:  0.0144, Adjusted R-squared:  0.00623
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
##  181 weights are ~= 1. The remaining 1890 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0426  0.8710  0.9500  0.9030  0.9840  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          4.83e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Sample size for the above analysis:  2071"
## [1] ""
## [1] ""
## [1] "####################################"
## [1] "Analysis of AJSC 1904"
## [1] "####################################"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 1129 1433 1331 1344 1253 1569 1330 1464 1318 1305 1424 1497 1538 1403 1396
## 2011 2012
## 1447 1547
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  707  812  759  771  634  789  847  874  828  791  873  926  899  887  819
## 2011 2012
```

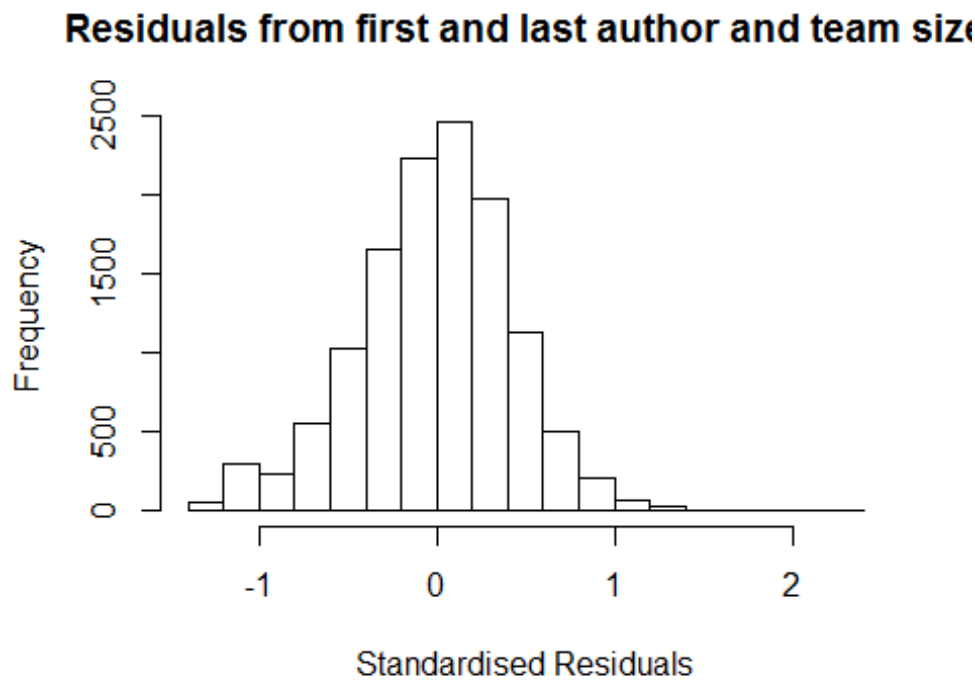
```
##  899  961
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  641  751  669  703  560  695  744  767  725  687  745  799  768  786  719
## 2011 2012
##  784  842
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
##  Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 67, df = 16, p-value = 4e-08
```




```
##
##  Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 0.05, df = 1, p-value = 0.8
```

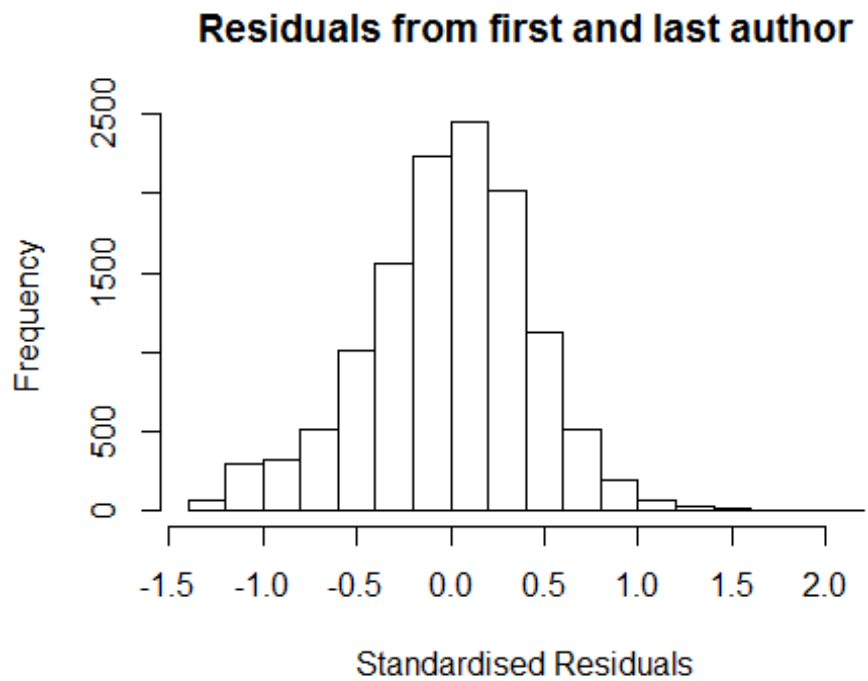



```
## [1] "Female first author team size 2018 geometric mean: 2.94525544439756"
## [1] "Male first author team size 2018 geometric mean: 2.66367514055109"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 84000, p-value = 0.2
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 2.68845623628813"
## [1] "Male last author team size 2018 geometric mean: 2.78363591060344"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 69000, p-value = 0.5
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.040  1           1.020
## LastAuthorFemale  1.036  1           1.018
## UniqueAuthors     1.080  4           1.010
## Year              1.087 16           1.003
```



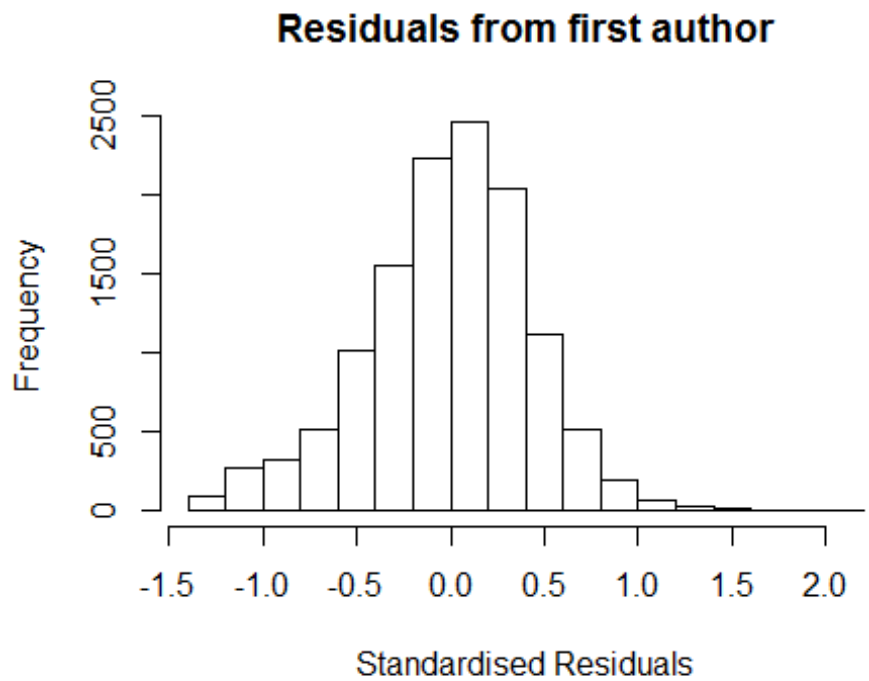

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
UniqueAuthors +
##     Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##     k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.3386 -0.2753  0.0136  0.2704  2.2629
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.16714    0.01974   59.12  < 2e-16 ***
## FirstAuthorFemale1  0.01198    0.00915    1.31  0.19056
## LastAuthorFemale1  -0.01194    0.01027   -1.16  0.24495
## UniqueAuthors2      0.11972    0.01189   10.07  < 2e-16 ***
## UniqueAuthors3      0.14325    0.01230   11.65  < 2e-16 ***
## UniqueAuthors4      0.17929    0.01356   13.22  < 2e-16 ***
## UniqueAuthors5      0.23456    0.01292   18.15  < 2e-16 ***
## Year1997           -0.06574    0.02473   -2.66  0.00786 **
## Year1998           -0.06369    0.02555   -2.49  0.01269 *
## Year1999           -0.07027    0.02440   -2.88  0.00399 **
```

```
## Year2000           -0.09202    0.02529   -3.64  0.00028 ***
## Year2001           -0.06312    0.02504   -2.52  0.01173 *
## Year2002           -0.12438    0.02453   -5.07  4.0e-07 ***
## Year2003           -0.12502    0.02334   -5.36  8.7e-08 ***
## Year2004           -0.14520    0.02374   -6.11  9.9e-10 ***
## Year2005           -0.09627    0.02385   -4.04  5.5e-05 ***
## Year2006           -0.12020    0.02446   -4.91  9.0e-07 ***
## Year2007           -0.08858    0.02298   -3.85  0.00012 ***
## Year2008           -0.13116    0.02341   -5.60  2.2e-08 ***
## Year2009           -0.09599    0.02354   -4.08  4.6e-05 ***
## Year2010           -0.14452    0.02471   -5.85  5.0e-09 ***
## Year2011           -0.17488    0.02485   -7.04  2.1e-12 ***
## Year2012           -0.13559    0.02368   -5.73  1.1e-08 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.402
## Multiple R-squared:  0.0405, Adjusted R-squared:  0.0388
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
##  observation 66 is an outlier with |weight| = 0 ( < 8.1e-06);
##  1087 weights are ~= 1. The remaining 11297 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0001  0.8650  0.9500  0.8940  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          8.07e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.033  1           1.017
## LastAuthorFemale  1.036  1           1.018
## Year              1.018 16           1.001
```



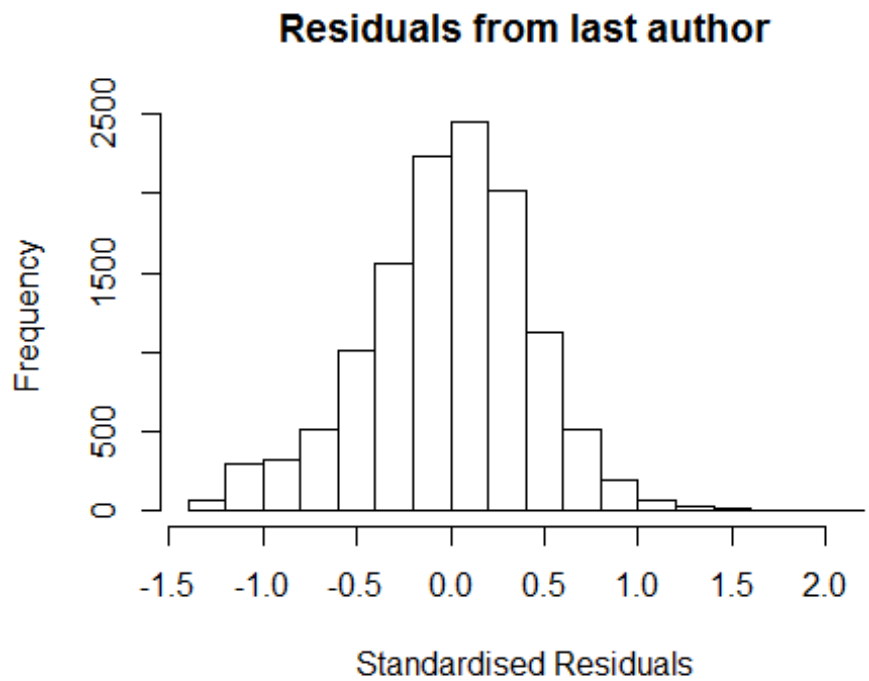

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.2667 -0.2778  0.0168  0.2742  2.1639
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)          1.26615    0.01833   69.06  < 2e-16 ***
## FirstAuthorFemale1  0.01824    0.00923    1.98  0.04818 *
## LastAuthorFemale1  -0.01765    0.01038   -1.70  0.08889 .
## Year1997           -0.07397    0.02493   -2.97  0.00301 **
## Year1998           -0.05492    0.02579   -2.13  0.03323 *
## Year1999           -0.06988    0.02442   -2.86  0.00422 **
## Year2000           -0.08321    0.02538   -3.28  0.00104 **
## Year2001           -0.04558    0.02521   -1.81  0.07068 .
## Year2002           -0.10190    0.02454   -4.15  3.3e-05 ***
## Year2003           -0.10064    0.02338   -4.30  1.7e-05 ***
## Year2004           -0.11637    0.02372   -4.91  9.4e-07 ***
## Year2005           -0.07409    0.02412   -3.07  0.00214 **
```

```
## Year2006           -0.08952    0.02441   -3.67  0.00025 ***
## Year2007           -0.05623    0.02303   -2.44  0.01465 *
## Year2008           -0.09907    0.02359   -4.20  2.7e-05 ***
## Year2009           -0.06781    0.02355   -2.88  0.00399 **
## Year2010           -0.11370    0.02483   -4.58  4.7e-06 ***
## Year2011           -0.14425    0.02502   -5.76  8.4e-09 ***
## Year2012           -0.10498    0.02371   -4.43  9.6e-06 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.406
## Multiple R-squared:  0.0059, Adjusted R-squared:  0.00445
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  observation 66 is an outlier with |weight| = 0 ( < 8.1e-06);
##  1006 weights are ~= 1. The remaining 11378 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0008  0.8660  0.9500  0.8940  0.9860  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          8.07e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.008  1           1.004
## Year              1.008 16           1.000
```




```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.2777 -0.2777  0.0171  0.2751  2.1661
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)           1.2639     0.0183   69.14  < 2e-16 ***
## FirstAuthorFemale1    0.0138     0.0092    1.50  0.13306
## Year1997             -0.0734     0.0249   -2.94  0.00324 **
## Year1998             -0.0549     0.0258   -2.13  0.03350 *
## Year1999             -0.0694     0.0244   -2.84  0.00449 **
## Year2000             -0.0833     0.0254   -3.28  0.00105 **
## Year2001             -0.0456     0.0252   -1.81  0.07085 .
## Year2002             -0.1014     0.0245   -4.13  3.7e-05 ***
## Year2003             -0.1006     0.0234   -4.30  1.7e-05 ***
## Year2004             -0.1163     0.0237   -4.90  9.8e-07 ***
## Year2005             -0.0741     0.0241   -3.07  0.00213 **
## Year2006             -0.0896     0.0244   -3.67  0.00025 ***
```

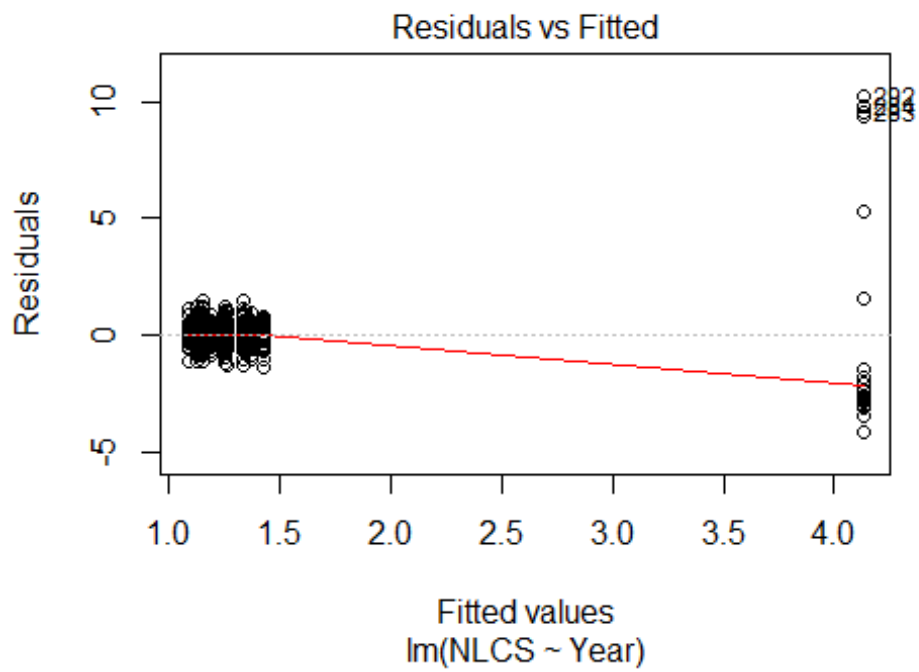
```
## Year2007            -0.0560     0.0230   -2.43  0.01507 *
## Year2008            -0.0996     0.0236   -4.22  2.5e-05 ***
## Year2009            -0.0679     0.0236   -2.88  0.00396 **
## Year2010            -0.1139     0.0248   -4.59  4.6e-06 ***
## Year2011            -0.1442     0.0250   -5.76  8.6e-09 ***
## Year2012            -0.1058     0.0237   -4.47  8.1e-06 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.406
## Multiple R-squared:  0.00567,    Adjusted R-squared:  0.0043
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  observation 66 is an outlier with |weight| = 0 ( < 8.1e-06);
##  998 weights are ~= 1. The remaining 11386 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0008  0.8650  0.9500  0.8940  0.9860  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          8.07e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##                  GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.01  1           1.005
## Year             1.01 16           1.000
```



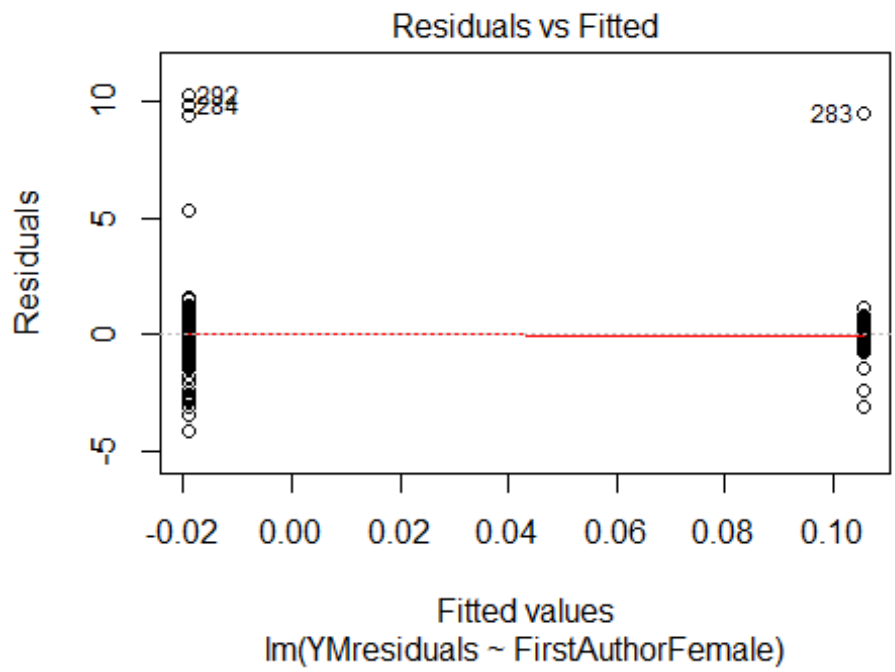

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.2680 -0.2790  0.0169  0.2744  2.1620
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.2680     0.0183   69.33  < 2e-16 ***
## LastAuthorFemale1  -0.0122     0.0103   -1.19   0.2357
## Year1997           -0.0731     0.0249   -2.94   0.0033 **
## Year1998           -0.0544     0.0258   -2.11   0.0350 *
## Year1999           -0.0687     0.0244   -2.82   0.0049 **
## Year2000           -0.0818     0.0253   -3.23   0.0013 **
## Year2001           -0.0445     0.0252   -1.77   0.0772 .
## Year2002           -0.1003     0.0245   -4.09  4.3e-05 ***
## Year2003           -0.0995     0.0234   -4.26  2.1e-05 ***
## Year2004           -0.1148     0.0237   -4.84  1.3e-06 ***
## Year2005           -0.0718     0.0241   -2.98   0.0028 **
## Year2006           -0.0881     0.0244   -3.61   0.0003 ***
```

```
## Year2007           -0.0546     0.0230   -2.37   0.0176 *
## Year2008           -0.0972     0.0236   -4.13  3.7e-05 ***
## Year2009           -0.0658     0.0235   -2.80   0.0052 **
## Year2010           -0.1118     0.0248   -4.51  6.5e-06 ***
## Year2011           -0.1423     0.0250   -5.70  1.3e-08 ***
## Year2012           -0.1030     0.0237   -4.35  1.4e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.406
## Multiple R-squared:  0.00561,    Adjusted R-squared:  0.00424
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  observation 66 is an outlier with |weight| = 0 ( < 8.1e-06);
##  1009 weights are ~= 1. The remaining 11375 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.001   0.865   0.950   0.894   0.986   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          8.07e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                    psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Sample size for the above analysis:  12385"
## [1] ""
## [1] ""
## [1] "####################################"
## [1] "Analysis of AJSC 1905"
## [1] "####################################"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##   71   49   65   41   50   56   40   59   62   68   46   29   56   55   47
## 2011 2012
##   62   66
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##   36   29   40   25   22   23   24   43   50   46   34   18   35   33   36
```

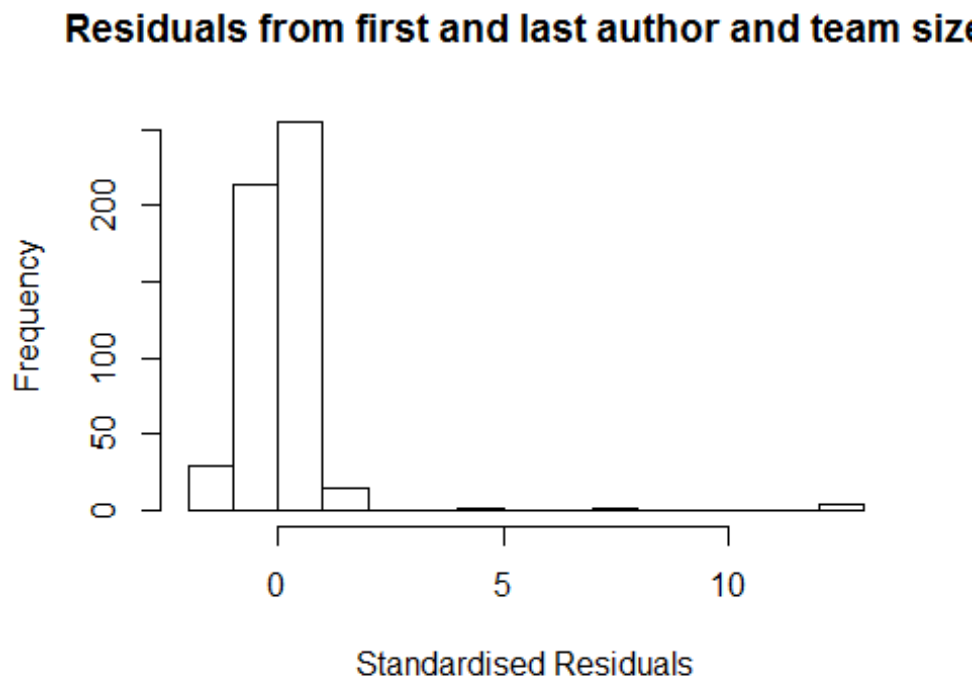
```
## 2011 2012
##   38   47
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##   31   28   35   23   20   20   20   40   49   37   33   13   33   28   32
## 2011 2012
##   33   43
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
##  Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 750, df = 16, p-value <2e-16
```




```
##
##  Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 0.47, df = 1, p-value = 0.5

## [1] "Female first author team size 2018 geometric mean: 3.37805401064092"
## [1] "Male first author team size 2018 geometric mean: 2.26594721802932"

## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties
```



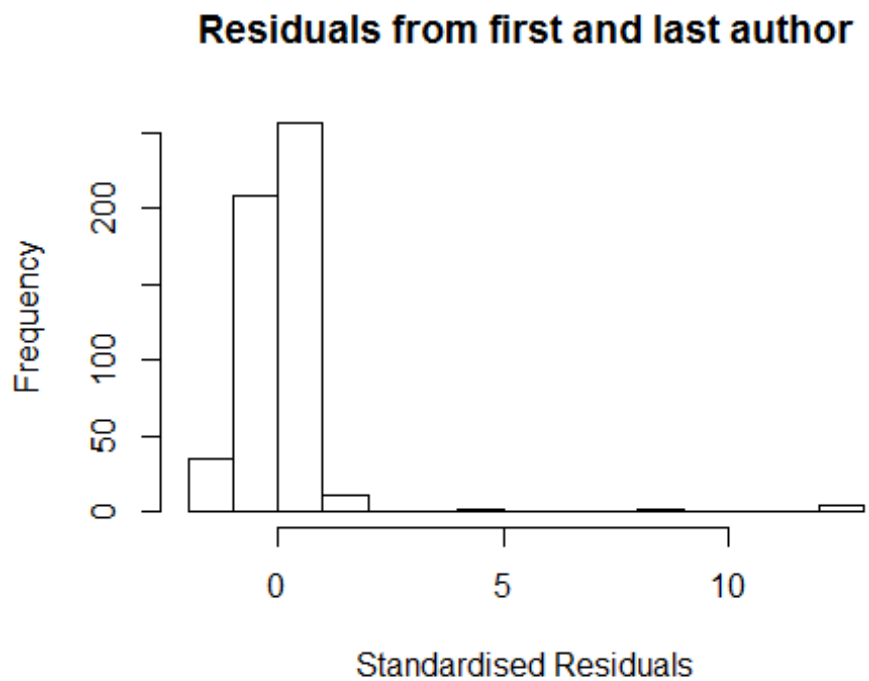

```
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 380, p-value = 0.09
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 2.68120563133006"
## [1] "Male last author team size 2018 geometric mean: 2.41131392761027"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 290, p-value = 0.4
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.491  1           1.221
## LastAuthorFemale  1.613  1           1.270
## UniqueAuthors     2.289  4           1.109
## Year              2.890 16           1.034
```




```
## [1] "List of  6 outliers with residuals above  2.5"
##        ScopusId   NLCS Year OneField Fields residuals
## 283  0042539614 13.671 2000     1905      1    12.144
## 284  0141533490 13.954 2000     1905      1    12.512
## 289 33750801920  9.447 2000     1905      1     7.983
## 291 33750819174  5.688 2000     1905      1     4.224
## 292 33750833421 14.345 2000     1905      1    12.967
## 293 33750834016 13.522 2000     1905      1    12.144
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
UniqueAuthors +
##     Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##     k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.4418 -0.3210  0.0404  0.3409 12.9665
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.09770    0.09416   11.66   <2e-16 ***
## FirstAuthorFemale1  0.06297    0.05899    1.07    0.286
## LastAuthorFemale1  -0.02259    0.07057   -0.32    0.749
## UniqueAuthors2      0.08591    0.07849    1.09    0.274
## UniqueAuthors3      0.07525    0.07884    0.95    0.340
```

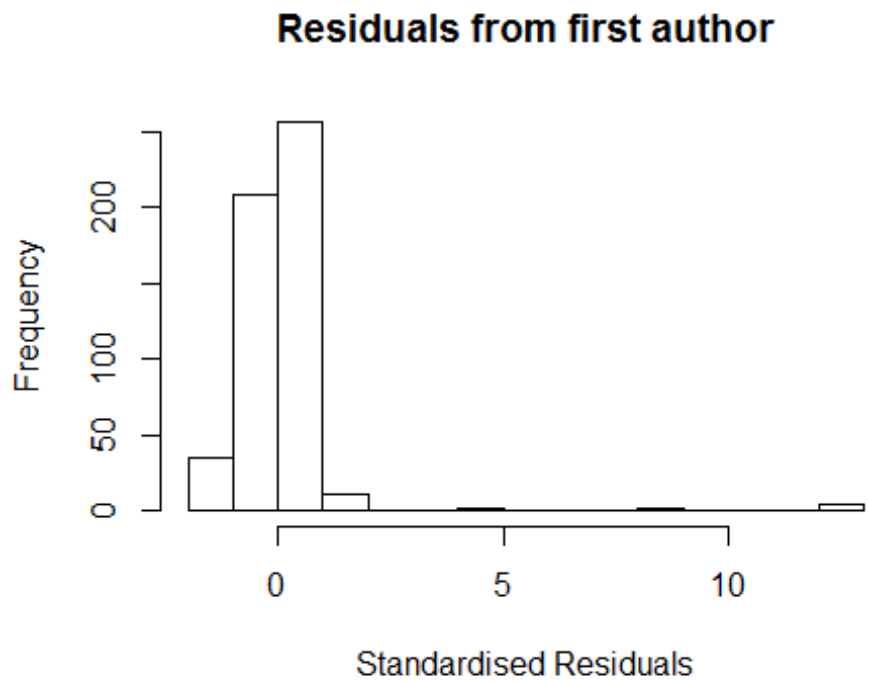
```
## UniqueAuthors4      0.12524    0.08647    1.45    0.148
## UniqueAuthors5      0.20461    0.08772    2.33    0.020 *
## Year1997            0.01607    0.12508    0.13    0.898
## Year1998           -0.01995    0.14397   -0.14    0.890
## Year1999           -0.01338    0.12709   -0.11    0.916
## Year2000            0.28076    0.22485    1.25    0.212
## Year2001            0.06790    0.13134    0.52    0.605
## Year2002            0.07915    0.16508    0.48    0.632
## Year2003            0.09928    0.12242    0.81    0.418
## Year2004            0.00111    0.15735    0.01    0.994
## Year2005           -0.14601    0.13268   -1.10    0.272
## Year2006           -0.09990    0.11402   -0.88    0.381
## Year2007           -0.07358    0.18771   -0.39    0.695
## Year2008            0.25119    0.13054    1.92    0.055 .
## Year2009            0.19833    0.10743    1.85    0.065 .
## Year2010            0.01229    0.11567    0.11    0.915
## Year2011            0.14904    0.10980    1.36    0.175
## Year2012            0.04519    0.11885    0.38    0.704
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.498
## Multiple R-squared:  0.0643, Adjusted R-squared:  0.0227
## Convergence in 18 IRWLS iterations
##
## Robustness weights:
##  6 observations c(118,119,121,122,123,124)
##   are outliers with |weight| = 0 ( < 0.00019);
##  40 weights are ~= 1. The remaining 472 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.350   0.860   0.952   0.896   0.987   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.93e-04          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                    psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##                     GVIF Df GVIF^(1/(2*Df))
```

```
## FirstAuthorFemale 1.462  1           1.209
## LastAuthorFemale  1.485  1           1.219
## Year              1.317 16           1.009
```



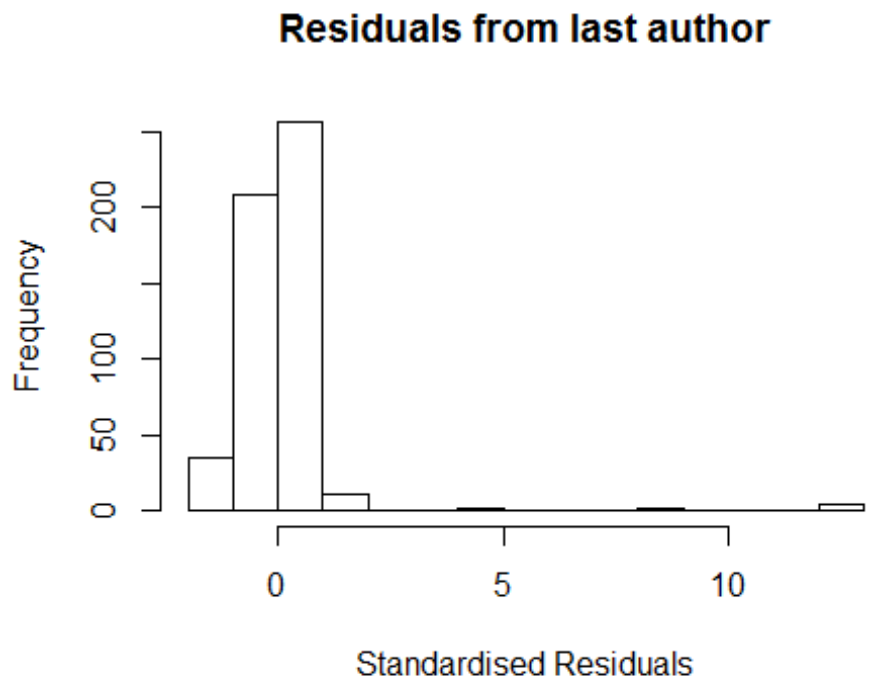

```
## [1] "List of  6 outliers with residuals above  2.5"
##        ScopusId   NLCS Year OneField Fields residuals
## 283  0042539614 13.671 2000     1905      1    12.154
## 284  0141533490 13.954 2000     1905      1    12.538
## 289 33750801920  9.447 2000     1905      1     8.011
## 291 33750819174  5.688 2000     1905      1     4.252
## 292 33750833421 14.345 2000     1905      1    12.909
## 293 33750834016 13.522 2000     1905      1    12.086
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##    Min     1Q Median     3Q    Max
## -1.465 -0.331  0.035  0.337 12.909
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.14403    0.08454   13.53   <2e-16 ***
## FirstAuthorFemale1  0.08092    0.06039    1.34    0.181
## LastAuthorFemale1  -0.02036    0.07033   -0.29    0.772
```

```
## Year1997            0.04024    0.12479    0.32    0.747
## Year1998            0.00923    0.14123    0.07    0.948
## Year1999            0.01951    0.12843    0.15    0.879
## Year2000            0.29216    0.22447    1.30    0.194
## Year2001            0.06107    0.13369    0.46    0.648
## Year2002            0.14238    0.15477    0.92    0.358
## Year2003            0.11439    0.12296    0.93    0.353
## Year2004            0.01846    0.15548    0.12    0.906
## Year2005           -0.10895    0.13221   -0.82    0.410
## Year2006           -0.07717    0.11390   -0.68    0.498
## Year2007           -0.07008    0.19301   -0.36    0.717
## Year2008            0.32075    0.12698    2.53    0.012 *
## Year2009            0.23900    0.10820    2.21    0.028 *
## Year2010            0.03711    0.11569    0.32    0.749
## Year2011            0.19687    0.10976    1.79    0.073 .
## Year2012            0.11785    0.11431    1.03    0.303
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.503
## Multiple R-squared:  0.0508, Adjusted R-squared:  0.0166
## Convergence in 16 IRWLS iterations
##
## Robustness weights:
##  6 observations c(118,119,121,122,123,124)
##   are outliers with |weight| = 0 ( < 0.00019);
##  44 weights are ~= 1. The remaining 468 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.352   0.859   0.954   0.896   0.986   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.93e-04          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.134  1           1.065
## Year              1.134 16           1.004
```




```
## [1] "List of  6 outliers with residuals above  2.5"
##        ScopusId   NLCS Year OneField Fields residuals
## 283  0042539614 13.671 2000     1905      1    12.154
## 284  0141533490 13.954 2000     1905      1    12.538
## 289 33750801920  9.447 2000     1905      1     8.011
## 291 33750819174  5.688 2000     1905      1     4.252
## 292 33750833421 14.345 2000     1905      1    12.909
## 293 33750834016 13.522 2000     1905      1    12.086
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.4635 -0.3279  0.0341  0.3388 12.9062
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)          1.14309    0.08454   13.52   <2e-16 ***
## FirstAuthorFemale1  0.07448    0.05371    1.39    0.166
## Year1997            0.04155    0.12474    0.33    0.739
## Year1998            0.00902    0.14148    0.06    0.949
## Year1999            0.01745    0.12892    0.14    0.892
## Year2000            0.29572    0.22307    1.33    0.186
## Year2001            0.06055    0.13403    0.45    0.652
```

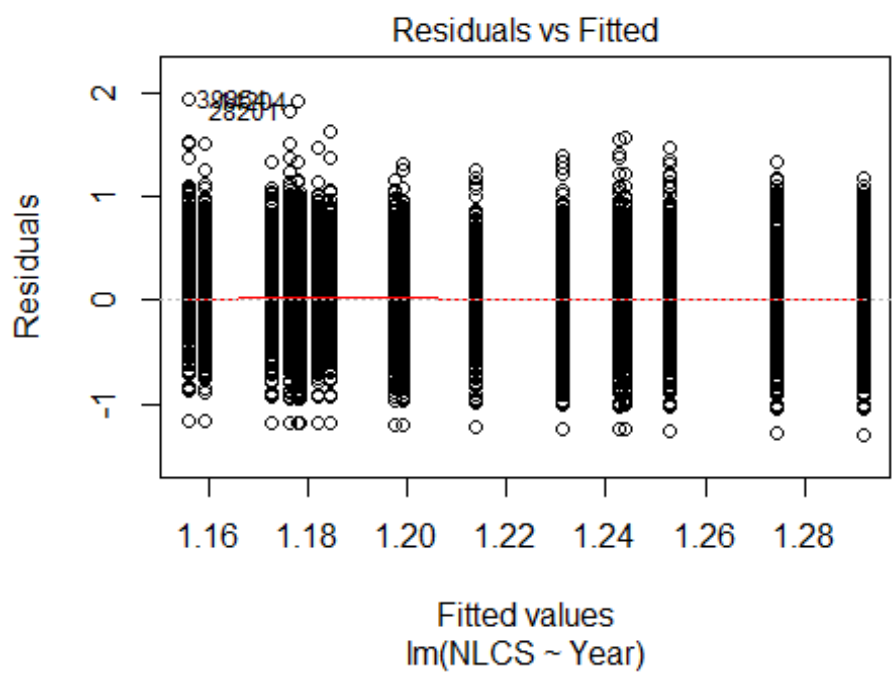
```
## Year2002            0.14410    0.15471    0.93    0.352
## Year2003            0.11309    0.12326    0.92    0.359
## Year2004            0.01946    0.15554    0.13    0.900
## Year2005           -0.10936    0.13231   -0.83    0.409
## Year2006           -0.07794    0.11370   -0.69    0.493
## Year2007           -0.07305    0.19164   -0.38    0.703
## Year2008            0.32036    0.12737    2.52    0.012 *
## Year2009            0.23861    0.10833    2.20    0.028 *
## Year2010            0.03674    0.11599    0.32    0.752
## Year2011            0.19612    0.10997    1.78    0.075 .
## Year2012            0.11432    0.11300    1.01    0.312
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.502
## Multiple R-squared:  0.0509, Adjusted R-squared:  0.0186
## Convergence in 16 IRWLS iterations
##
## Robustness weights:
##  6 observations c(118,119,121,122,123,124)
##   are outliers with |weight| = 0 ( < 0.00019);
##  41 weights are ~= 1. The remaining 471 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.363   0.861   0.954   0.896   0.987   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.93e-04          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##                   GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.149  1           1.072
## Year             1.149 16           1.004
```



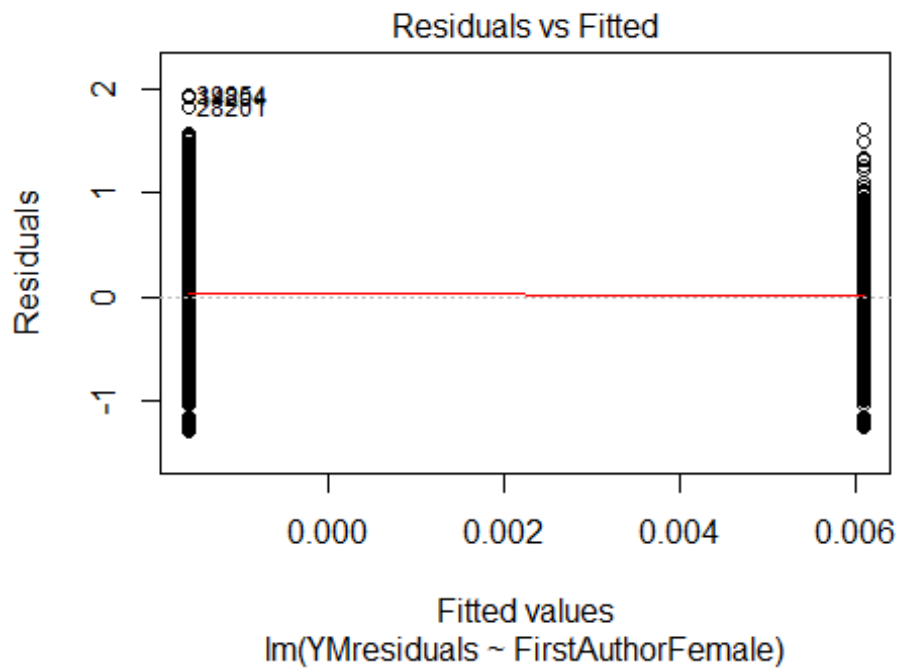

```
## [1] "List of  6 outliers with residuals above  2.5"
##        ScopusId   NLCS Year OneField Fields residuals
## 283  0042539614 13.671 2000     1905      1    12.154
## 284  0141533490 13.954 2000     1905      1    12.538
## 289 33750801920  9.447 2000     1905      1     8.011
## 291 33750819174  5.688 2000     1905      1     4.252
## 292 33750833421 14.345 2000     1905      1    12.909
## 293 33750834016 13.522 2000     1905      1    12.086
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.4783 -0.3237  0.0298  0.3366 12.8967
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)          1.1482     0.0843   13.61   <2e-16 ***
## LastAuthorFemale1    0.0119     0.0629    0.19   0.8496
## Year1997             0.0482     0.1229    0.39   0.6950
## Year1998             0.0153     0.1421    0.11   0.9146
## Year1999             0.0258     0.1272    0.20   0.8395
## Year2000             0.3001     0.2268    1.32   0.1864
## Year2001             0.0605     0.1340    0.45   0.6521
```

```
## Year2002            0.1599     0.1510    1.06   0.2901
## Year2003            0.1151     0.1224    0.94   0.3474
## Year2004            0.0360     0.1560    0.23   0.8175
## Year2005           -0.1043     0.1319   -0.79   0.4293
## Year2006           -0.0713     0.1136   -0.63   0.5303
## Year2007           -0.0545     0.1968   -0.28   0.7819
## Year2008            0.3302     0.1276    2.59   0.0099 **
## Year2009            0.2449     0.1080    2.27   0.0238 *
## Year2010            0.0446     0.1156    0.39   0.6997
## Year2011            0.2048     0.1099    1.86   0.0630 .
## Year2012            0.1178     0.1144    1.03   0.3039
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.502
## Multiple R-squared:  0.0483, Adjusted R-squared:  0.0159
## Convergence in 15 IRWLS iterations
##
## Robustness weights:
##  6 observations c(118,119,121,122,123,124)
##   are outliers with |weight| = 0 ( < 0.00019);
##  41 weights are ~= 1. The remaining 471 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.366   0.866   0.953   0.895   0.987   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.93e-04          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Sample size for the above analysis:  518"
## [1] ""
## [1] ""
## [1] "####################################"
## [1] "Analysis of AJSC 1906"
## [1] "####################################"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
```

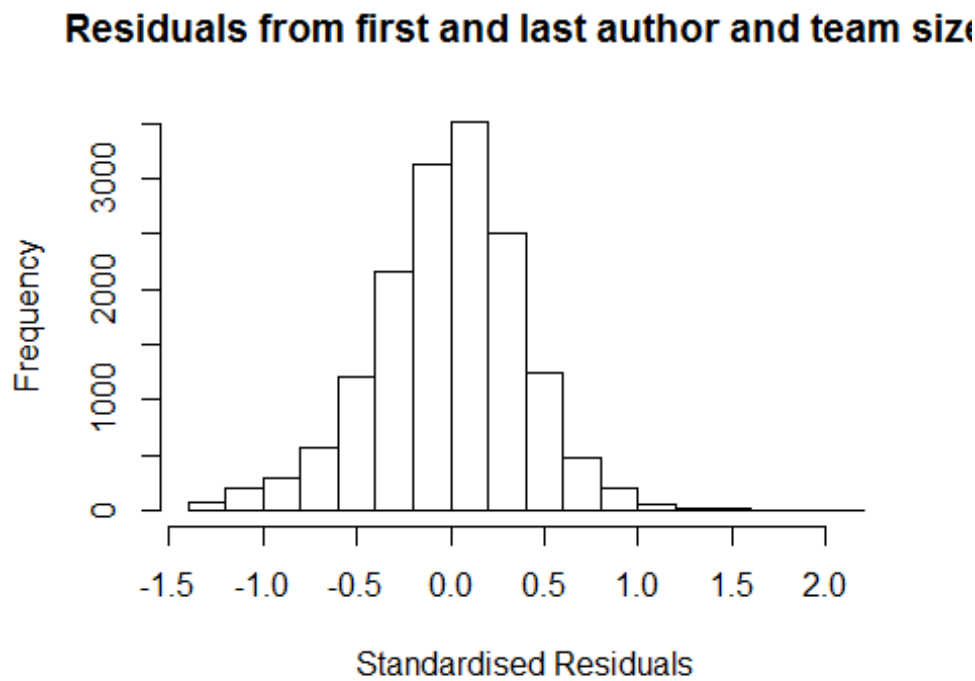
```
## 2005 1961 1983 1938 1741 1971 1895 2242 1852 1693 1997 1902 1927 1770 1757
## 2011 2012
## 1728 1780
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 1135 1096 1109 1047  851  876 1137 1301 1110 1002 1225 1160 1124 1063  991
## 2011 2012
##  988 1035
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  989  983  942  921  725  750  975 1120  951  830 1033  997  942  931  855
## 2011 2012
##  839  869
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
##  Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 39, df = 16, p-value = 0.001
```




```
##
##  Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 66, df = 1, p-value = 5e-16
```

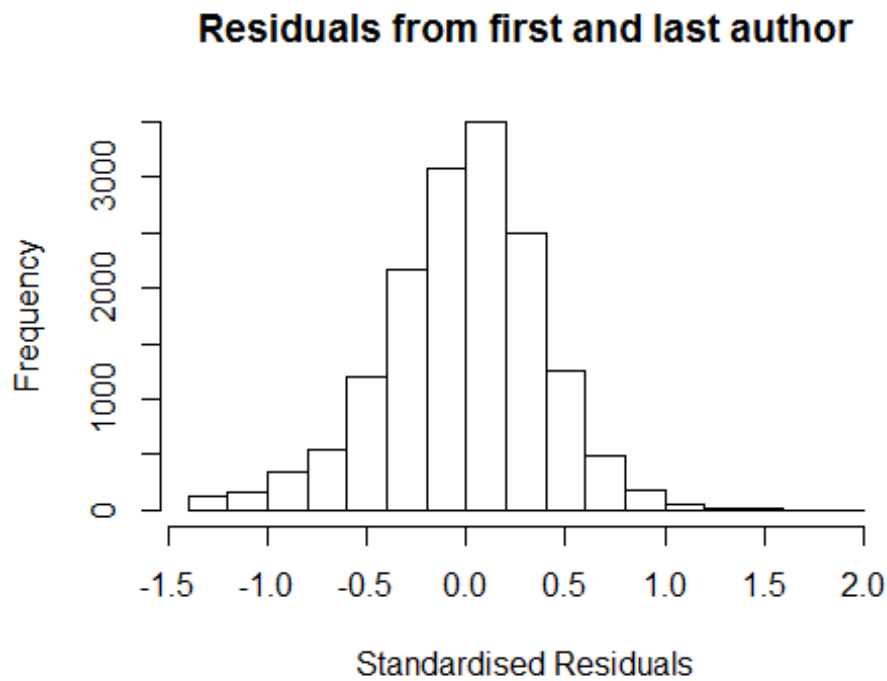



```
## [1] "Female first author team size 2018 geometric mean: 3.53893818130429"
## [1] "Male first author team size 2018 geometric mean: 3.05700659873634"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 37000, p-value = 0.04
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 3.46798500664273"
## [1] "Male last author team size 2018 geometric mean: 3.13177522956585"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 26000, p-value = 0.1
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.044  1           1.022
## LastAuthorFemale  1.030  1           1.015
## UniqueAuthors     1.076  4           1.009
## Year              1.079 16           1.002
```



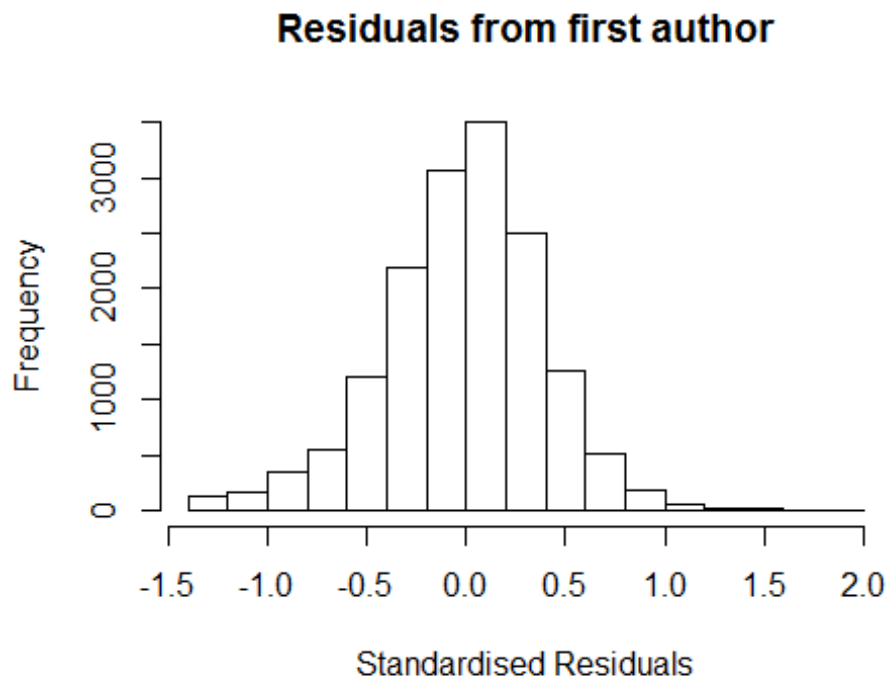

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
UniqueAuthors +
##     Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##     k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.3811 -0.2460  0.0104  0.2401  2.0043
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.22113    0.01550   78.77  < 2e-16 ***
## FirstAuthorFemale1 -0.00658    0.00721   -0.91  0.36137
## LastAuthorFemale1  -0.01578    0.00866   -1.82  0.06832 .
## UniqueAuthors2      0.09476    0.01099    8.62  < 2e-16 ***
## UniqueAuthors3      0.10938    0.01129    9.68  < 2e-16 ***
## UniqueAuthors4      0.13445    0.01230   10.93  < 2e-16 ***
## UniqueAuthors5      0.17390    0.01195   14.56  < 2e-16 ***
## Year1997           -0.01388    0.01838   -0.76  0.45006
## Year1998           -0.04009    0.01882   -2.13  0.03323 *
## Year1999           -0.05296    0.01861   -2.85  0.00443 **
```

```
## Year2000           -0.06887    0.01867   -3.69  0.00023 ***
## Year2001           -0.06398    0.01952   -3.28  0.00105 **
## Year2002           -0.11845    0.01880   -6.30  3.0e-10 ***
## Year2003           -0.11079    0.01722   -6.43  1.3e-10 ***
## Year2004           -0.09579    0.01782   -5.37  7.8e-08 ***
## Year2005           -0.09571    0.01906   -5.02  5.2e-07 ***
## Year2006           -0.12954    0.01775   -7.30  3.0e-13 ***
## Year2007           -0.13844    0.01804   -7.68  1.7e-14 ***
## Year2008           -0.12874    0.01796   -7.17  7.9e-13 ***
## Year2009           -0.12520    0.01794   -6.98  3.1e-12 ***
## Year2010           -0.12665    0.01868   -6.78  1.3e-11 ***
## Year2011           -0.16321    0.01856   -8.79  < 2e-16 ***
## Year2012           -0.15971    0.01866   -8.56  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.361
## Multiple R-squared:  0.0295, Adjusted R-squared:  0.0281
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
##  3 observations c(5452,10917,15176)
##   are outliers with |weight| = 0 ( < 6.4e-06);
##  1359 weights are ~= 1. The remaining 14290 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0016  0.8640  0.9500  0.8930  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          6.39e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.034  1           1.017
## LastAuthorFemale  1.024  1           1.012
## Year              1.020 16           1.001
```



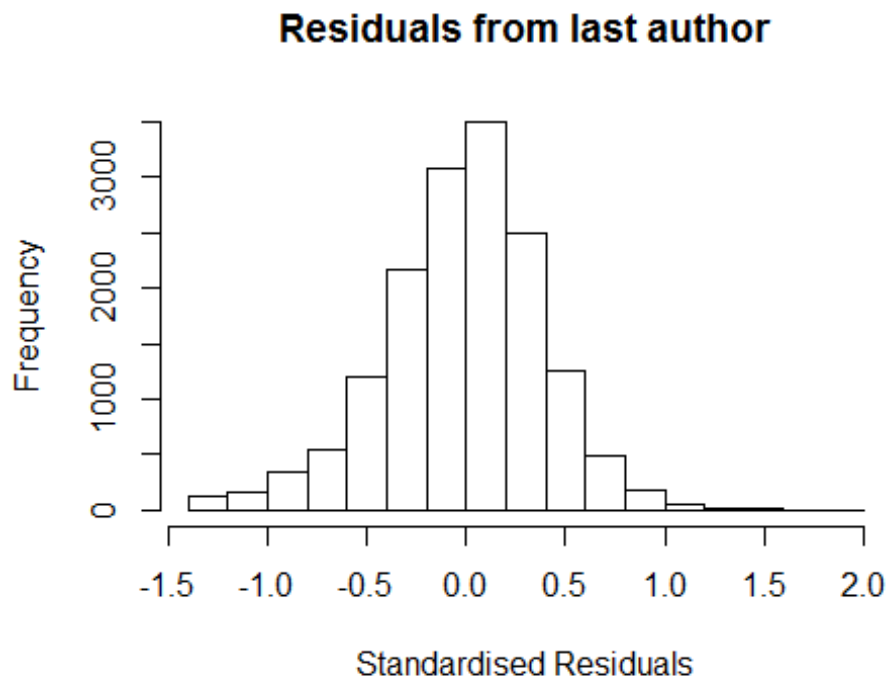

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.3048 -0.2492  0.0119  0.2434  1.9334
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)          1.30476    0.01324   98.52  < 2e-16 ***
## FirstAuthorFemale1  0.00276    0.00722    0.38   0.7022
## LastAuthorFemale1  -0.01005    0.00869   -1.16   0.2479
## Year1997           -0.01586    0.01839   -0.86   0.3884
## Year1998           -0.03758    0.01886   -1.99   0.0464 *
## Year1999           -0.04527    0.01858   -2.44   0.0148 *
## Year2000           -0.05815    0.01873   -3.10   0.0019 **
## Year2001           -0.04722    0.01959   -2.41   0.0159 *
## Year2002           -0.10255    0.01872   -5.48  4.4e-08 ***
## Year2003           -0.09228    0.01716   -5.38  7.6e-08 ***
## Year2004           -0.07350    0.01772   -4.15  3.4e-05 ***
## Year2005           -0.07649    0.01906   -4.01  6.0e-05 ***
```

```
## Year2006           -0.10852    0.01767   -6.14  8.3e-10 ***
## Year2007           -0.11957    0.01808   -6.61  3.9e-11 ***
## Year2008           -0.10630    0.01782   -5.97  2.5e-09 ***
## Year2009           -0.10257    0.01787   -5.74  9.7e-09 ***
## Year2010           -0.10347    0.01868   -5.54  3.1e-08 ***
## Year2011           -0.13722    0.01853   -7.40  1.4e-13 ***
## Year2012           -0.13415    0.01861   -7.21  5.8e-13 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.363
## Multiple R-squared:  0.011,  Adjusted R-squared:  0.00986
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  3 observations c(5452,10917,15176)
##   are outliers with |weight| = 0 ( < 6.4e-06);
##  1320 weights are ~= 1. The remaining 14329 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0118  0.8640  0.9500  0.8920  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          6.39e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.015  1           1.008
## Year              1.015 16           1.000
```




```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.3037 -0.2493  0.0121  0.2437  1.9251
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)          1.30374    0.01322   98.60  < 2e-16 ***
## FirstAuthorFemale1  0.00143    0.00716    0.20   0.8419
## Year1997           -0.01574    0.01839   -0.86   0.3921
## Year1998           -0.03752    0.01886   -1.99   0.0467 *
## Year1999           -0.04538    0.01857   -2.44   0.0146 *
## Year2000           -0.05818    0.01874   -3.11   0.0019 **
## Year2001           -0.04731    0.01959   -2.42   0.0157 *
## Year2002           -0.10273    0.01872   -5.49  4.1e-08 ***
## Year2003           -0.09247    0.01716   -5.39  7.2e-08 ***
## Year2004           -0.07366    0.01772   -4.16  3.2e-05 ***
## Year2005           -0.07656    0.01906   -4.02  5.9e-05 ***
## Year2006           -0.10850    0.01767   -6.14  8.4e-10 ***
```

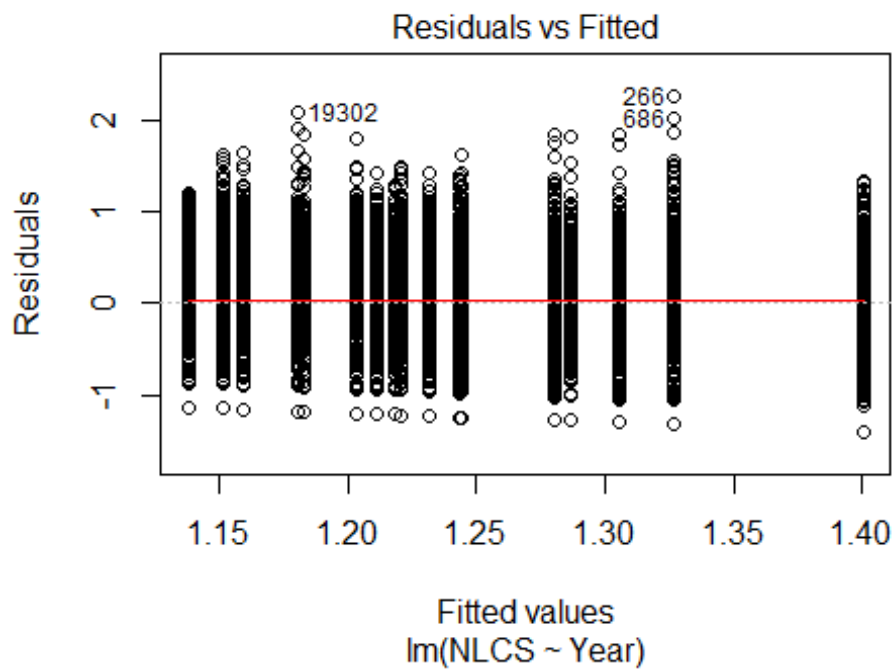
```
## Year2007           -0.11962    0.01808   -6.61  3.8e-11 ***
## Year2008           -0.10638    0.01782   -5.97  2.4e-09 ***
## Year2009           -0.10284    0.01787   -5.75  8.9e-09 ***
## Year2010           -0.10379    0.01867   -5.56  2.8e-08 ***
## Year2011           -0.13766    0.01852   -7.43  1.1e-13 ***
## Year2012           -0.13479    0.01859   -7.25  4.4e-13 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.363
## Multiple R-squared:  0.0109, Adjusted R-squared:  0.00984
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  3 observations c(5452,10917,15176)
##   are outliers with |weight| = 0 ( < 6.4e-06);
##  1313 weights are ~= 1. The remaining 14336 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0112  0.8640  0.9500  0.8920  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          6.39e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##                   GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.006  1           1.003
## Year             1.006 16           1.000
```



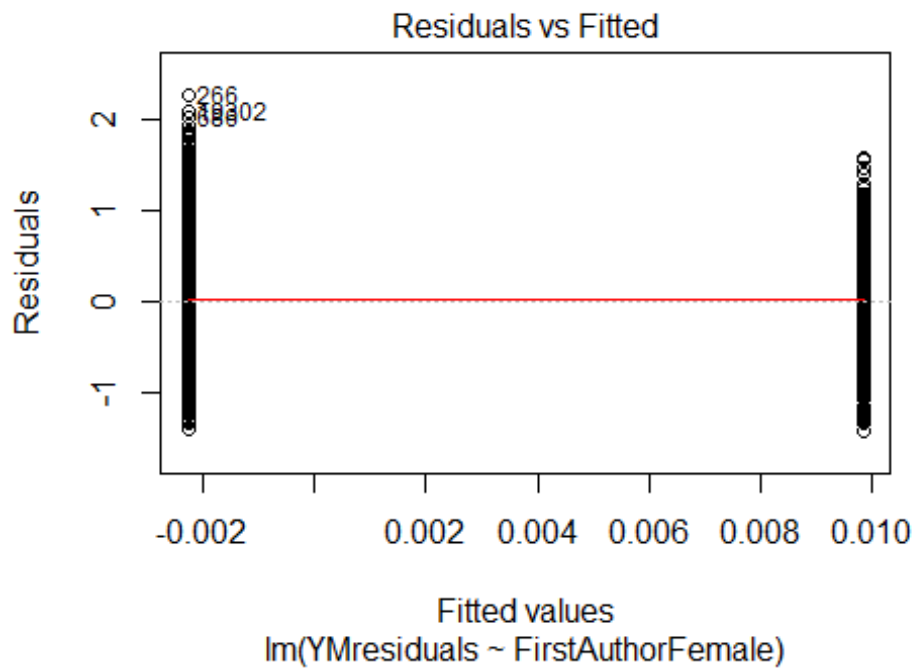

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.3051 -0.2493  0.0119  0.2432  1.9322
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.30515    0.01320   98.89  < 2e-16 ***
## LastAuthorFemale1 -0.00954    0.00862   -1.11   0.2684
## Year1997          -0.01579    0.01839   -0.86   0.3905
## Year1998          -0.03753    0.01886   -1.99   0.0466 *
## Year1999          -0.04522    0.01857   -2.43   0.0149 *
## Year2000          -0.05804    0.01873   -3.10   0.0019 **
## Year2001          -0.04719    0.01958   -2.41   0.0160 *
## Year2002          -0.10242    0.01872   -5.47  4.5e-08 ***
## Year2003          -0.09220    0.01715   -5.37  7.8e-08 ***
## Year2004          -0.07330    0.01770   -4.14  3.5e-05 ***
## Year2005          -0.07632    0.01904   -4.01  6.1e-05 ***
## Year2006          -0.10835    0.01766   -6.14  8.7e-10 ***
```

```
## Year2007          -0.11935    0.01806   -6.61  4.0e-11 ***
## Year2008          -0.10606    0.01780   -5.96  2.6e-09 ***
## Year2009          -0.10235    0.01787   -5.73  1.0e-08 ***
## Year2010          -0.10320    0.01865   -5.53  3.2e-08 ***
## Year2011          -0.13687    0.01850   -7.40  1.4e-13 ***
## Year2012          -0.13382    0.01858   -7.20  6.2e-13 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.363
## Multiple R-squared:  0.011,  Adjusted R-squared:  0.00992
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  3 observations c(5452,10917,15176)
##   are outliers with |weight| = 0 ( < 6.4e-06);
##  1317 weights are ~= 1. The remaining 14332 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0113  0.8640  0.9500  0.8920  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          6.39e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                  psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Sample size for the above analysis:  15652"
## [1] ""
## [1] ""
## [1] "##################################"
## [1] "Analysis of AJSC 1907"
## [1] "##################################"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  972  916  960  919  930  923 1041  865  853  852  927  912  886  914  897
## 2011 2012
##  919  840
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
```

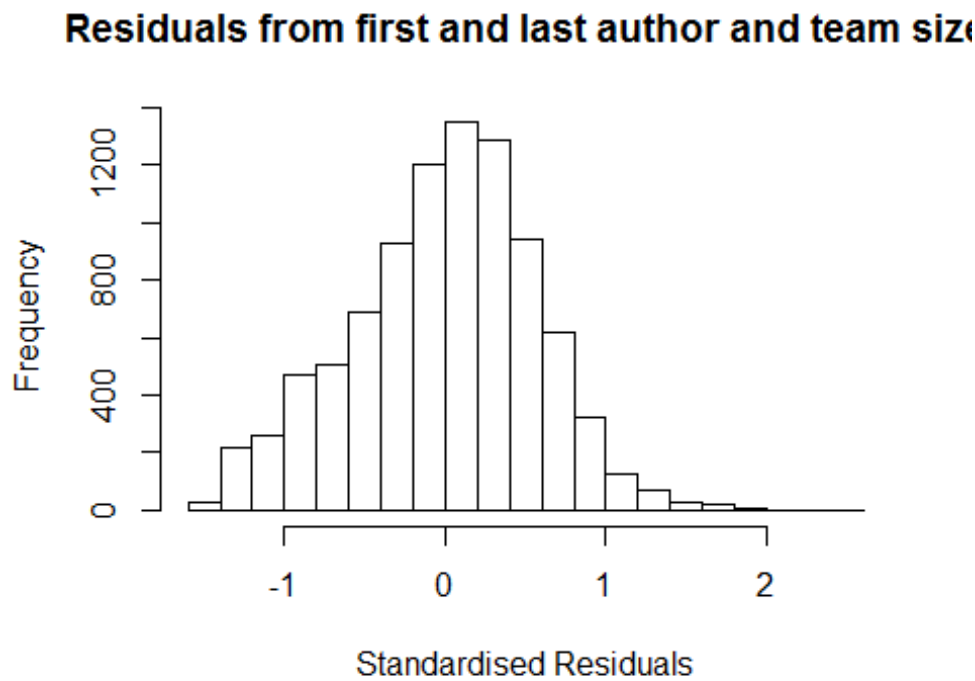
```
##  622  631  675  635  352  418  700  622  622  615  660  628  598  602  621
## 2011 2012
##  647  598
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  547  576  595  576  310  357  628  560  561  553  575  565  521  537  544
## 2011 2012
##  566  515
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
##  Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 180, df = 16, p-value <2e-16
```




```
##
##  Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 33, df = 1, p-value = 9e-09
```

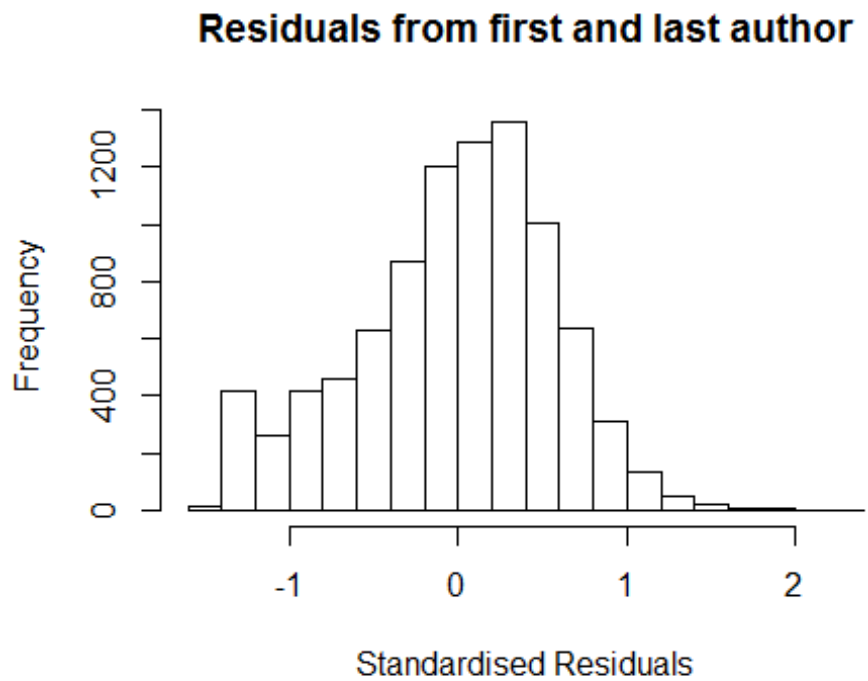



```
## [1] "Female first author team size 2018 geometric mean: 3.3460839694304"
## [1] "Male first author team size 2018 geometric mean: 2.90410840536789"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 41000, p-value = 0.01
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 3.00116621900742"
## [1] "Male last author team size 2018 geometric mean: 3.02172454593913"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 28000, p-value = 0.6
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.035  1           1.017
## LastAuthorFemale  1.027  1           1.014
## UniqueAuthors     1.078  4           1.009
## Year              1.095 16           1.003
```



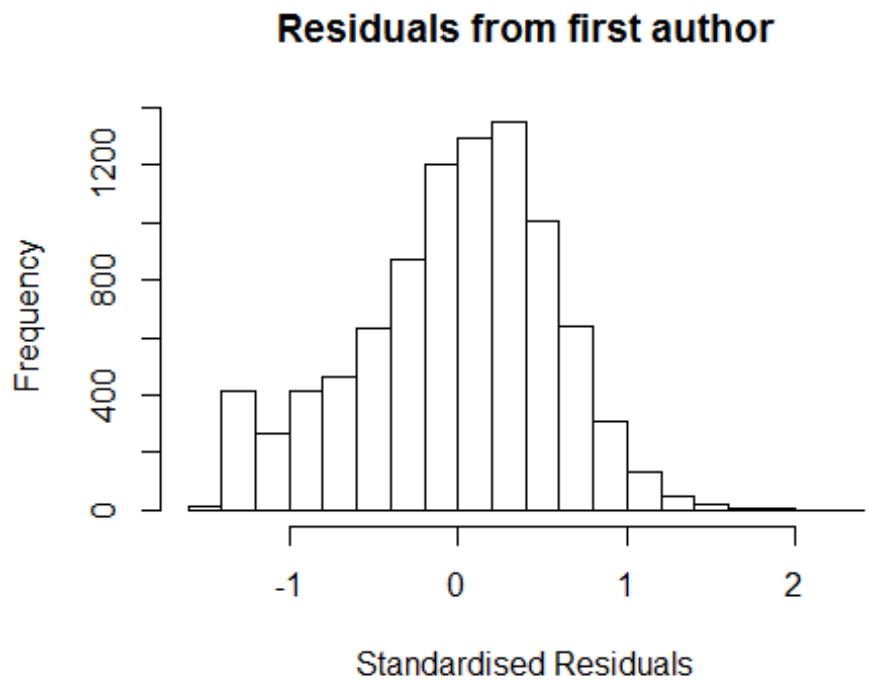

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
UniqueAuthors +
##     Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##     k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.5759 -0.3764  0.0355  0.3751  2.4717
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.10733    0.03527   31.40  < 2e-16 ***
## FirstAuthorFemale1 -0.01200    0.01477   -0.81  0.41660
## LastAuthorFemale1  -0.00557    0.01647   -0.34  0.73530
## UniqueAuthors2      0.26906    0.02057   13.08  < 2e-16 ***
## UniqueAuthors3      0.32883    0.02129   15.45  < 2e-16 ***
## UniqueAuthors4      0.37356    0.02285   16.35  < 2e-16 ***
## UniqueAuthors5      0.42631    0.02210   19.29  < 2e-16 ***
## Year1997           -0.09133    0.04077   -2.24  0.02511 *
## Year1998           -0.15591    0.04306   -3.62  0.00029 ***
## Year1999           -0.16593    0.04007   -4.14  3.5e-05 ***
```

```
## Year2000           -0.12802    0.04812   -2.66  0.00782 **
## Year2001            0.04225    0.04296    0.98  0.32548
## Year2002           -0.05340    0.03960   -1.35  0.17747
## Year2003           -0.03817    0.03885   -0.98  0.32596
## Year2004           -0.10112    0.03990   -2.53  0.01129 *
## Year2005           -0.12469    0.04105   -3.04  0.00239 **
## Year2006           -0.21171    0.03978   -5.32  1.0e-07 ***
## Year2007           -0.15966    0.03964   -4.03  5.7e-05 ***
## Year2008           -0.14864    0.04038   -3.68  0.00023 ***
## Year2009           -0.09753    0.03727   -2.62  0.00889 **
## Year2010           -0.16173    0.03741   -4.32  1.6e-05 ***
## Year2011           -0.15896    0.03734   -4.26  2.1e-05 ***
## Year2012           -0.24029    0.03809   -6.31  2.9e-10 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.545
## Multiple R-squared:  0.0738, Adjusted R-squared:  0.0716
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
##  754 weights are ~= 1. The remaining 8332 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.004   0.859   0.948   0.899   0.985   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.10e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.029  1           1.015
## LastAuthorFemale  1.025  1           1.013
## Year              1.031 16           1.001
```



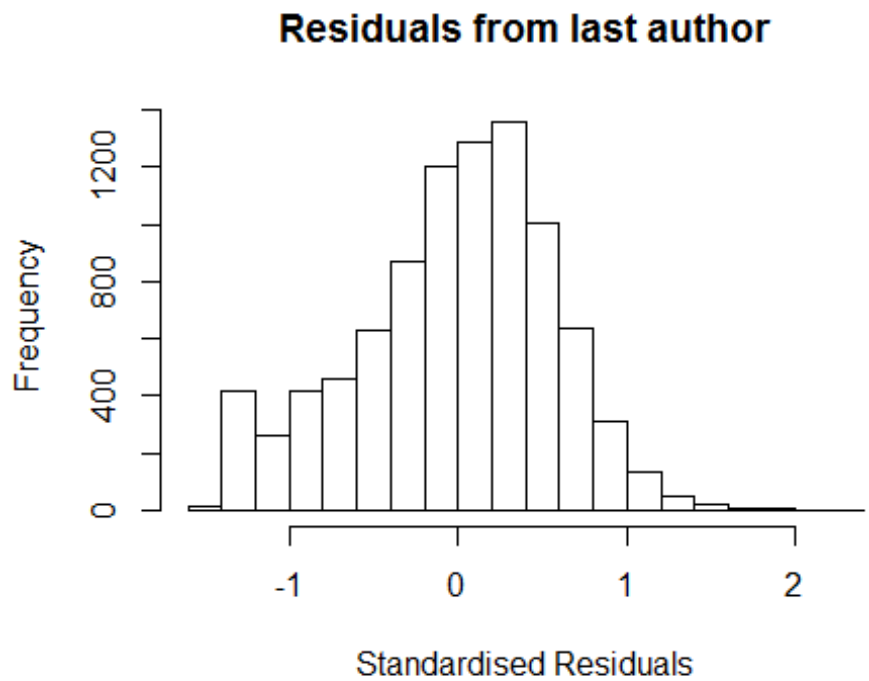

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.4328 -0.3828  0.0403  0.3841  2.2278
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)          1.35122    0.03254   41.53  < 2e-16 ***
## FirstAuthorFemale1   0.01388    0.01517    0.91  0.36030
## LastAuthorFemale1    0.00999    0.01700    0.59  0.55696
## Year1997            -0.11565    0.04175   -2.77  0.00562 **
## Year1998            -0.18527    0.04459   -4.15  3.3e-05 ***
## Year1999            -0.17517    0.04231   -4.14  3.5e-05 ***
## Year2000            -0.14359    0.05039   -2.85  0.00439 **
## Year2001             0.05767    0.04392    1.31  0.18925
## Year2002            -0.03601    0.04111   -0.88  0.38109
## Year2003            -0.02323    0.04016   -0.58  0.56306
## Year2004            -0.09377    0.04141   -2.26  0.02358 *
## Year2005            -0.10755    0.04293   -2.51  0.01225 *
```

```
## Year2006           -0.21008    0.04177   -5.03  5.0e-07 ***
## Year2007           -0.14656    0.04141   -3.54  0.00040 ***
## Year2008           -0.12034    0.04209   -2.86  0.00425 **
## Year2009           -0.06434    0.03876   -1.66  0.09696 .
## Year2010           -0.12855    0.03878   -3.31  0.00092 ***
## Year2011           -0.11538    0.03878   -2.98  0.00293 **
## Year2012           -0.18440    0.03953   -4.67  3.1e-06 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.552
## Multiple R-squared:  0.0133, Adjusted R-squared:  0.0113
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
##  739 weights are ~= 1. The remaining 8347 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0661  0.8570  0.9470  0.8960  0.9840  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.10e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.018  1           1.009
## Year              1.018 16           1.001
```




```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.4256 -0.3841  0.0403  0.3847  2.2267
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)           1.3523     0.0325   41.58  < 2e-16 ***
## FirstAuthorFemale1    0.0158     0.0152    1.04  0.29839
## Year1997             -0.1157     0.0418   -2.77  0.00560 **
## Year1998             -0.1849     0.0446   -4.15  3.4e-05 ***
## Year1999             -0.1749     0.0423   -4.14  3.6e-05 ***
## Year2000             -0.1441     0.0504   -2.86  0.00425 **
## Year2001              0.0576     0.0439    1.31  0.18981
## Year2002             -0.0358     0.0411   -0.87  0.38425
## Year2003             -0.0228     0.0402   -0.57  0.57013
## Year2004             -0.0934     0.0414   -2.26  0.02408 *
## Year2005             -0.1072     0.0429   -2.50  0.01251 *
## Year2006             -0.2095     0.0417   -5.02  5.3e-07 ***
```

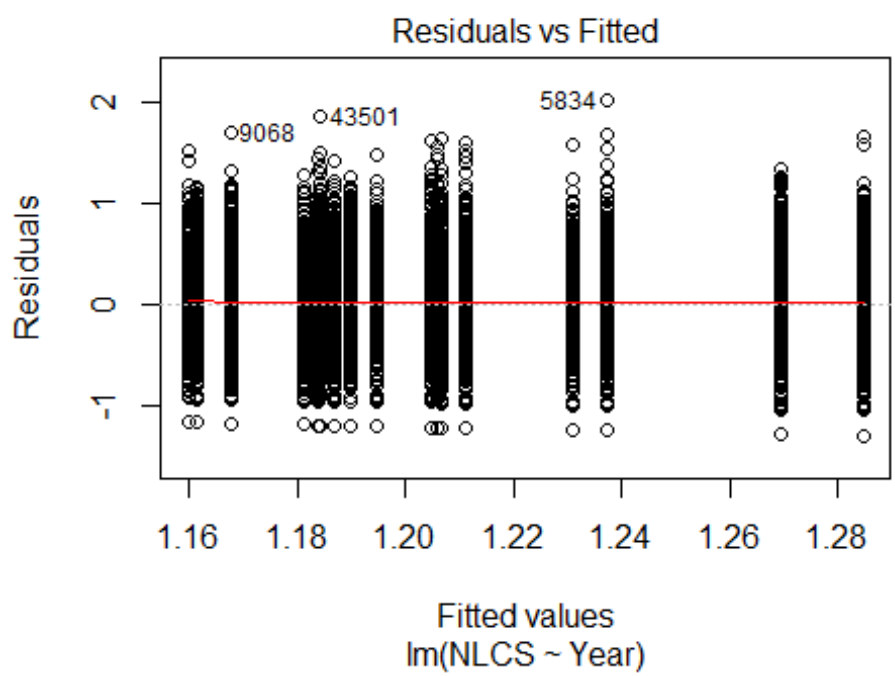
```
## Year2007            -0.1466     0.0414   -3.54  0.00040 ***
## Year2008            -0.1201     0.0421   -2.85  0.00434 **
## Year2009            -0.0640     0.0388   -1.65  0.09857 .
## Year2010            -0.1283     0.0388   -3.31  0.00094 ***
## Year2011            -0.1151     0.0388   -2.97  0.00300 **
## Year2012            -0.1841     0.0395   -4.66  3.2e-06 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.552
## Multiple R-squared:  0.0132, Adjusted R-squared:  0.0114
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
##  739 weights are ~= 1. The remaining 8347 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0664  0.8560  0.9470  0.8960  0.9840  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.10e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##                   GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.014  1           1.007
## Year             1.014 16           1.000
```



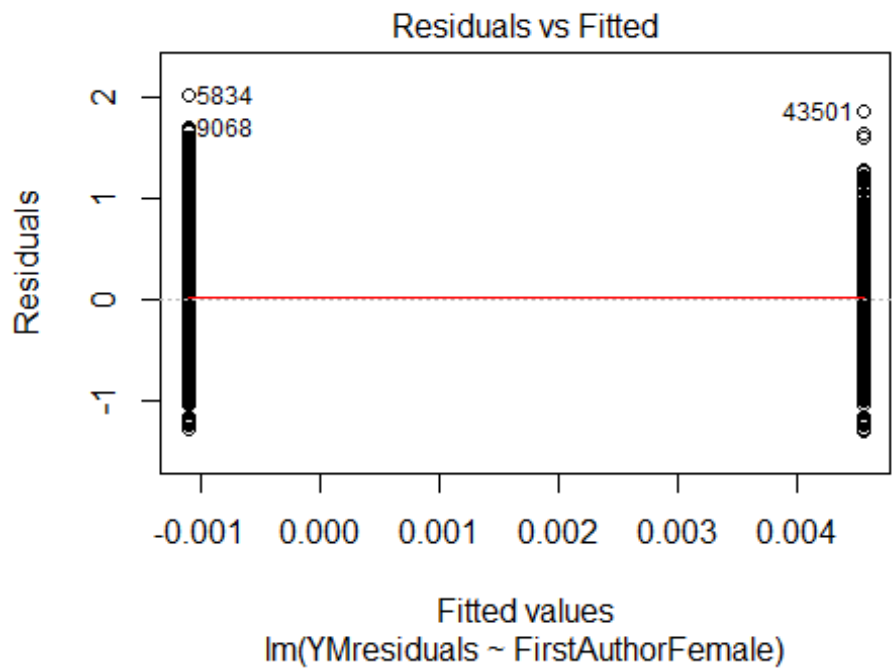

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.4239 -0.3825  0.0404  0.3845  2.2260
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)          1.3530     0.0325   41.65  < 2e-16 ***
## LastAuthorFemale1    0.0130     0.0170    0.77  0.44187
## Year1997            -0.1154     0.0418   -2.76  0.00571 **
## Year1998            -0.1855     0.0446   -4.16  3.2e-05 ***
## Year1999            -0.1751     0.0423   -4.14  3.5e-05 ***
## Year2000            -0.1435     0.0504   -2.85  0.00444 **
## Year2001             0.0578     0.0439    1.32  0.18794
## Year2002            -0.0355     0.0411   -0.86  0.38811
## Year2003            -0.0230     0.0402   -0.57  0.56617
## Year2004            -0.0932     0.0414   -2.25  0.02441 *
## Year2005            -0.1069     0.0429   -2.49  0.01276 *
## Year2006            -0.2097     0.0418   -5.02  5.3e-07 ***
```

```
## Year2007           -0.1463     0.0414   -3.53  0.00041 ***
## Year2008           -0.1193     0.0421   -2.84  0.00456 **
## Year2009           -0.0634     0.0387   -1.64  0.10168
## Year2010           -0.1275     0.0388   -3.29  0.00101 **
## Year2011           -0.1144     0.0388   -2.95  0.00317 **
## Year2012           -0.1834     0.0395   -4.64  3.5e-06 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.552
## Multiple R-squared:  0.0132, Adjusted R-squared:  0.0114
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
##  741 weights are ~= 1. The remaining 8345 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0666  0.8560  0.9470  0.8960  0.9840  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.10e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Sample size for the above analysis:  9086"
## [1] ""
## [1] ""
## [1] "####################################"
## [1] "Analysis of AJSC 1908"
## [1] "####################################"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 1873 2062 2250 2063 1838 2067 2117 2458 2116 1833 2041 1997 2125 1947 1912
## 2011 2012
## 1873 2039
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 1050 1131 1205 1102  812  804 1235 1378 1170 1046 1189 1147 1189 1103 1034
## 2011 2012
```

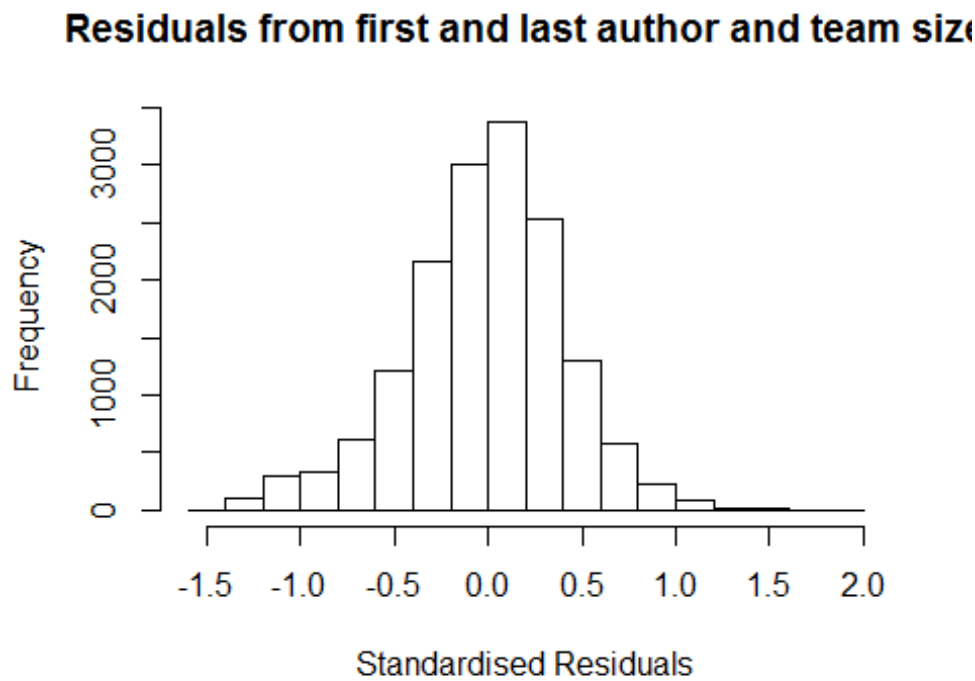
```
## 1012 1119
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  921 1005 1031  950  698  680 1055 1182  993  836  985  939  960  959  883
## 2011 2012
##  853  932
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
##  Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 100, df = 16, p-value = 5e-15
```




```
##
##  Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 48, df = 1, p-value = 4e-12
```

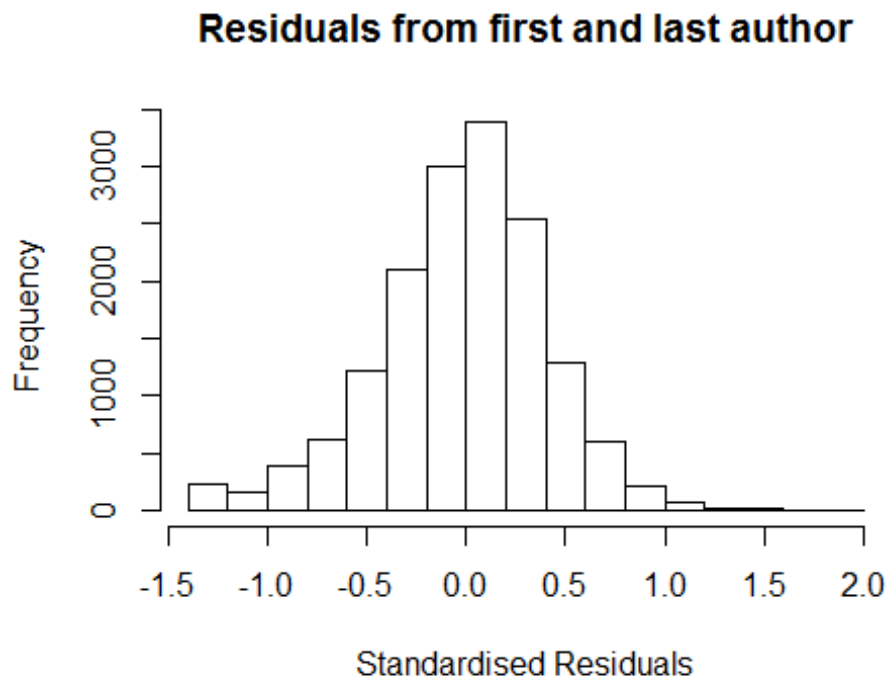



```
## [1] "Female first author team size 2018 geometric mean: 3.39437389425837"
## [1] "Male first author team size 2018 geometric mean: 2.89707513843651"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 43000, p-value = 0.02
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 3.23100237479629"
## [1] "Male last author team size 2018 geometric mean: 2.99143137522552"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 30000, p-value = 0.3
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.042  1           1.021
## LastAuthorFemale  1.032  1           1.016
## UniqueAuthors     1.075  4           1.009
## Year              1.079 16           1.002
```



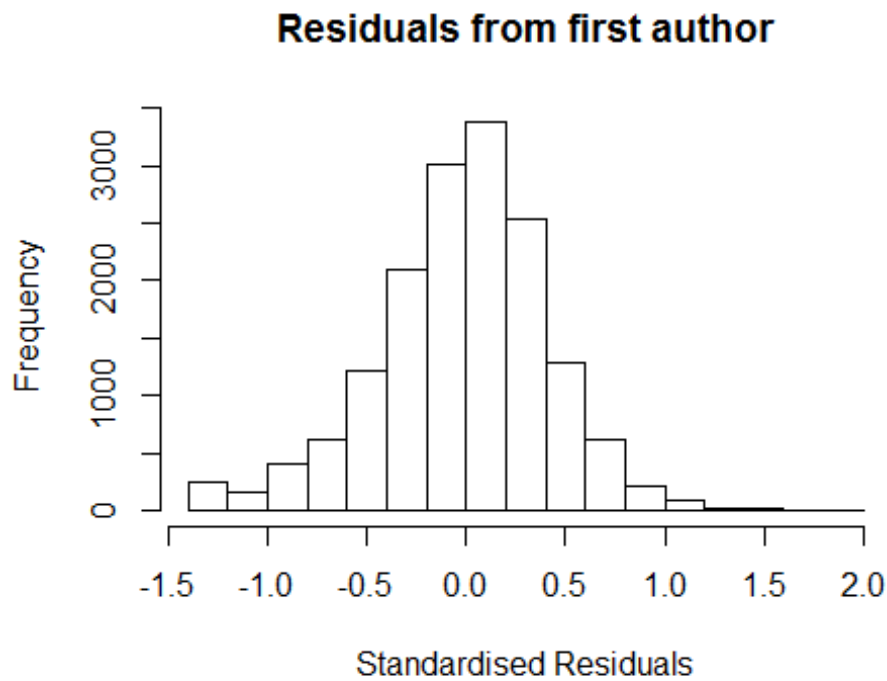

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
UniqueAuthors +
##     Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##     k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.4142 -0.2587  0.0116  0.2567  1.9668
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.19984    0.01629   73.67  < 2e-16 ***
## FirstAuthorFemale1 -0.01153    0.00781   -1.48   0.1402
## LastAuthorFemale1  -0.02799    0.00945   -2.96   0.0030 **
## UniqueAuthors2      0.13744    0.01130   12.17  < 2e-16 ***
## UniqueAuthors3      0.14911    0.01172   12.72  < 2e-16 ***
## UniqueAuthors4      0.17325    0.01285   13.48  < 2e-16 ***
## UniqueAuthors5      0.23380    0.01231   18.99  < 2e-16 ***
## Year1997           -0.01943    0.01956   -0.99   0.3207
## Year1998           -0.05603    0.01983   -2.83   0.0047 **
## Year1999           -0.10756    0.02106   -5.11  3.3e-07 ***
```

```
## Year2000           -0.11765    0.02068   -5.69  1.3e-08 ***
## Year2001           -0.09193    0.02081   -4.42  1.0e-05 ***
## Year2002           -0.14379    0.01959   -7.34  2.2e-13 ***
## Year2003           -0.15943    0.01845   -8.64  < 2e-16 ***
## Year2004           -0.14668    0.01877   -7.81  5.9e-15 ***
## Year2005           -0.13238    0.02016   -6.57  5.3e-11 ***
## Year2006           -0.14062    0.01868   -7.53  5.5e-14 ***
## Year2007           -0.12868    0.01936   -6.65  3.1e-11 ***
## Year2008           -0.11984    0.01920   -6.24  4.4e-10 ***
## Year2009           -0.13631    0.01884   -7.23  4.9e-13 ***
## Year2010           -0.13545    0.01994   -6.79  1.1e-11 ***
## Year2011           -0.13161    0.01986   -6.63  3.6e-11 ***
## Year2012           -0.14571    0.01957   -7.44  1.0e-13 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.38
## Multiple R-squared:  0.0381, Adjusted R-squared:  0.0368
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
##  2 observations c(2365,15485) are outliers with |weight| = 0 ( < 6.3e-06);
##  1346 weights are ~= 1. The remaining 14514 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0014  0.8630  0.9500  0.8910  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          6.30e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.032  1           1.016
## LastAuthorFemale  1.030  1           1.015
## Year              1.018 16           1.001
```



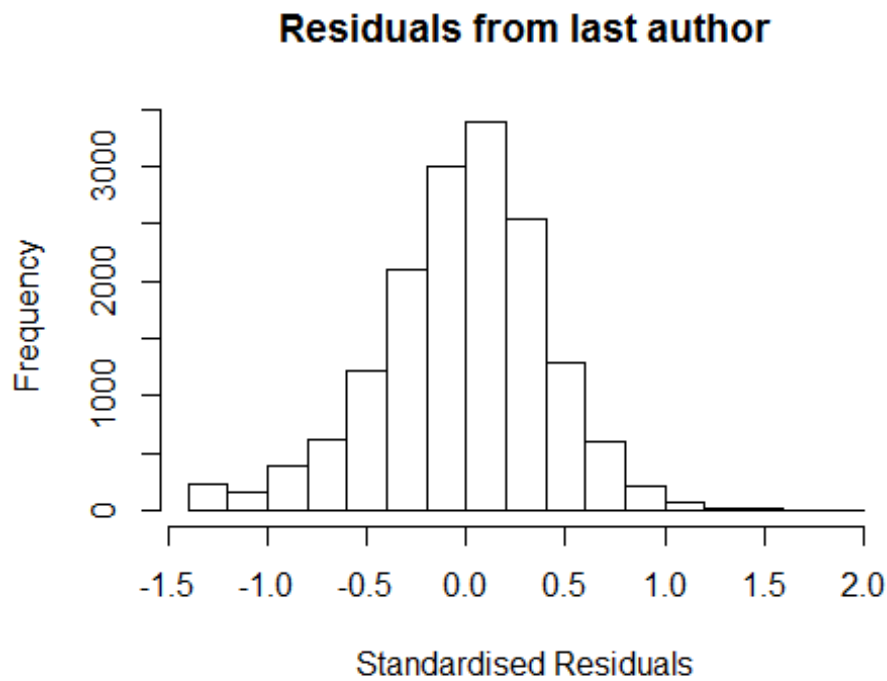

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.3103 -0.2590  0.0126  0.2564  1.9852
##
## Coefficients:
##                      Estimate Std. Error t value Pr(>|t|)
## (Intercept)          1.310275   0.014177   92.42  < 2e-16 ***
## FirstAuthorFemale1  0.000783   0.007825    0.10   0.9203
## LastAuthorFemale1  -0.023129   0.009506   -2.43   0.0150 *
## Year1997           -0.022987   0.019580   -1.17   0.2404
## Year1998           -0.047451   0.019781   -2.40   0.0165 *
## Year1999           -0.098559   0.021000   -4.69  2.7e-06 ***
## Year2000           -0.099872   0.020626   -4.84  1.3e-06 ***
## Year2001           -0.061121   0.020910   -2.92   0.0035 **
## Year2002           -0.118953   0.019403   -6.13  9.0e-10 ***
## Year2003           -0.129780   0.018318   -7.08  1.5e-12 ***
## Year2004           -0.114880   0.018656   -6.16  7.6e-10 ***
## Year2005           -0.104742   0.020139   -5.20  2.0e-07 ***
```

```
## Year2006           -0.108651   0.018607   -5.84  5.4e-09 ***
## Year2007           -0.100567   0.019288   -5.21  1.9e-07 ***
## Year2008           -0.086853   0.019041   -4.56  5.1e-06 ***
## Year2009           -0.100482   0.018765   -5.35  8.7e-08 ***
## Year2010           -0.103255   0.019898   -5.19  2.1e-07 ***
## Year2011           -0.092699   0.019943   -4.65  3.4e-06 ***
## Year2012           -0.111234   0.019591   -5.68  1.4e-08 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.382
## Multiple R-squared:  0.00796,    Adjusted R-squared:  0.00683
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  2 observations c(2365,15485) are outliers with |weight| = 0 ( < 6.3e-06);
##  1337 weights are ~= 1. The remaining 14523 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0256  0.8620  0.9500  0.8890  0.9860  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          6.30e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##                   GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.01  1           1.005
## Year              1.01 16           1.000
```




```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.3080 -0.2585  0.0128  0.2560  1.9874
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.30799    0.01415   92.44  < 2e-16 ***
## FirstAuthorFemale1 -0.00259    0.00776   -0.33   0.7385
## Year1997           -0.02302    0.01958   -1.18   0.2397
## Year1998           -0.04741    0.01978   -2.40   0.0166 *
## Year1999           -0.09900    0.02100   -4.72  2.4e-06 ***
## Year2000           -0.09982    0.02063   -4.84  1.3e-06 ***
## Year2001           -0.06128    0.02092   -2.93   0.0034 **
## Year2002           -0.11920    0.01942   -6.14  8.6e-10 ***
## Year2003           -0.13019    0.01834   -7.10  1.3e-12 ***
## Year2004           -0.11501    0.01867   -6.16  7.5e-10 ***
## Year2005           -0.10485    0.02015   -5.20  2.0e-07 ***
## Year2006           -0.10858    0.01862   -5.83  5.6e-09 ***
```

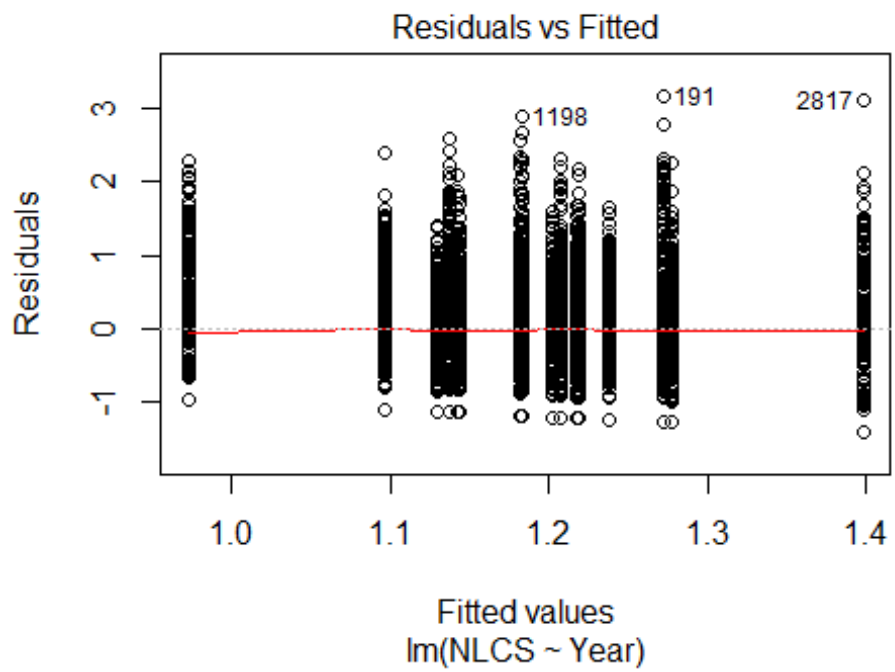
```
## Year2007           -0.10032    0.01929   -5.20  2.0e-07 ***
## Year2008           -0.08679    0.01906   -4.55  5.3e-06 ***
## Year2009           -0.10107    0.01879   -5.38  7.6e-08 ***
## Year2010           -0.10374    0.01989   -5.22  1.9e-07 ***
## Year2011           -0.09320    0.01995   -4.67  3.0e-06 ***
## Year2012           -0.11295    0.01958   -5.77  8.2e-09 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.383
## Multiple R-squared:  0.00757,    Adjusted R-squared:  0.0065
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  2 observations c(2365,15485) are outliers with |weight| = 0 ( < 6.3e-06);
##  1332 weights are ~= 1. The remaining 14528 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.025   0.863   0.950   0.889   0.986   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          6.30e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##                   GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.008  1           1.004
## Year             1.008 16           1.000
```



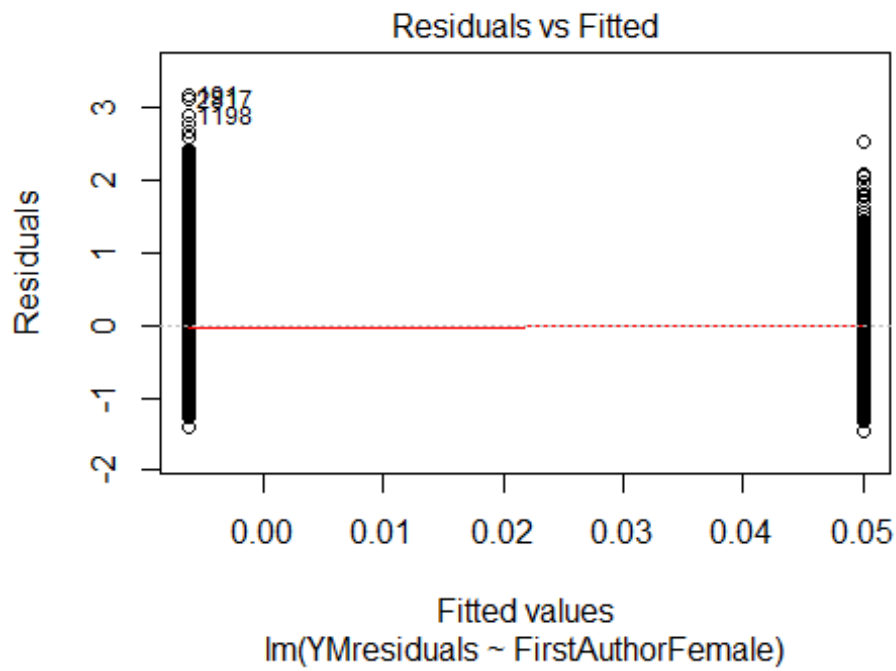

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.3104 -0.2589  0.0128  0.2567  1.9851
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.3104     0.0141   92.68  < 2e-16 ***
## LastAuthorFemale1  -0.0230     0.0094   -2.44   0.0146 *
## Year1997           -0.0230     0.0196   -1.17   0.2407
## Year1998           -0.0474     0.0198   -2.40   0.0165 *
## Year1999           -0.0985     0.0210   -4.69  2.7e-06 ***
## Year2000           -0.0998     0.0206   -4.84  1.3e-06 ***
## Year2001           -0.0611     0.0209   -2.92   0.0035 **
## Year2002           -0.1189     0.0194   -6.13  9.0e-10 ***
## Year2003           -0.1298     0.0183   -7.08  1.5e-12 ***
## Year2004           -0.1148     0.0186   -6.16  7.5e-10 ***
## Year2005           -0.1047     0.0201   -5.20  2.0e-07 ***
## Year2006           -0.1086     0.0186   -5.84  5.3e-09 ***
```

```
## Year2007           -0.1005     0.0193   -5.22  1.8e-07 ***
## Year2008           -0.0868     0.0190   -4.56  5.1e-06 ***
## Year2009           -0.1004     0.0188   -5.35  8.7e-08 ***
## Year2010           -0.1032     0.0199   -5.19  2.1e-07 ***
## Year2011           -0.0926     0.0199   -4.65  3.3e-06 ***
## Year2012           -0.1112     0.0196   -5.68  1.4e-08 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.382
## Multiple R-squared:  0.00796,    Adjusted R-squared:  0.00689
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  2 observations c(2365,15485) are outliers with |weight| = 0 ( < 6.3e-06);
##  1327 weights are ~= 1. The remaining 14533 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0256  0.8620  0.9500  0.8890  0.9860  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          6.30e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Sample size for the above analysis:  15862"
## [1] ""
## [1] ""
## [1] "####################################"
## [1] "Analysis of AJSC 1909"
## [1] "####################################"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  498  561  460  477  407  401  455  427  404  438  517  503  476  511  474
## 2011 2012
##  494  454
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  264  303  202  246  140  167  244  222  233  235  304  274  254  255  271
```

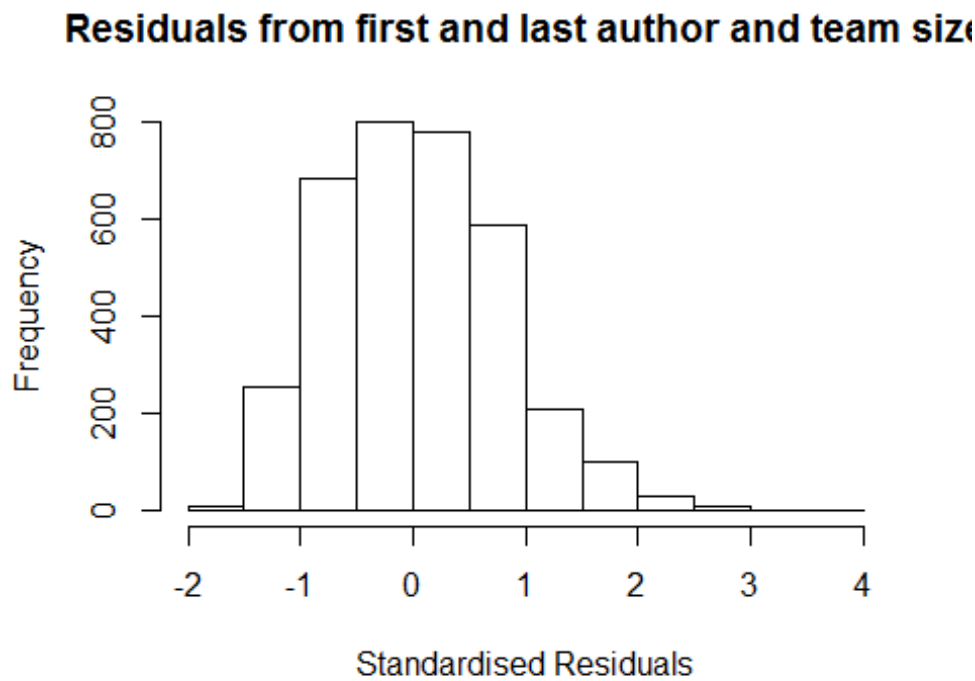
```
## 2011 2012
##  274  262
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  234  272  170  217  115  142  208  191  194  177  239  222  208  207  231
## 2011 2012
##  224  211
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
##  Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 210, df = 16, p-value <2e-16
```




```
##
##  Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 1.4, df = 1, p-value = 0.2
```



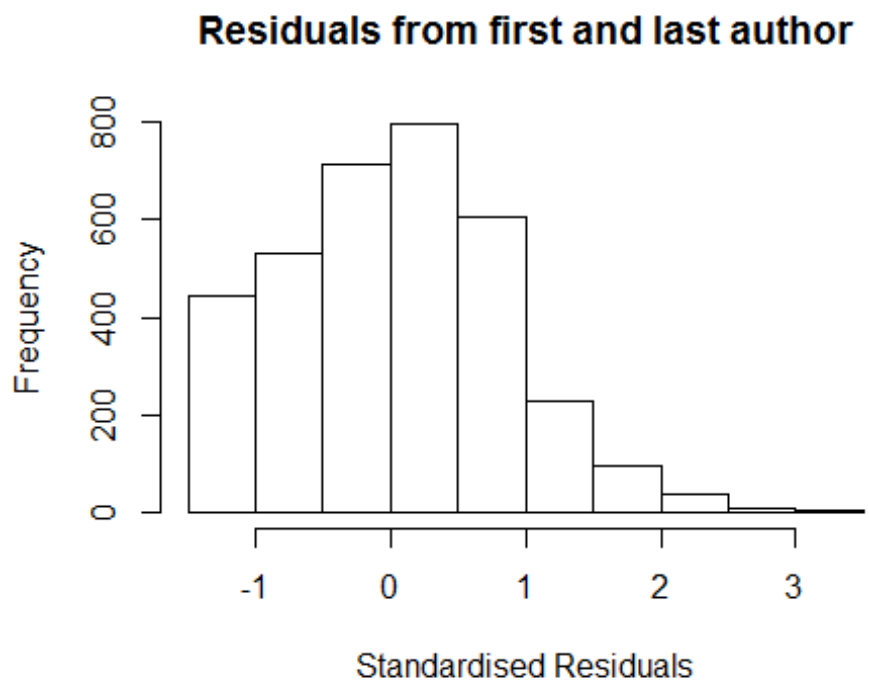

```
## [1] "Female first author team size 2018 geometric mean: 2.71097112891173"
## [1] "Male first author team size 2018 geometric mean: 2.70514922567681"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 8900, p-value = 0.9
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 2.71290033210501"
## [1] "Male last author team size 2018 geometric mean: 2.70493283910101"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 8100, p-value = 0.8
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.104  1           1.051
## LastAuthorFemale  1.106  1           1.052
## UniqueAuthors     1.179  4           1.021
## Year              1.193 16           1.006
```




```
## [1] "List of  10 outliers with residuals above  2.5"
##         ScopusId  NLCS Year OneField Fields residuals
## 55    0030406862 4.040 1996     1909      2     3.195
## 140   0030265174 3.579 1996     1907      2     2.734
## 149   0030230285 4.050 1996     1909      1     2.669
## 191   0001690750 4.435 1996     1909      2     3.590
## 1198  0005194452 4.068 1998     1909      1     2.763
## 1451  0002459140 3.851 1998     1909      1     2.546
## 2817  0035492228 4.506 2001     1909      2     2.993
## 3440  0036009678 3.374 2002     1909      1     2.621
## 4049  0038537573 3.410 2003     1909      2     2.684
## 9239 84861812124 3.270 2012     1111      4     2.548
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
UniqueAuthors +
##     Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##     k.max = 1000))
##  \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.52065 -0.56228 -0.00924  0.53869  3.58956
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)           0.8454     0.0669   12.64   <2e-16 ***
```

```
## FirstAuthorFemale1   0.0212     0.0444    0.48   0.6330
## LastAuthorFemale1   -0.0236     0.0450   -0.53   0.5991
## UniqueAuthors2       0.5435     0.0376   14.44   <2e-16 ***
## UniqueAuthors3       0.4973     0.0406   12.25   <2e-16 ***
## UniqueAuthors4       0.5358     0.0491   10.91   <2e-16 ***
## UniqueAuthors5       0.4898     0.0574    8.53   <2e-16 ***
## Year1997            -0.1317     0.0880   -1.50   0.1346
## Year1998            -0.0376     0.0950   -0.40   0.6922
## Year1999            -0.2832     0.0889   -3.18   0.0015 **
## Year2000            -0.1100     0.1012   -1.09   0.2773
## Year2001             0.1317     0.1043    1.26   0.2070
## Year2002            -0.0924     0.0896   -1.03   0.3025
## Year2003            -0.1190     0.0904   -1.32   0.1880
## Year2004            -0.0874     0.0874   -1.00   0.3176
## Year2005            -0.0438     0.0876   -0.50   0.6175
## Year2006            -0.1877     0.0847   -2.22   0.0267 *
## Year2007            -0.0271     0.0792   -0.34   0.7326
## Year2008            -0.0597     0.0806   -0.74   0.4590
## Year2009            -0.1195     0.0778   -1.54   0.1247
## Year2010            -0.0394     0.0753   -0.52   0.6004
## Year2011            -0.1468     0.0766   -1.92   0.0554 .
## Year2012            -0.1238     0.0785   -1.58   0.1151
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.748
## Multiple R-squared:  0.0997, Adjusted R-squared:  0.094
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
##  observation 104 is an outlier with |weight| = 0 ( < 2.9e-05);
##  260 weights are ~= 1. The remaining 3201 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0284  0.8830  0.9420  0.9040  0.9830  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          2.89e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                    psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
```

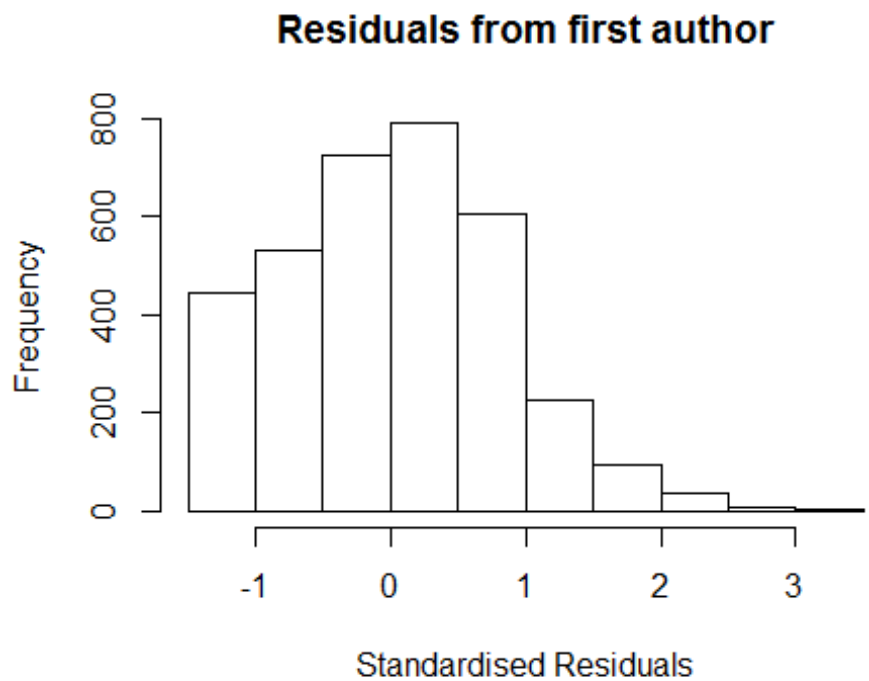
```
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##                     GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.054  1           1.027
## LastAuthorFemale  1.053  1           1.026
## Year              1.034 16           1.001
```




```
## [1] "List of  9 outliers with residuals above  2.5"
##        ScopusId  NLCS Year OneField Fields residuals
## 55   0030406862 4.040 1996     1909      2     2.885
## 149  0030230285 4.050 1996     1909      1     2.895
## 191  0001690750 4.435 1996     1909      2     3.280
## 581  0031429740 3.721 1997     1909      1     2.724
## 781  0031191272 3.548 1997     1909      1     2.551
## 1198 0005194452 4.068 1998     1909      1     2.984
## 1451 0002459140 3.851 1998     1909      1     2.767
## 2817 0035492228 4.506 2001     1909      2     3.169
## 3291 0036748892 3.751 2002     1909      1     2.572
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.4343 -0.5872  0.0161  0.5648  3.2797
```

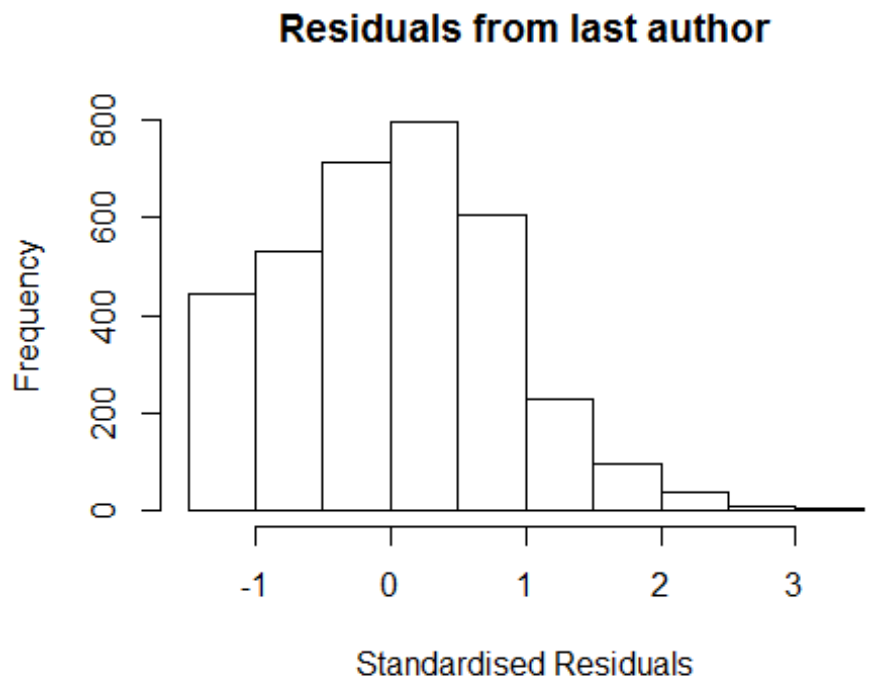
```
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)          1.1553     0.0684   16.90   <2e-16 ***
## FirstAuthorFemale1   0.0792     0.0453    1.75    0.080 .
## LastAuthorFemale1    0.0180     0.0466    0.39    0.699
## Year1997            -0.1587     0.0945   -1.68    0.093 .
## Year1998            -0.0708     0.1019   -0.70    0.487
## Year1999            -0.2887     0.0932   -3.10    0.002 **
## Year2000            -0.1096     0.1038   -1.06    0.291
## Year2001             0.1816     0.1064    1.71    0.088 .
## Year2002            -0.0558     0.0946   -0.59    0.556
## Year2003            -0.0511     0.0929   -0.55    0.582
## Year2004            -0.0170     0.0922   -0.18    0.854
## Year2005             0.0534     0.0915    0.58    0.560
## Year2006            -0.1106     0.0880   -1.26    0.209
## Year2007             0.0521     0.0844    0.62    0.537
## Year2008             0.0295     0.0857    0.34    0.730
## Year2009            -0.0159     0.0817   -0.19    0.846
## Year2010             0.0716     0.0802    0.89    0.372
## Year2011            -0.0296     0.0805   -0.37    0.713
## Year2012            -0.0136     0.0831   -0.16    0.870
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.824
## Multiple R-squared:  0.016,  Adjusted R-squared:  0.0108
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
##  300 weights are ~= 1. The remaining 3162 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0775  0.8660  0.9460  0.9120  0.9840  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          2.89e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                  psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
```

```
##                     GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.016  1           1.008
## Year              1.016 16           1.001
```




```
## [1] "List of  9 outliers with residuals above  2.5"
##        ScopusId  NLCS Year OneField Fields residuals
## 55   0030406862 4.040 1996     1909      2     2.885
## 149  0030230285 4.050 1996     1909      1     2.895
## 191  0001690750 4.435 1996     1909      2     3.280
## 581  0031429740 3.721 1997     1909      1     2.724
## 781  0031191272 3.548 1997     1909      1     2.551
## 1198 0005194452 4.068 1998     1909      1     2.984
## 1451 0002459140 3.851 1998     1909      1     2.767
## 2817 0035492228 4.506 2001     1909      2     3.169
## 3291 0036748892 3.751 2002     1909      1     2.572
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.4207 -0.5881  0.0149  0.5639  3.2784
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
```

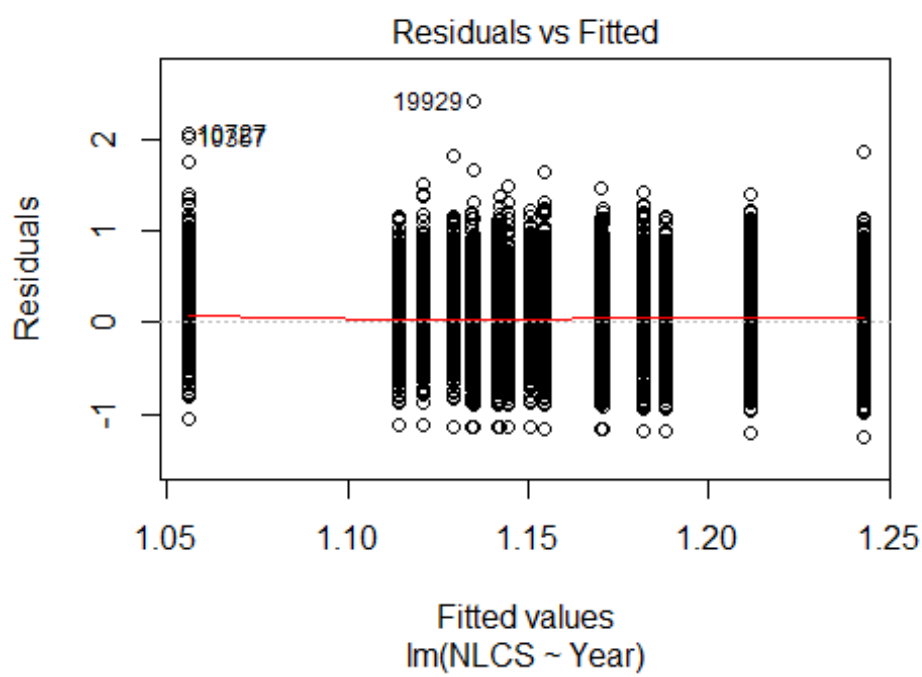
```
## (Intercept)          1.1566     0.0683   16.92   <2e-16 ***
## FirstAuthorFemale1   0.0830     0.0444    1.87    0.062 .
## Year1997            -0.1582     0.0945   -1.67    0.094 .
## Year1998            -0.0709     0.1019   -0.70    0.486
## Year1999            -0.2878     0.0931   -3.09    0.002 **
## Year2000            -0.1096     0.1038   -1.06    0.291
## Year2001             0.1810     0.1065    1.70    0.089 .
## Year2002            -0.0555     0.0942   -0.59    0.556
## Year2003            -0.0518     0.0928   -0.56    0.577
## Year2004            -0.0174     0.0922   -0.19    0.851
## Year2005             0.0531     0.0915    0.58    0.561
## Year2006            -0.1104     0.0880   -1.26    0.209
## Year2007             0.0524     0.0843    0.62    0.534
## Year2008             0.0298     0.0857    0.35    0.728
## Year2009            -0.0159     0.0817   -0.19    0.846
## Year2010             0.0720     0.0802    0.90    0.369
## Year2011            -0.0295     0.0805   -0.37    0.714
## Year2012            -0.0139     0.0832   -0.17    0.867
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.825
## Multiple R-squared:  0.0159, Adjusted R-squared:  0.0111
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
##  305 weights are ~= 1. The remaining 3157 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0787  0.8660  0.9460  0.9120  0.9840  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          2.89e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##                   GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.016  1           1.008
## Year             1.016 16           1.000
```



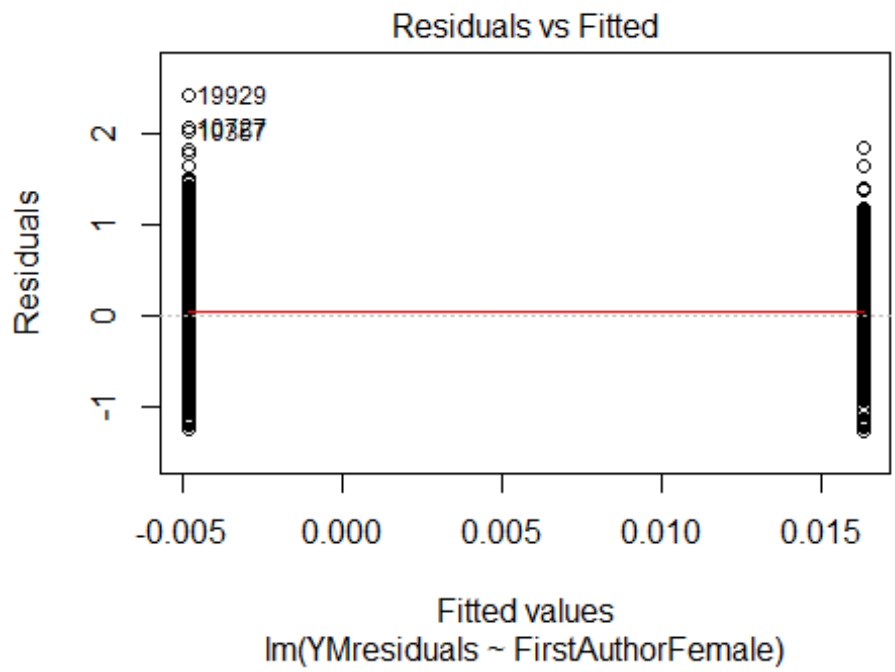

```
## [1] "List of  9 outliers with residuals above  2.5"
##        ScopusId  NLCS Year OneField Fields residuals
## 55   0030406862 4.040 1996     1909      2     2.885
## 149  0030230285 4.050 1996     1909      1     2.895
## 191  0001690750 4.435 1996     1909      2     3.280
## 581  0031429740 3.721 1997     1909      1     2.724
## 781  0031191272 3.548 1997     1909      1     2.551
## 1198 0005194452 4.068 1998     1909      1     2.984
## 1451 0002459140 3.851 1998     1909      1     2.767
## 2817 0035492228 4.506 2001     1909      2     3.169
## 3291 0036748892 3.751 2002     1909      1     2.572
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.3817 -0.5912  0.0164  0.5734  3.2752
##
## Coefficients:
##                   Estimate Std. Error t value Pr(>|t|)
## (Intercept)        1.15984    0.06824   17.00   <2e-16 ***
## LastAuthorFemale1  0.03708    0.04548    0.82   0.4150
## Year1997          -0.15798    0.09447   -1.67   0.0946 .
## Year1998          -0.07159    0.10180   -0.70   0.4820
```

```
## Year1999          -0.28919    0.09315   -3.10   0.0019 **
## Year2000          -0.10980    0.10384   -1.06   0.2904
## Year2001           0.18479    0.10661    1.73   0.0831 .
## Year2002          -0.05480    0.09435   -0.58   0.5614
## Year2003          -0.04764    0.09299   -0.51   0.6085
## Year2004          -0.01062    0.09206   -0.12   0.9082
## Year2005           0.05826    0.09118    0.64   0.5229
## Year2006          -0.10743    0.08809   -1.22   0.2227
## Year2007           0.05638    0.08434    0.67   0.5039
## Year2008           0.03367    0.08555    0.39   0.6939
## Year2009          -0.01630    0.08162   -0.20   0.8417
## Year2010           0.07201    0.08028    0.90   0.3698
## Year2011          -0.02421    0.08046   -0.30   0.7635
## Year2012          -0.00786    0.08312   -0.09   0.9246
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.825
## Multiple R-squared:  0.0151, Adjusted R-squared:  0.0103
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
##  303 weights are ~= 1. The remaining 3159 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0793  0.8650  0.9450  0.9120  0.9840  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          2.89e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                    psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Sample size for the above analysis:  3462"
## [1] ""
## [1] ""
## [1] "####################################"
## [1] "Analysis of AJSC 1910"
## [1] "####################################"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
```

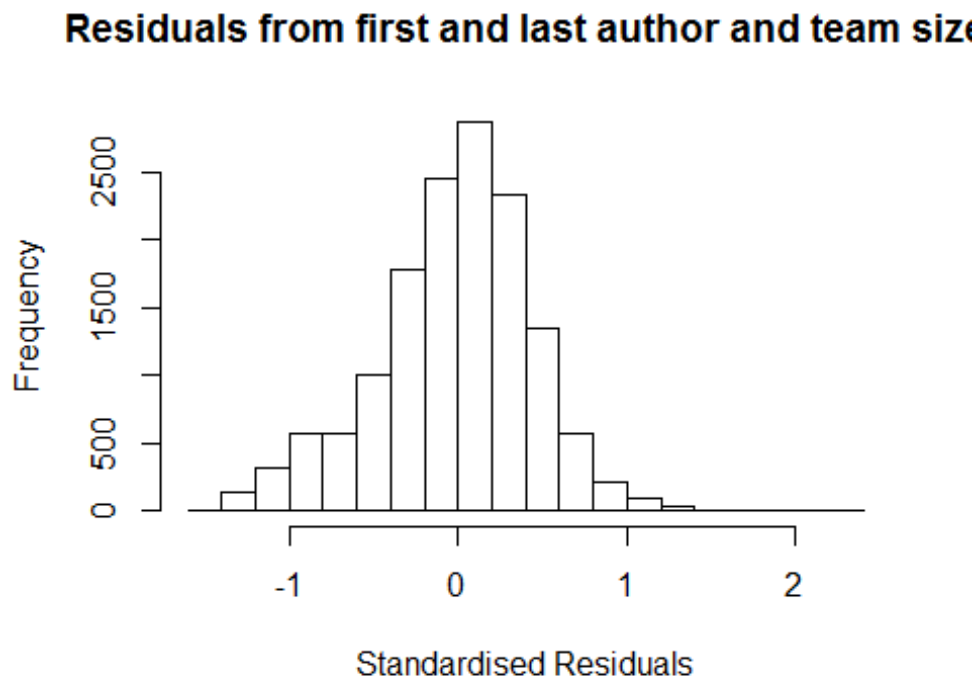
```
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 1603 1742 1820 1704 1725 1983 1727 1930 1529 1551 1640 1667 1730 1436 1445
## 2011 2012
## 1399 1410
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 1007  977  991  964  832  934 1077 1141  949  965 1031 1061 1061  904  870
## 2011 2012
##  856  870
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  915  892  864  858  721  800  945  974  810  800  882  902  895  788  754
## 2011 2012
##  737  750
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
##  Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 290, df = 16, p-value <2e-16
```




```
##
##  Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 68, df = 1, p-value <2e-16
```

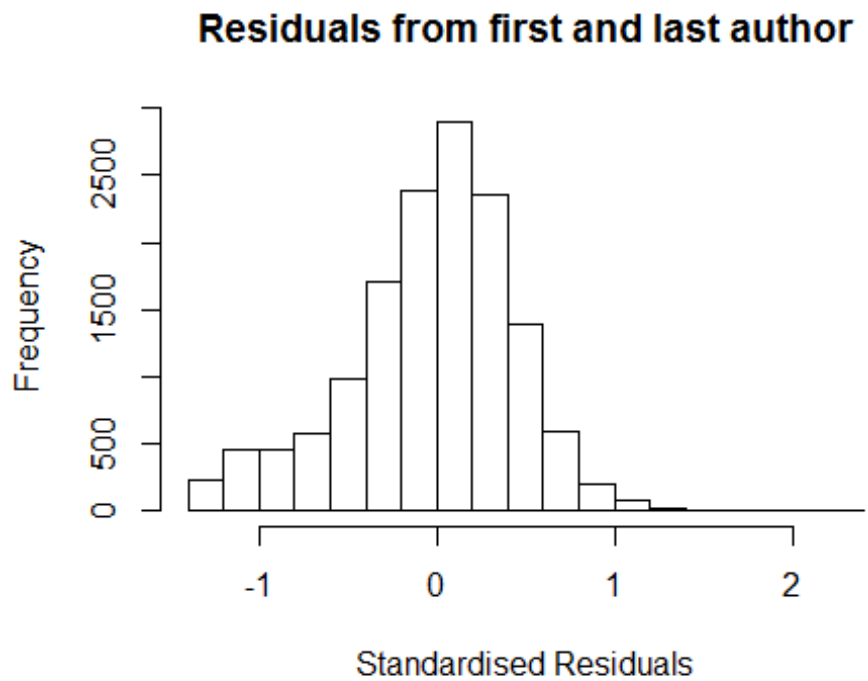



```
## [1] "Female first author team size 2018 geometric mean: 3.87268288467375"
## [1] "Male first author team size 2018 geometric mean: 3.29557489719588"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 55000, p-value = 0.03
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 3.84085371104486"
## [1] "Male last author team size 2018 geometric mean: 3.39278892601779"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 47000, p-value = 0.04
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.032  1           1.016
## LastAuthorFemale  1.020  1           1.010
## UniqueAuthors     1.071  4           1.009
## Year              1.074 16           1.002
```



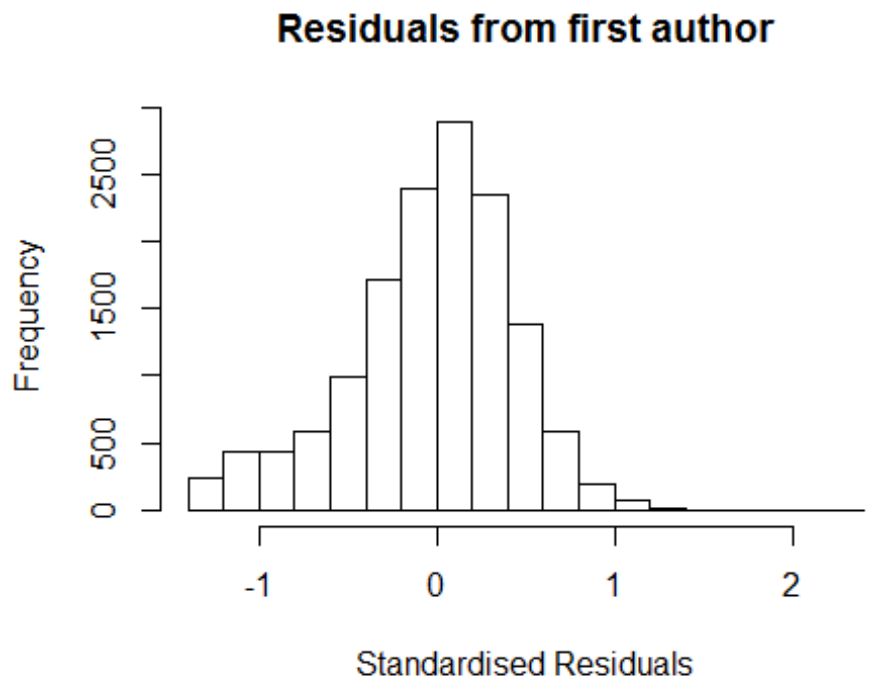

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
UniqueAuthors +
##     Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##     k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.4132 -0.2790  0.0198  0.2769  2.3853
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.103166   0.018251   60.44  < 2e-16 ***
## FirstAuthorFemale1  0.000732   0.008301    0.09   0.9298
## LastAuthorFemale1  -0.021743   0.009903   -2.20   0.0281 *
## UniqueAuthors2      0.214374   0.013462   15.92  < 2e-16 ***
## UniqueAuthors3      0.248225   0.013977   17.76  < 2e-16 ***
## UniqueAuthors4      0.284138   0.015000   18.94  < 2e-16 ***
## UniqueAuthors5      0.337998   0.014257   23.71  < 2e-16 ***
## Year1997           -0.027978   0.022433   -1.25   0.2123
## Year1998           -0.059808   0.023439   -2.55   0.0107 *
## Year1999           -0.066252   0.022126   -2.99   0.0028 **
```

```
## Year2000           -0.141181   0.024055   -5.87  4.5e-09 ***
## Year2001           -0.226651   0.026065   -8.70  < 2e-16 ***
## Year2002           -0.110537   0.022672   -4.88  1.1e-06 ***
## Year2003           -0.153086   0.020424   -7.50  7.0e-14 ***
## Year2004           -0.164371   0.021007   -7.82  5.4e-15 ***
## Year2005           -0.158883   0.021061   -7.54  4.8e-14 ***
## Year2006           -0.195638   0.020418   -9.58  < 2e-16 ***
## Year2007           -0.134700   0.020052   -6.72  1.9e-11 ***
## Year2008           -0.169474   0.020411   -8.30  < 2e-16 ***
## Year2009           -0.183931   0.021127   -8.71  < 2e-16 ***
## Year2010           -0.150564   0.021470   -7.01  2.4e-12 ***
## Year2011           -0.175580   0.021487   -8.17  3.3e-16 ***
## Year2012           -0.196090   0.022351   -8.77  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.405
## Multiple R-squared:  0.0713, Adjusted R-squared:  0.0699
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
##  4 observations c(580,4438,4779,8316)
##   are outliers with |weight| = 0 ( < 7e-06);
##  1256 weights are ~= 1. The remaining 13027 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0321  0.8580  0.9480  0.8870  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          7.00e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.023  1           1.011
## LastAuthorFemale  1.016  1           1.008
## Year              1.014 16           1.000
```



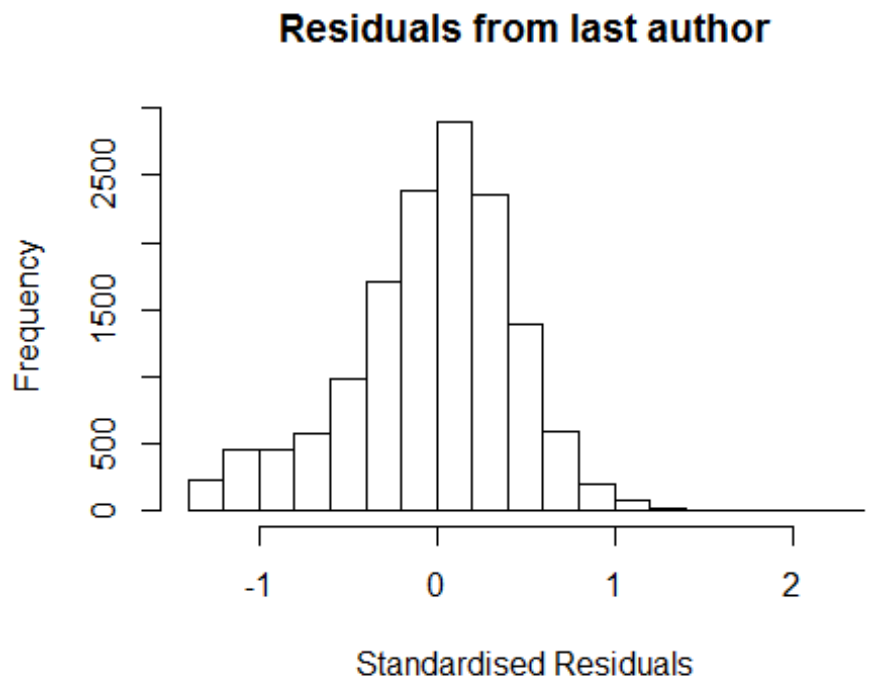

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##    Min     1Q Median     3Q    Max
## -1.304 -0.282  0.027  0.279  2.392
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)          1.28339    0.01575   81.50  < 2e-16 ***
## FirstAuthorFemale1  0.02088    0.00844    2.47  0.01338 *
## LastAuthorFemale1  -0.01660    0.01014   -1.64  0.10156
## Year1997           -0.03517    0.02284   -1.54  0.12359
## Year1998           -0.04836    0.02387   -2.03  0.04281 *
## Year1999           -0.05684    0.02252   -2.52  0.01163 *
## Year2000           -0.12552    0.02459   -5.10  3.4e-07 ***
## Year2001           -0.20110    0.02730   -7.37  1.8e-13 ***
## Year2002           -0.07860    0.02273   -3.46  0.00055 ***
## Year2003           -0.11017    0.02063   -5.34  9.4e-08 ***
## Year2004           -0.12357    0.02125   -5.81  6.2e-09 ***
## Year2005           -0.13151    0.02143   -6.14  8.7e-10 ***
```

```
## Year2006           -0.15860    0.02085   -7.61  3.0e-14 ***
## Year2007           -0.09666    0.02039   -4.74  2.2e-06 ***
## Year2008           -0.12556    0.02080   -6.04  1.6e-09 ***
## Year2009           -0.13393    0.02156   -6.21  5.3e-10 ***
## Year2010           -0.10580    0.02182   -4.85  1.3e-06 ***
## Year2011           -0.12159    0.02183   -5.57  2.6e-08 ***
## Year2012           -0.14089    0.02272   -6.20  5.7e-10 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.41
## Multiple R-squared:  0.012,  Adjusted R-squared:  0.0107
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
##  2 observations c(4438,8316) are outliers with |weight| = 0 ( < 7e-06);
##  1143 weights are ~= 1. The remaining 13142 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0116  0.8580  0.9490  0.8850  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          7.00e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##                   GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.01  1           1.005
## Year              1.01 16           1.000
```




```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.2996 -0.2826  0.0271  0.2790  2.3943
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)          1.28122    0.01568   81.73  < 2e-16 ***
## FirstAuthorFemale1   0.01836    0.00842    2.18  0.02912 *
## Year1997            -0.03487    0.02283   -1.53  0.12668
## Year1998            -0.04848    0.02386   -2.03  0.04223 *
## Year1999            -0.05690    0.02252   -2.53  0.01153 *
## Year2000            -0.12542    0.02459   -5.10  3.4e-07 ***
## Year2001            -0.20086    0.02730   -7.36  2.0e-13 ***
## Year2002            -0.07844    0.02273   -3.45  0.00056 ***
## Year2003            -0.11020    0.02063   -5.34  9.4e-08 ***
## Year2004            -0.12373    0.02127   -5.82  6.2e-09 ***
## Year2005            -0.13156    0.02146   -6.13  8.9e-10 ***
## Year2006            -0.15836    0.02085   -7.60  3.2e-14 ***
```

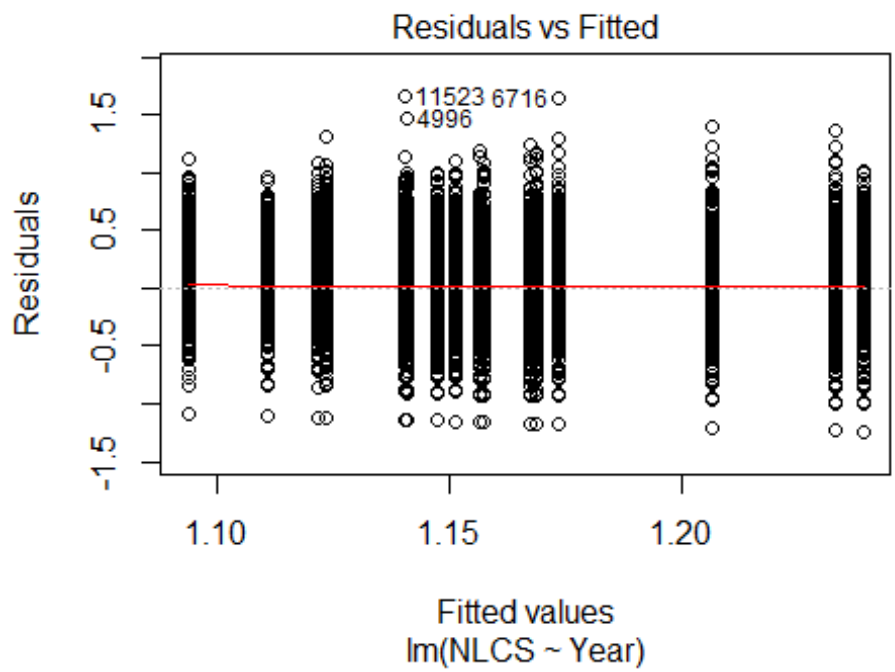
```
## Year2007           -0.09669    0.02040   -4.74  2.2e-06 ***
## Year2008           -0.12572    0.02079   -6.05  1.5e-09 ***
## Year2009           -0.13405    0.02157   -6.21  5.3e-10 ***
## Year2010           -0.10611    0.02182   -4.86  1.2e-06 ***
## Year2011           -0.12185    0.02182   -5.59  2.4e-08 ***
## Year2012           -0.14161    0.02271   -6.24  4.6e-10 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.41
## Multiple R-squared:  0.0118, Adjusted R-squared:  0.0106
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
##  2 observations c(4438,8316) are outliers with |weight| = 0 ( < 7e-06);
##  1151 weights are ~= 1. The remaining 13134 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0147  0.8580  0.9490  0.8850  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          7.00e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##                   GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.004  1           1.002
## Year             1.004 16           1.000
```



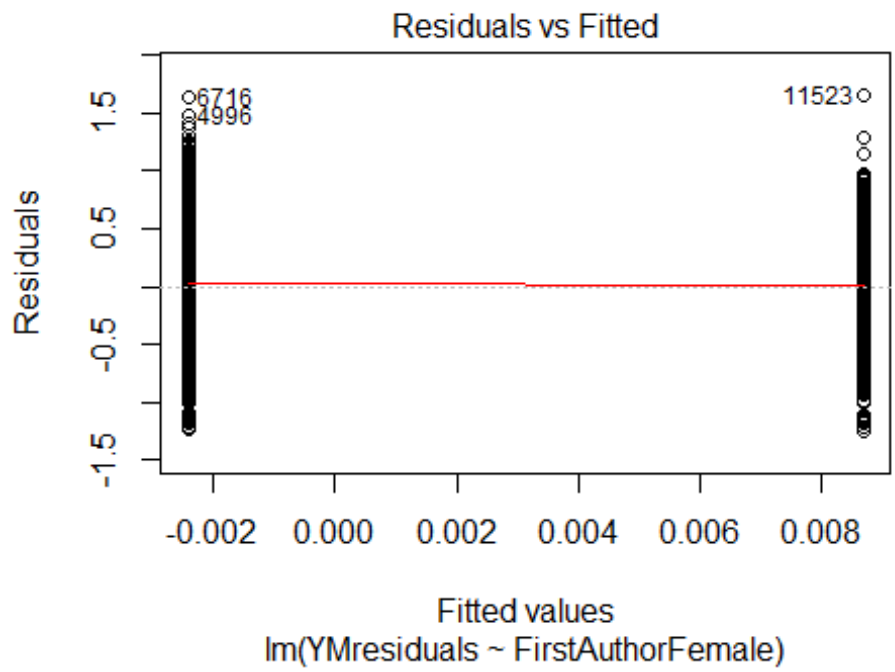

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.2862 -0.2810  0.0275  0.2793  2.3873
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.2862     0.0157   82.13  < 2e-16 ***
## LastAuthorFemale1  -0.0125     0.0101   -1.24  0.21520
## Year1997           -0.0347     0.0228   -1.52  0.12920
## Year1998           -0.0475     0.0239   -1.99  0.04658 *
## Year1999           -0.0565     0.0225   -2.51  0.01206 *
## Year2000           -0.1243     0.0246   -5.06  4.3e-07 ***
## Year2001           -0.2004     0.0273   -7.33  2.4e-13 ***
## Year2002           -0.0769     0.0227   -3.39  0.00071 ***
## Year2003           -0.1091     0.0206   -5.29  1.2e-07 ***
## Year2004           -0.1218     0.0212   -5.73  1.0e-08 ***
## Year2005           -0.1295     0.0214   -6.05  1.5e-09 ***
## Year2006           -0.1562     0.0208   -7.50  6.7e-14 ***
```

```
## Year2007           -0.0946     0.0204   -4.64  3.5e-06 ***
## Year2008           -0.1232     0.0208   -5.93  3.1e-09 ***
## Year2009           -0.1314     0.0215   -6.11  1.0e-09 ***
## Year2010           -0.1033     0.0218   -4.74  2.1e-06 ***
## Year2011           -0.1190     0.0218   -5.47  4.7e-08 ***
## Year2012           -0.1379     0.0226   -6.09  1.2e-09 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.41
## Multiple R-squared:  0.0116, Adjusted R-squared:  0.0104
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
##  2 observations c(4438,8316) are outliers with |weight| = 0 ( < 7e-06);
##  1136 weights are ~= 1. The remaining 13149 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0097  0.8580  0.9490  0.8850  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          7.00e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                    psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Sample size for the above analysis:  14287"
## [1] ""
## [1] ""
## [1] "####################################"
## [1] "Analysis of AJSC 1911"
## [1] "####################################"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  928 1069  916 1019  887 1088  894 1075  852  850  886  988 1003  869  932
## 2011 2012
##  871  911
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  579  565  507  572  450  497  545  590  488  511  540  608  563  516  512
```

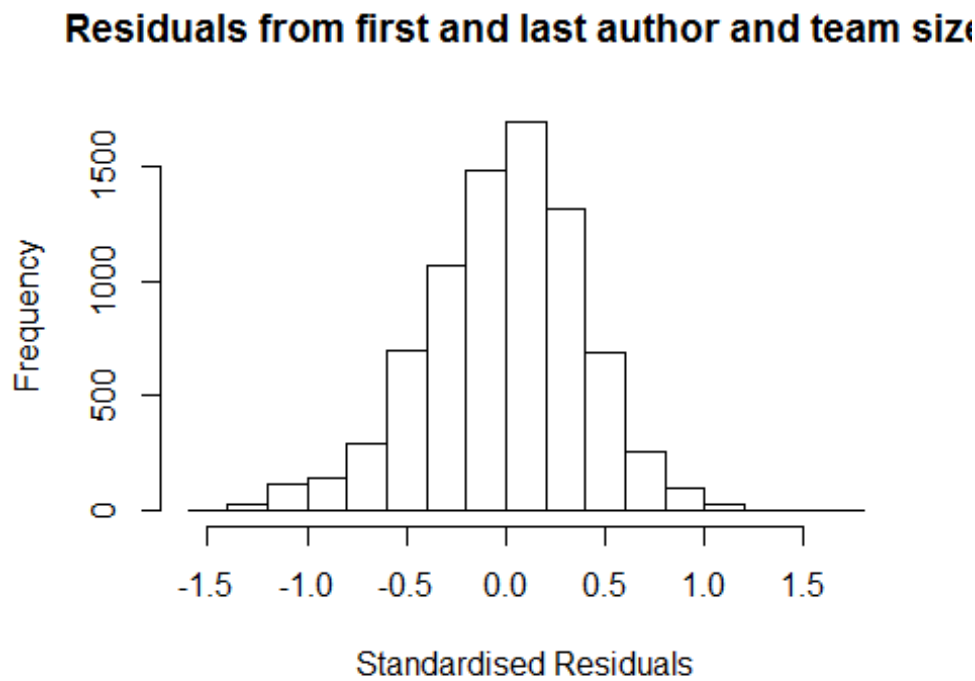
```
## 2011 2012
##  496  512
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  522  521  446  516  400  425  472  514  420  437  453  517  468  457  444
## 2011 2012
##  436  445
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
##  Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 72, df = 16, p-value = 5e-09
```




```
##
##  Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 18, df = 1, p-value = 2e-05
```

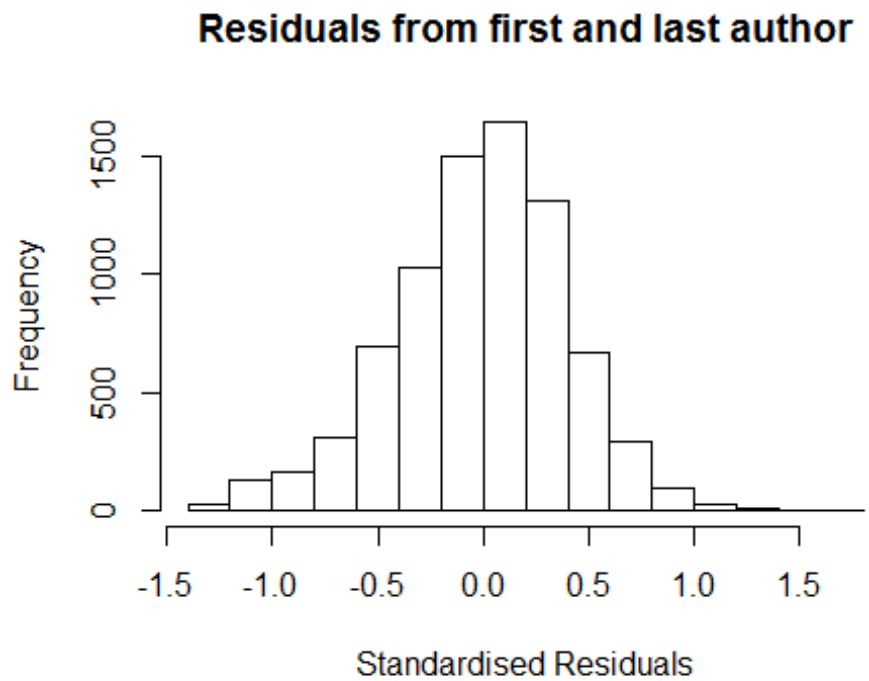



```
## [1] "Female first author team size 2018 geometric mean: 3.61270900180009"
## [1] "Male first author team size 2018 geometric mean: 2.96180762713727"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 24000, p-value = 0.01
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 3.32784065119325"
## [1] "Male last author team size 2018 geometric mean: 3.13354027729228"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 17000, p-value = 0.4
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##                     GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.068  1           1.034
## LastAuthorFemale  1.048  1           1.023
## UniqueAuthors     1.088  4           1.011
## Year              1.092 16           1.003
```



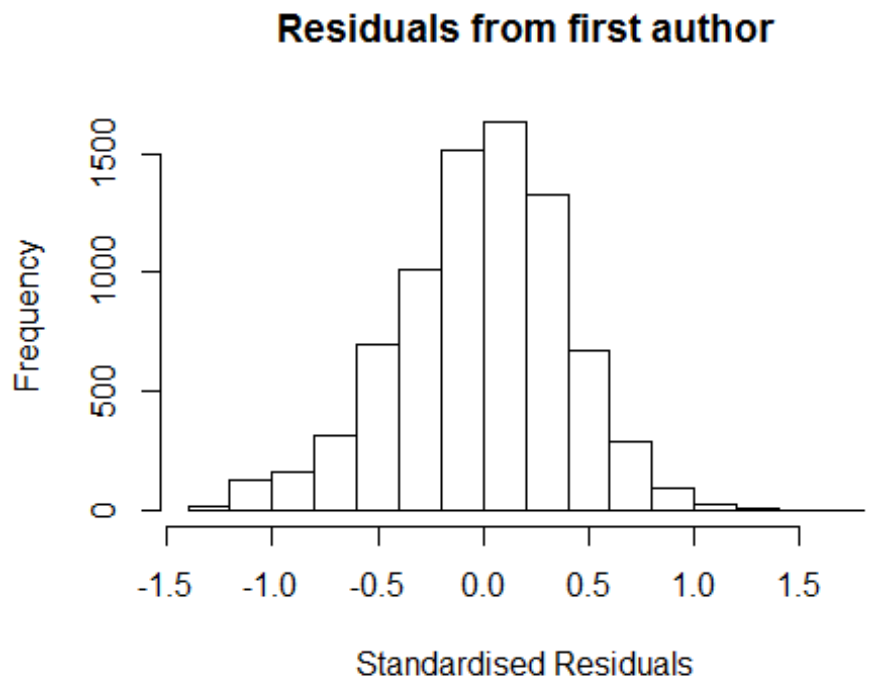

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
UniqueAuthors +
##     Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##     k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.4022 -0.2554  0.0133  0.2533  1.7601
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.15335    0.01921   60.04  < 2e-16 ***
## FirstAuthorFemale1  0.00976    0.01034    0.94  0.34566
## LastAuthorFemale1  -0.04376    0.01203   -3.64  0.00028 ***
## UniqueAuthors2      0.10272    0.01330    7.72  1.3e-14 ***
## UniqueAuthors3      0.17094    0.01445   11.83  < 2e-16 ***
## UniqueAuthors4      0.21155    0.01573   13.45  < 2e-16 ***
## UniqueAuthors5      0.27245    0.01464   18.61  < 2e-16 ***
## Year1997           -0.02360    0.02543   -0.93  0.35329
## Year1998           -0.01288    0.02649   -0.49  0.62684
## Year1999           -0.07970    0.02631   -3.03  0.00246 **
```

```
## Year2000           -0.11494    0.02660   -4.32  1.6e-05 ***
## Year2001           -0.09847    0.02721   -3.62  0.00030 ***
## Year2002           -0.16017    0.02790   -5.74  9.8e-09 ***
## Year2003           -0.13498    0.02417   -5.58  2.4e-08 ***
## Year2004           -0.13407    0.02497   -5.37  8.2e-08 ***
## Year2005           -0.12697    0.02472   -5.14  2.9e-07 ***
## Year2006           -0.15836    0.02421   -6.54  6.5e-11 ***
## Year2007           -0.12779    0.02380   -5.37  8.1e-08 ***
## Year2008           -0.12371    0.02486   -4.98  6.6e-07 ***
## Year2009           -0.15233    0.02485   -6.13  9.2e-10 ***
## Year2010           -0.18617    0.02542   -7.32  2.6e-13 ***
## Year2011           -0.17099    0.02450   -6.98  3.2e-12 ***
## Year2012           -0.16961    0.02586   -6.56  5.8e-11 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.375
## Multiple R-squared:  0.0651, Adjusted R-squared:  0.0625
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  observation 2709 is an outlier with |weight| = 0 ( < 1.3e-05);
##  686 weights are ~= 1. The remaining 7206 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0037  0.8690  0.9500  0.8990  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.27e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.059  1           1.029
## LastAuthorFemale  1.047  1           1.023
## Year              1.027 16           1.001
```



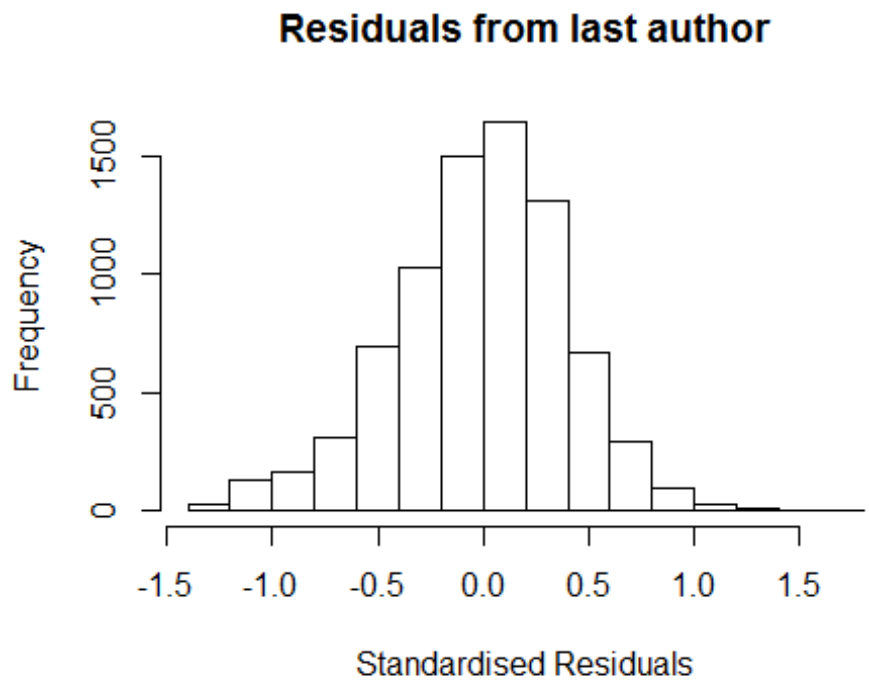

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.2833 -0.2596  0.0155  0.2585  1.6719
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)          1.25327    0.01749   71.64  < 2e-16 ***
## FirstAuthorFemale1  0.02138    0.01054    2.03  0.04242 *
## LastAuthorFemale1  -0.04424    0.01226   -3.61  0.00031 ***
## Year1997           -0.02529    0.02556   -0.99  0.32253
## Year1998            0.00869    0.02631    0.33  0.74111
## Year1999           -0.06702    0.02640   -2.54  0.01115 *
## Year2000           -0.10111    0.02708   -3.73  0.00019 ***
## Year2001           -0.06424    0.02756   -2.33  0.01976 *
## Year2002           -0.12463    0.02809   -4.44  9.2e-06 ***
## Year2003           -0.09366    0.02409   -3.89  0.00010 ***
## Year2004           -0.09409    0.02504   -3.76  0.00017 ***
## Year2005           -0.10230    0.02514   -4.07  4.7e-05 ***
```

```
## Year2006           -0.11647    0.02461   -4.73  2.3e-06 ***
## Year2007           -0.08565    0.02417   -3.54  0.00040 ***
## Year2008           -0.08190    0.02510   -3.26  0.00111 **
## Year2009           -0.10946    0.02517   -4.35  1.4e-05 ***
## Year2010           -0.13704    0.02540   -5.39  7.1e-08 ***
## Year2011           -0.12276    0.02494   -4.92  8.7e-07 ***
## Year2012           -0.11260    0.02606   -4.32  1.6e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.382
## Multiple R-squared:  0.0128, Adjusted R-squared:  0.0106
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  612 weights are ~= 1. The remaining 7281 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0165  0.8690  0.9500  0.8980  0.9860  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.27e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##                   GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.02  1           1.010
## Year              1.02 16           1.001
```




```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.2700 -0.2588  0.0151  0.2596  1.6413
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)          1.24689    0.01735   71.88  < 2e-16 ***
## FirstAuthorFemale1   0.01319    0.01039    1.27  0.20428
## Year1997            -0.02365    0.02552   -0.93  0.35426
## Year1998             0.00994    0.02633    0.38  0.70591
## Year1999            -0.06572    0.02634   -2.50  0.01261 *
## Year2000            -0.09917    0.02709   -3.66  0.00025 ***
## Year2001            -0.06296    0.02758   -2.28  0.02246 *
## Year2002            -0.12395    0.02808   -4.41  1.0e-05 ***
## Year2003            -0.09244    0.02413   -3.83  0.00013 ***
## Year2004            -0.09346    0.02509   -3.72  0.00020 ***
## Year2005            -0.10139    0.02516   -4.03  5.7e-05 ***
## Year2006            -0.11488    0.02460   -4.67  3.1e-06 ***
```

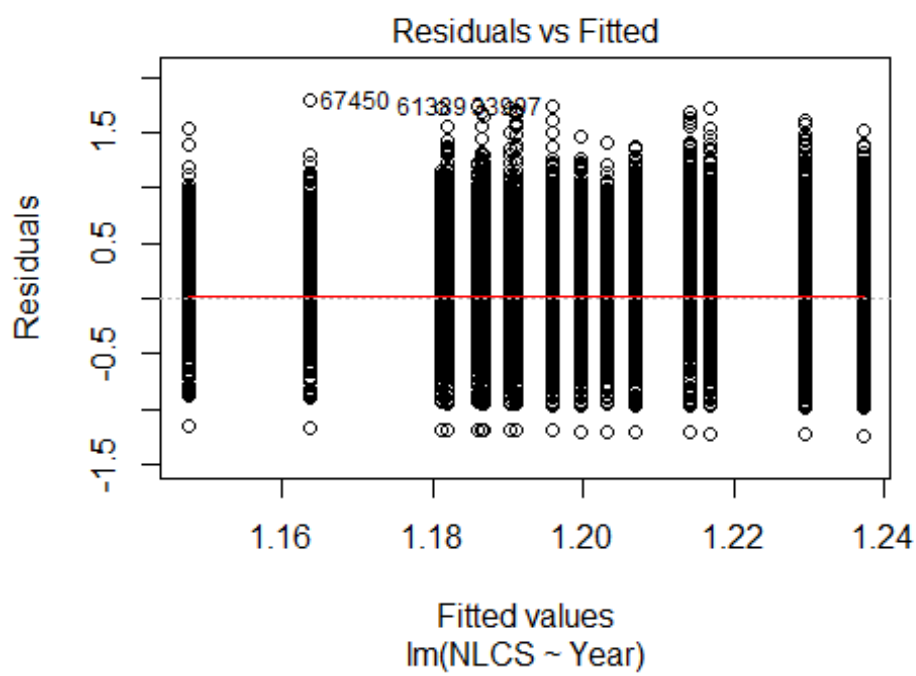
```
## Year2007           -0.08501    0.02421   -3.51  0.00045 ***
## Year2008           -0.08217    0.02508   -3.28  0.00106 **
## Year2009           -0.10850    0.02519   -4.31  1.7e-05 ***
## Year2010           -0.13672    0.02541   -5.38  7.6e-08 ***
## Year2011           -0.12115    0.02491   -4.86  1.2e-06 ***
## Year2012           -0.11323    0.02598   -4.36  1.3e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.383
## Multiple R-squared:  0.0111, Adjusted R-squared:  0.00899
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  610 weights are ~= 1. The remaining 7283 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0268  0.8700  0.9500  0.8980  0.9860  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.27e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##                   GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.009  1           1.004
## Year             1.009 16           1.000
```



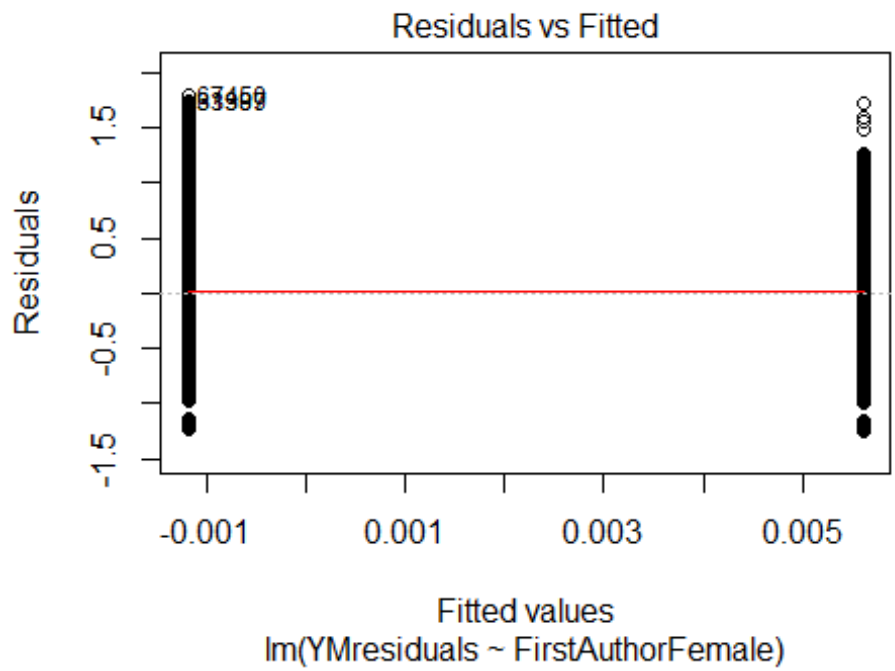

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.2648 -0.2593  0.0144  0.2589  1.6834
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.25586    0.01742   72.08  < 2e-16 ***
## LastAuthorFemale1 -0.03911    0.01202   -3.25  0.00114 **
## Year1997          -0.02477    0.02557   -0.97  0.33279
## Year1998           0.00893    0.02632    0.34  0.73454
## Year1999          -0.06618    0.02635   -2.51  0.01205 *
## Year2000          -0.09978    0.02707   -3.69  0.00023 ***
## Year2001          -0.06373    0.02755   -2.31  0.02071 *
## Year2002          -0.12207    0.02805   -4.35  1.4e-05 ***
## Year2003          -0.09238    0.02407   -3.84  0.00012 ***
## Year2004          -0.09183    0.02503   -3.67  0.00025 ***
## Year2005          -0.10013    0.02508   -3.99  6.6e-05 ***
## Year2006          -0.11463    0.02459   -4.66  3.2e-06 ***
```

```
## Year2007          -0.08382    0.02414   -3.47  0.00052 ***
## Year2008          -0.07920    0.02505   -3.16  0.00157 **
## Year2009          -0.10730    0.02512   -4.27  2.0e-05 ***
## Year2010          -0.13539    0.02537   -5.34  9.7e-08 ***
## Year2011          -0.12054    0.02491   -4.84  1.3e-06 ***
## Year2012          -0.10964    0.02595   -4.22  2.4e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.382
## Multiple R-squared:  0.0123, Adjusted R-squared:  0.0102
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  615 weights are ~= 1. The remaining 7278 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0136  0.8690  0.9500  0.8980  0.9860  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.27e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Sample size for the above analysis:  7893"
## [1] ""
## [1] ""
## [1] "####################################"
## [1] "Analysis of AJSC 1912"
## [1] "####################################"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 3407 3462 3336 3171 3150 3463 3092 3302 3071 2832 3177 3292 3161 3201 3162
## 2011 2012
## 3094 3078
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 1873 1854 1793 1753 1772 1773 1691 1784 1631 1570 1833 1935 1736 1806 1780
## 2011 2012
```

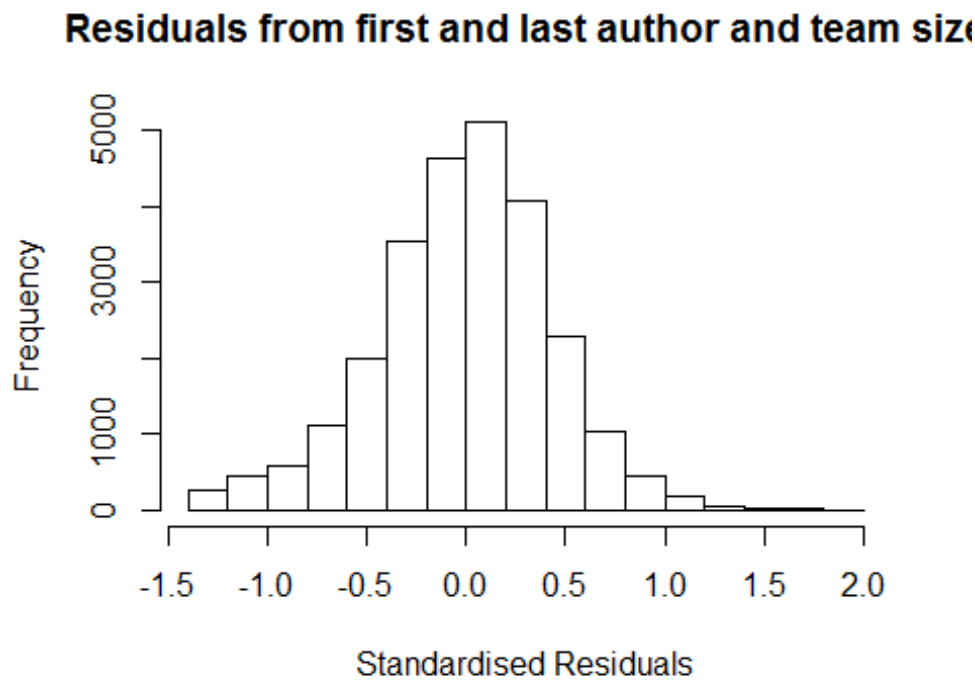
```
## 1775 1784
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 1613 1610 1552 1526 1537 1534 1434 1490 1386 1326 1560 1670 1455 1550 1522
## 2011 2012
## 1534 1516
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
##  Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 190, df = 16, p-value <2e-16
```




```
##
##  Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 86, df = 1, p-value <2e-16
```

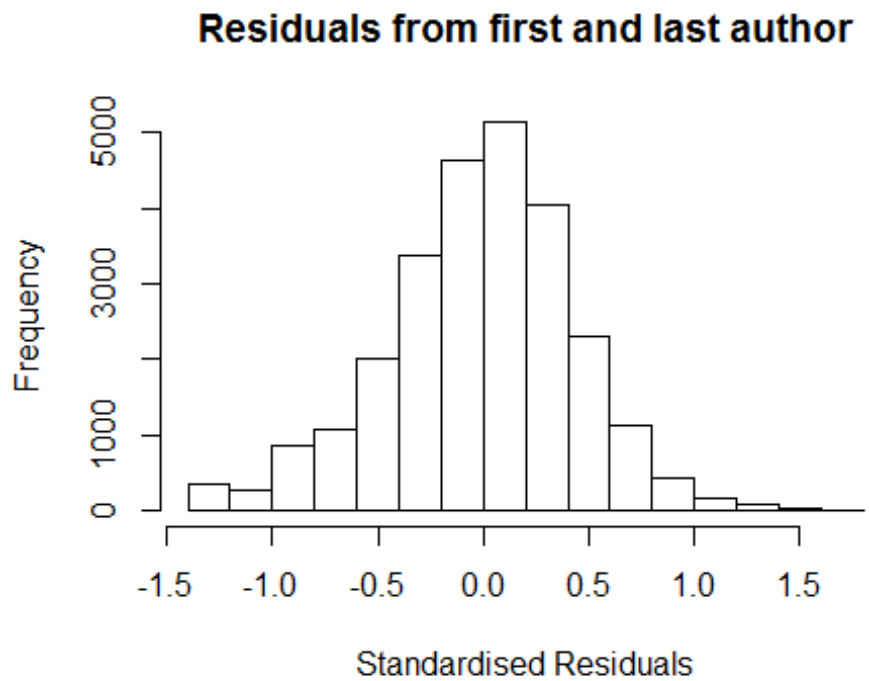



```
## [1] "Female first author team size 2018 geometric mean: 3.75502975067264"
## [1] "Male first author team size 2018 geometric mean: 2.92624460575213"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 110000, p-value = 9e-07
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 3.5531746400606"
## [1] "Male last author team size 2018 geometric mean: 3.01433481078231"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 88000, p-value = 0.002
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.034  1           1.017
## LastAuthorFemale  1.023  1           1.012
## UniqueAuthors     1.050  4           1.006
## Year              1.042 16           1.001
```



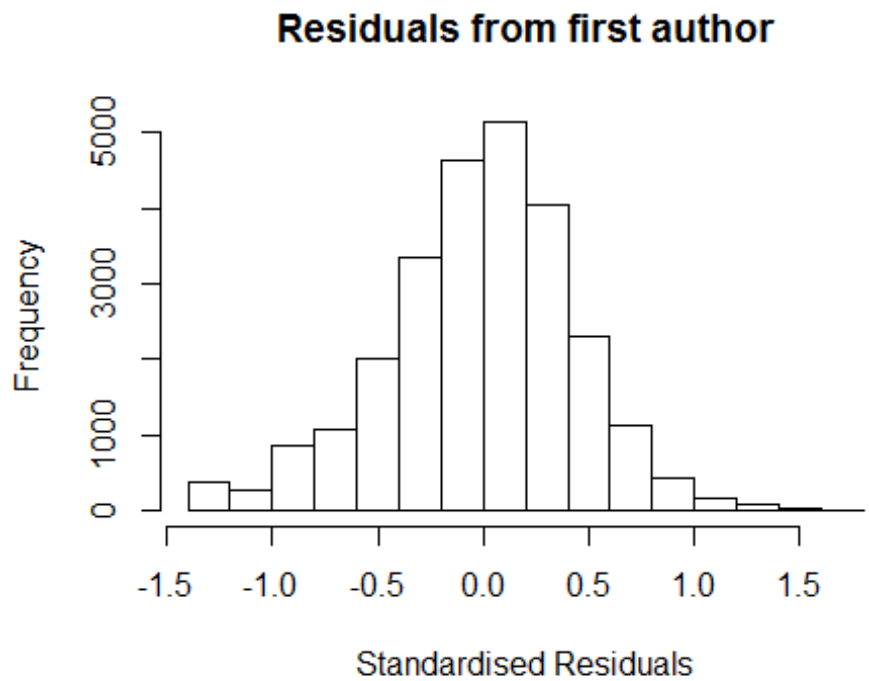

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
UniqueAuthors +
##     Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##     k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.3546 -0.2768  0.0128  0.2716  1.8553
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.12434    0.01321   85.10  < 2e-16 ***
## FirstAuthorFemale1 -0.00845    0.00671   -1.26  0.20747
## LastAuthorFemale1  -0.01868    0.00777   -2.41  0.01617 *
## UniqueAuthors2      0.15919    0.00886   17.97  < 2e-16 ***
## UniqueAuthors3      0.18241    0.00931   19.60  < 2e-16 ***
## UniqueAuthors4      0.19001    0.01032   18.42  < 2e-16 ***
## UniqueAuthors5      0.23021    0.00955   24.12  < 2e-16 ***
## Year1997           -0.00567    0.01646   -0.34  0.73045
## Year1998           -0.03414    0.01676   -2.04  0.04168 *
## Year1999           -0.03114    0.01669   -1.87  0.06208 .
```

```
## Year2000           -0.07495    0.01658   -4.52  6.2e-06 ***
## Year2001           -0.05340    0.01678   -3.18  0.00146 **
## Year2002           -0.06331    0.01690   -3.75  0.00018 ***
## Year2003           -0.06305    0.01606   -3.93  8.6e-05 ***
## Year2004           -0.07767    0.01579   -4.92  8.8e-07 ***
## Year2005           -0.06575    0.01633   -4.03  5.7e-05 ***
## Year2006           -0.08255    0.01567   -5.27  1.4e-07 ***
## Year2007           -0.08249    0.01583   -5.21  1.9e-07 ***
## Year2008           -0.06707    0.01573   -4.26  2.0e-05 ***
## Year2009           -0.08465    0.01568   -5.40  6.8e-08 ***
## Year2010           -0.08571    0.01596   -5.37  8.0e-08 ***
## Year2011           -0.10788    0.01607   -6.71  1.9e-11 ***
## Year2012           -0.11436    0.01608   -7.11  1.2e-12 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.406
## Multiple R-squared:  0.0345, Adjusted R-squared:  0.0337
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
##  2192 weights are ~= 1. The remaining 23623 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0024  0.8640  0.9500  0.8940  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          3.87e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.021  1            1.01
## LastAuthorFemale  1.020  1            1.01
## Year              1.010 16            1.00
```



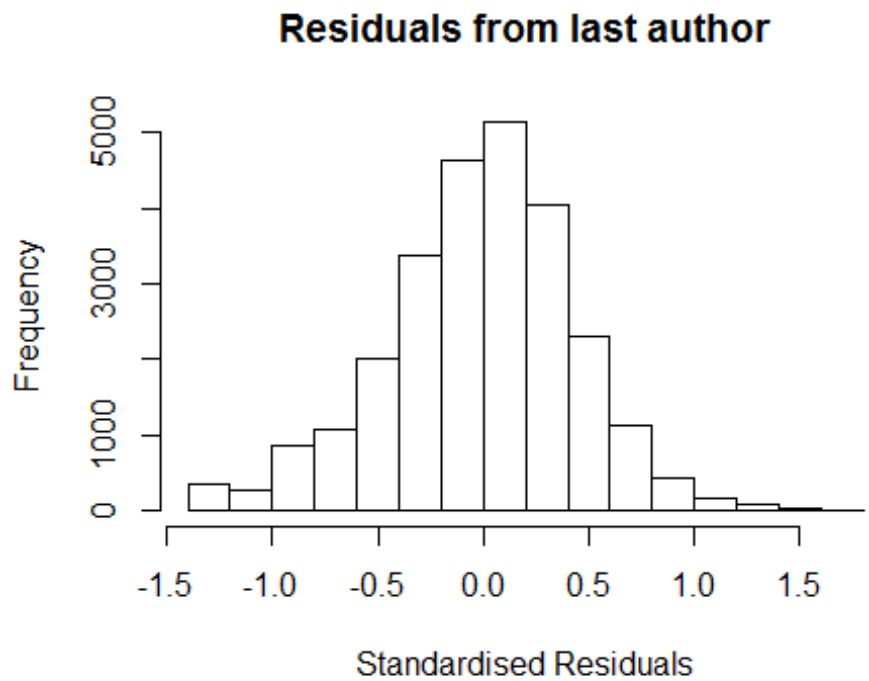

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.2553 -0.2734  0.0145  0.2745  1.7795
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)          1.24652    0.01176  106.00  < 2e-16 ***
## FirstAuthorFemale1  0.01010    0.00672    1.50  0.13299
## LastAuthorFemale1  -0.01137    0.00784   -1.45  0.14680
## Year1997           -0.00134    0.01646   -0.08  0.93536
## Year1998           -0.02549    0.01681   -1.52  0.12940
## Year1999           -0.02015    0.01677   -1.20  0.22941
## Year2000           -0.05520    0.01656   -3.33  0.00086 ***
## Year2001           -0.02785    0.01682   -1.66  0.09786 .
## Year2002           -0.03865    0.01686   -2.29  0.02186 *
## Year2003           -0.03890    0.01603   -2.43  0.01524 *
## Year2004           -0.04908    0.01573   -3.12  0.00181 **
## Year2005           -0.03897    0.01633   -2.39  0.01702 *
```

```
## Year2006           -0.05194    0.01576   -3.30  0.00098 ***
## Year2007           -0.05168    0.01590   -3.25  0.00115 **
## Year2008           -0.03382    0.01580   -2.14  0.03227 *
## Year2009           -0.05327    0.01575   -3.38  0.00072 ***
## Year2010           -0.05003    0.01606   -3.12  0.00184 **
## Year2011           -0.07105    0.01608   -4.42  1.0e-05 ***
## Year2012           -0.08186    0.01611   -5.08  3.8e-07 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.409
## Multiple R-squared:  0.00257,    Adjusted R-squared:  0.00187
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  2135 weights are ~= 1. The remaining 23680 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0191  0.8620  0.9510  0.8930  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          3.87e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.006  1           1.003
## Year              1.006 16           1.000
```




```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.2538 -0.2738  0.0137  0.2751  1.7809
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)          1.24543    0.01172  106.23  < 2e-16 ***
## FirstAuthorFemale1  0.00833    0.00669    1.25  0.21308
## Year1997           -0.00136    0.01646   -0.08  0.93426
## Year1998           -0.02548    0.01681   -1.52  0.12950
## Year1999           -0.02023    0.01676   -1.21  0.22748
## Year2000           -0.05513    0.01656   -3.33  0.00087 ***
## Year2001           -0.02780    0.01682   -1.65  0.09849 .
## Year2002           -0.03860    0.01686   -2.29  0.02206 *
## Year2003           -0.03894    0.01603   -2.43  0.01515 *
## Year2004           -0.04921    0.01573   -3.13  0.00176 **
## Year2005           -0.03903    0.01633   -2.39  0.01683 *
## Year2006           -0.05186    0.01576   -3.29  0.00100 **
```

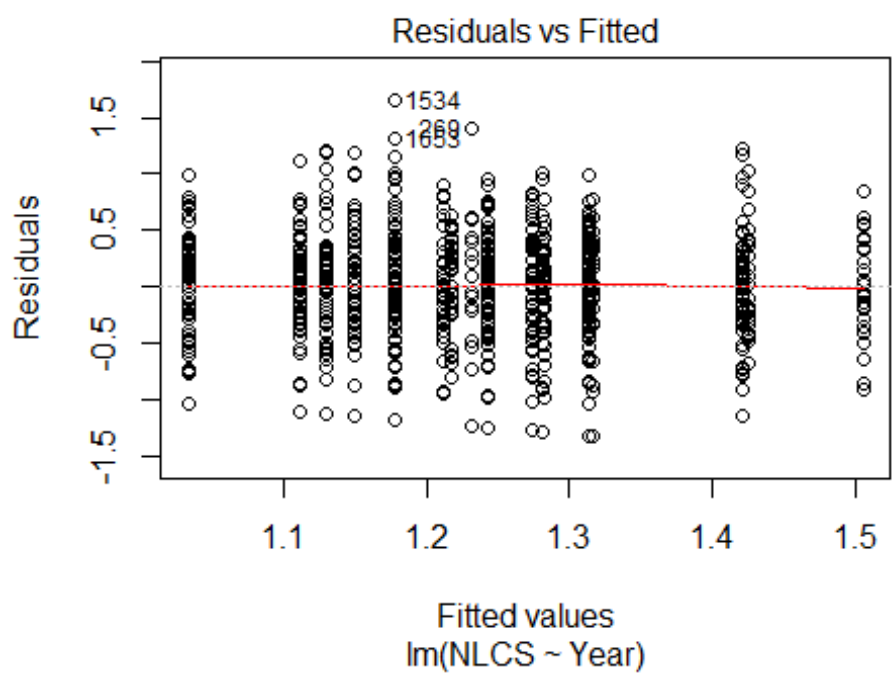
```
## Year2007           -0.05171    0.01589   -3.25  0.00114 **
## Year2008           -0.03398    0.01580   -2.15  0.03149 *
## Year2009           -0.05350    0.01575   -3.40  0.00068 ***
## Year2010           -0.05038    0.01605   -3.14  0.00170 **
## Year2011           -0.07134    0.01608   -4.44  9.1e-06 ***
## Year2012           -0.08231    0.01610   -5.11  3.2e-07 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.409
## Multiple R-squared:  0.0025, Adjusted R-squared:  0.00184
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  2143 weights are ~= 1. The remaining 23672 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0188  0.8620  0.9510  0.8930  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          3.87e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##                   GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.005  1           1.003
## Year             1.005 16           1.000
```



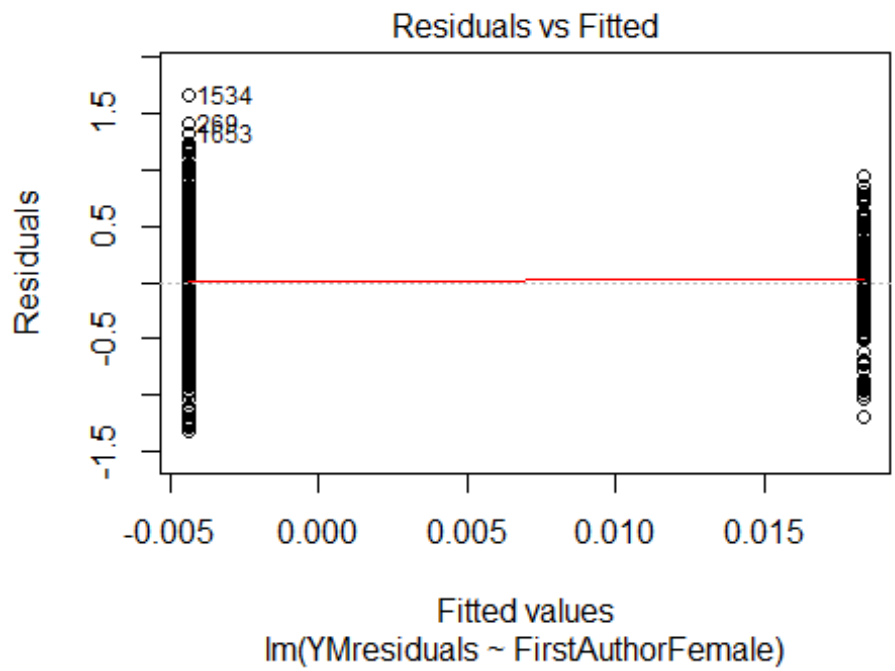

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.2476 -0.2737  0.0145  0.2745  1.7776
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.24762    0.01172  106.41  < 2e-16 ***
## LastAuthorFemale1 -0.00932    0.00779   -1.20  0.23135
## Year1997          -0.00117    0.01647   -0.07  0.94356
## Year1998          -0.02531    0.01682   -1.50  0.13235
## Year1999          -0.01982    0.01677   -1.18  0.23716
## Year2000          -0.05489    0.01656   -3.31  0.00092 ***
## Year2001          -0.02768    0.01683   -1.64  0.10000
## Year2002          -0.03811    0.01686   -2.26  0.02383 *
## Year2003          -0.03854    0.01603   -2.40  0.01623 *
## Year2004          -0.04852    0.01573   -3.08  0.00204 **
## Year2005          -0.03851    0.01633   -2.36  0.01835 *
## Year2006          -0.05130    0.01576   -3.25  0.00114 **
```

```
## Year2007          -0.05111    0.01590   -3.21  0.00131 **
## Year2008          -0.03311    0.01579   -2.10  0.03603 *
## Year2009          -0.05264    0.01575   -3.34  0.00083 ***
## Year2010          -0.04916    0.01605   -3.06  0.00219 **
## Year2011          -0.07020    0.01607   -4.37  1.3e-05 ***
## Year2012          -0.08104    0.01610   -5.03  4.9e-07 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.409
## Multiple R-squared:  0.0025, Adjusted R-squared:  0.00184
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  2124 weights are ~= 1. The remaining 23691 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0196  0.8620  0.9510  0.8930  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          3.87e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Sample size for the above analysis:  25815"
## [1] ""
## [1] ""
## [1] "####################################"
## [1] "Analysis of AJSC 1913"
## [1] "####################################"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##   58   57   41   42   36   50   56   62   62   58   82   92  105   86  100
## 2011 2012
##  133  171
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##   32   38   26   28   17   27   38   48   53   44   69   62   79   58   75
## 2011 2012
```

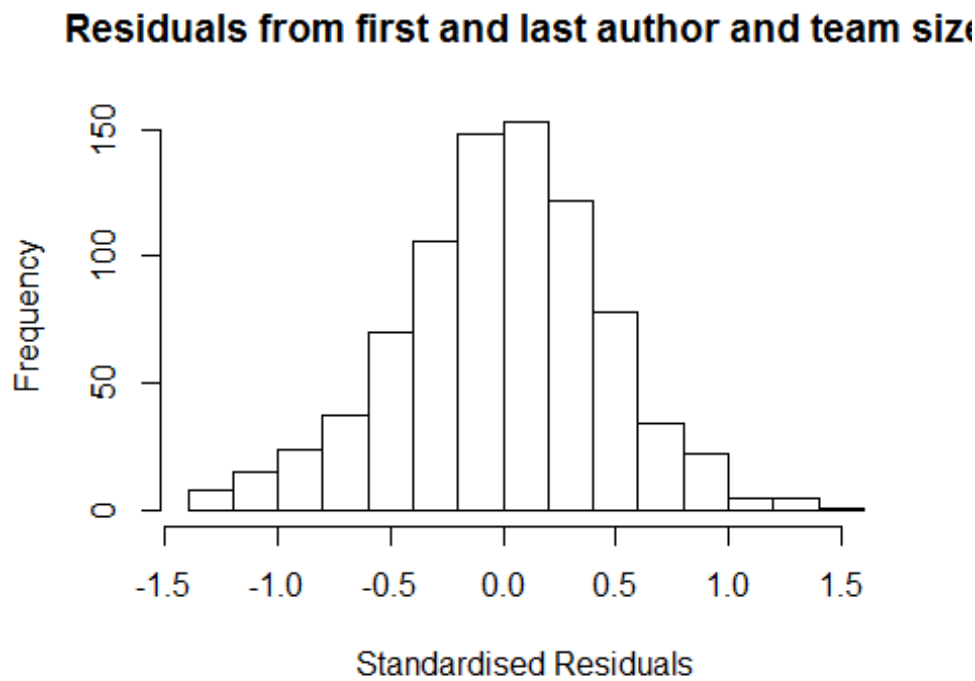
```
##  102  130
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##   30   36   26   25   15   21   33   46   48   37   63   58   69   49   67
## 2011 2012
##   92  113
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
##  Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 17, df = 16, p-value = 0.4
```




```
##
##  Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 7.8, df = 1, p-value = 0.005
```

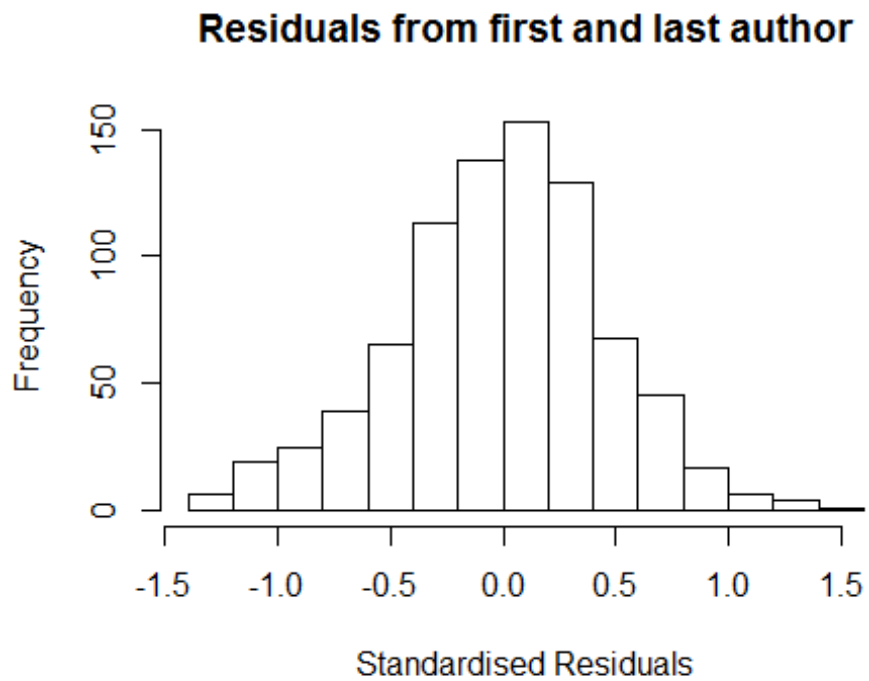



```
## [1] "Female first author team size 2018 geometric mean: 4.29075960707389"
## [1] "Male first author team size 2018 geometric mean: 2.74455033772913"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 1400, p-value = 8e-04
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 3.27343563610951"
## [1] "Male last author team size 2018 geometric mean: 3.10938992920549"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 760, p-value = 0.8
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.197  1           1.094
## LastAuthorFemale  1.234  1           1.111
## UniqueAuthors     1.430  4           1.046
## Year              1.591 16           1.015
```



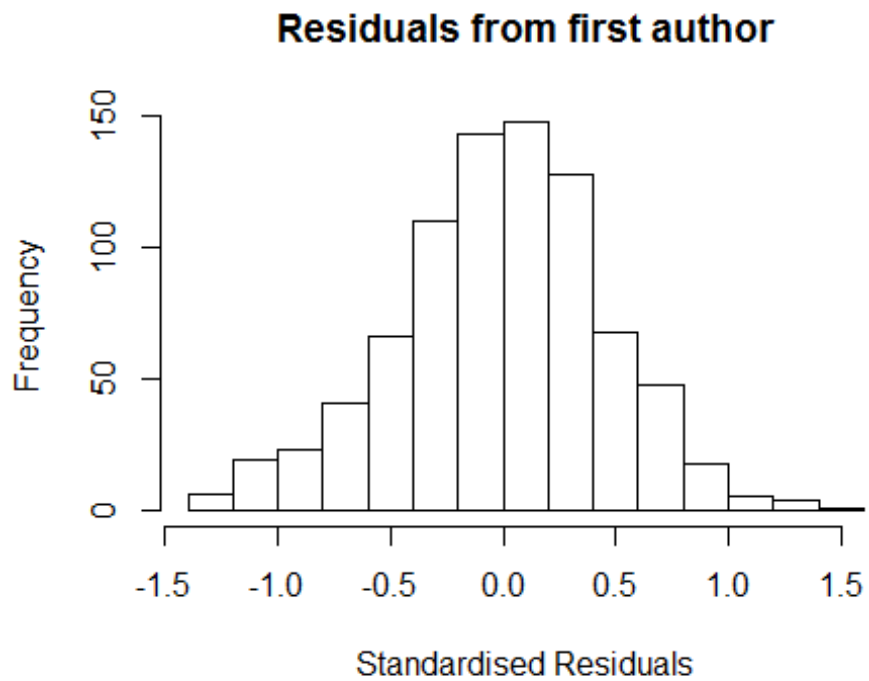

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
UniqueAuthors +
##     Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##     k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.3025 -0.2811  0.0143  0.2919  1.4999
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.26707    0.09215   13.75   <2e-16 ***
## FirstAuthorFemale1  0.03757    0.04053    0.93   0.3543
## LastAuthorFemale1  -0.02526    0.04151   -0.61   0.5431
## UniqueAuthors2      0.03952    0.04839    0.82   0.4144
## UniqueAuthors3      0.02067    0.05545    0.37   0.7095
## UniqueAuthors4      0.01498    0.05646    0.27   0.7908
## UniqueAuthors5      0.16400    0.05729    2.86   0.0043 **
## Year1997           -0.07157    0.10865   -0.66   0.5102
## Year1998            0.21948    0.12414    1.77   0.0774 .
## Year1999            0.09049    0.12074    0.75   0.4538
```

```
## Year2000           -0.13900    0.17789   -0.78   0.4348
## Year2001           -0.00230    0.12538   -0.02   0.9854
## Year2002            0.01937    0.11081    0.17   0.8613
## Year2003           -0.01311    0.12063   -0.11   0.9135
## Year2004            0.13687    0.11431    1.20   0.2315
## Year2005           -0.10902    0.11253   -0.97   0.3329
## Year2006           -0.19357    0.10049   -1.93   0.0544 .
## Year2007           -0.20458    0.10717   -1.91   0.0566 .
## Year2008           -0.02979    0.10441   -0.29   0.7755
## Year2009           -0.15415    0.11412   -1.35   0.1772
## Year2010           -0.26192    0.10274   -2.55   0.0110 *
## Year2011           -0.00408    0.09331   -0.04   0.9651
## Year2012           -0.17898    0.09386   -1.91   0.0569 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.44
## Multiple R-squared:  0.0767, Adjusted R-squared:  0.0515
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
##  69 weights are ~= 1. The remaining 759 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.220   0.866   0.953   0.899   0.985   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.21e-04          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.217  1           1.103
## LastAuthorFemale  1.233  1           1.110
## Year              1.122 16           1.004
```



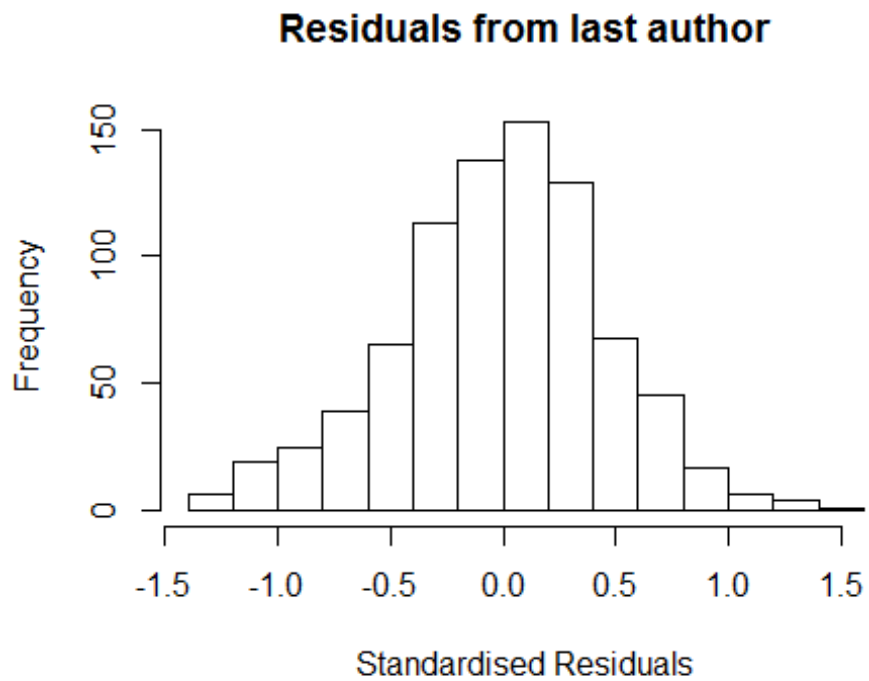

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.3198 -0.2879  0.0117  0.2917  1.4741
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.29493    0.08505   15.23   <2e-16 ***
## FirstAuthorFemale1  0.03653    0.04163    0.88    0.381
## LastAuthorFemale1  -0.02559    0.04205   -0.61    0.543
## Year1997           -0.07873    0.10903   -0.72    0.470
## Year1998            0.22204    0.12433    1.79    0.075 .
## Year1999            0.08647    0.11888    0.73    0.467
## Year2000           -0.14103    0.18291   -0.77    0.441
## Year2001            0.00890    0.12667    0.07    0.944
## Year2002            0.02482    0.11022    0.23    0.822
## Year2003           -0.02054    0.12132   -0.17    0.866
## Year2004            0.14974    0.11430    1.31    0.191
## Year2005           -0.10947    0.11283   -0.97    0.332
```

```
## Year2006           -0.17811    0.10095   -1.76    0.078 .
## Year2007           -0.19279    0.10752   -1.79    0.073 .
## Year2008            0.00062    0.10509    0.01    0.995
## Year2009           -0.12322    0.11420   -1.08    0.281
## Year2010           -0.25393    0.10393   -2.44    0.015 *
## Year2011            0.01560    0.09446    0.17    0.869
## Year2012           -0.14658    0.09432   -1.55    0.121
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.44
## Multiple R-squared:  0.0648, Adjusted R-squared:  0.044
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
##  66 weights are ~= 1. The remaining 762 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.240   0.865   0.954   0.898   0.986   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.21e-04          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##                   GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.05  1           1.025
## Year              1.05 16           1.002
```




```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.31629 -0.28531  0.00656  0.29355  1.47079
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.29238    0.08440   15.31   <2e-16 ***
## FirstAuthorFemale1  0.02867    0.03881    0.74    0.460
## Year1997           -0.07803    0.10849   -0.72    0.472
## Year1998            0.22132    0.12419    1.78    0.075 .
## Year1999            0.08737    0.11873    0.74    0.462
## Year2000           -0.13518    0.18292   -0.74    0.460
## Year2001            0.00905    0.12680    0.07    0.943
## Year2002            0.02390    0.11005    0.22    0.828
## Year2003           -0.02156    0.12119   -0.18    0.859
## Year2004            0.14997    0.11383    1.32    0.188
## Year2005           -0.11045    0.11312   -0.98    0.329
## Year2006           -0.18053    0.10041   -1.80    0.073 .
```

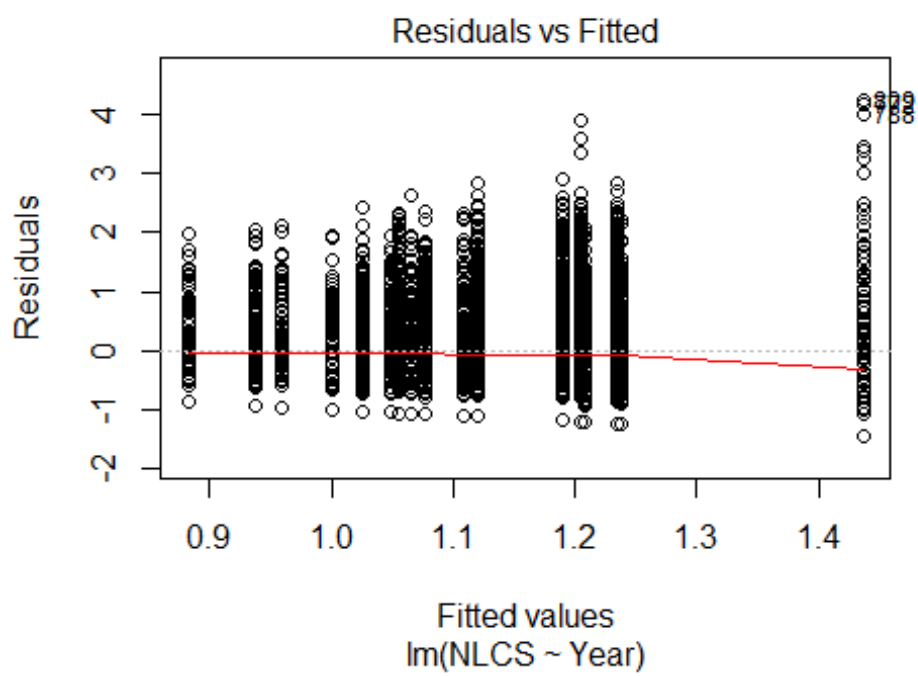
```
## Year2007           -0.19253    0.10715   -1.80    0.073 .
## Year2008           -0.00105    0.10491   -0.01    0.992
## Year2009           -0.12362    0.11411   -1.08    0.279
## Year2010           -0.25451    0.10366   -2.46    0.014 *
## Year2011            0.01409    0.09418    0.15    0.881
## Year2012           -0.14818    0.09394   -1.58    0.115
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.443
## Multiple R-squared:  0.0642, Adjusted R-squared:  0.0446
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
##  67 weights are ~= 1. The remaining 761 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.248   0.865   0.954   0.899   0.986   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.21e-04          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##                   GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.065  1           1.032
## Year             1.065 16           1.002
```



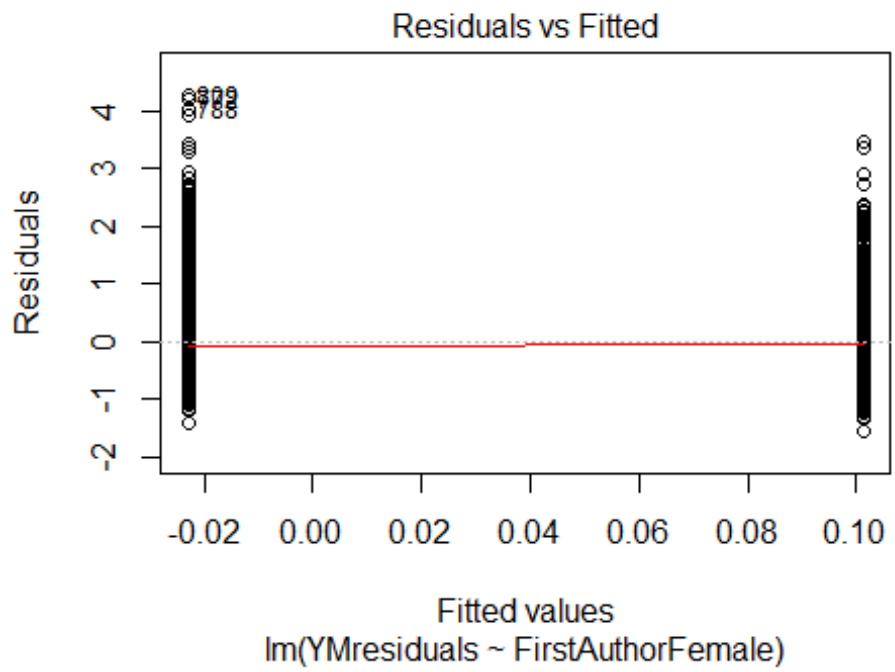

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.32786 -0.28670  0.00995  0.29351  1.46477
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.29728    0.08473   15.31   <2e-16 ***
## LastAuthorFemale1 -0.01392    0.03916   -0.36    0.722
## Year1997          -0.07570    0.10934   -0.69    0.489
## Year1998           0.22091    0.12433    1.78    0.076 .
## Year1999           0.08909    0.11812    0.75    0.451
## Year2000          -0.13405    0.18489   -0.73    0.469
## Year2001           0.01444    0.12503    0.12    0.908
## Year2002           0.03058    0.10947    0.28    0.780
## Year2003          -0.02068    0.12110   -0.17    0.864
## Year2004           0.15629    0.11395    1.37    0.171
## Year2005          -0.10599    0.11292   -0.94    0.348
## Year2006          -0.17630    0.10063   -1.75    0.080 .
```

```
## Year2007          -0.19140    0.10728   -1.78    0.075 .
## Year2008           0.00325    0.10482    0.03    0.975
## Year2009          -0.12343    0.11423   -1.08    0.280
## Year2010          -0.25194    0.10358   -2.43    0.015 *
## Year2011           0.01919    0.09379    0.20    0.838
## Year2012          -0.14306    0.09397   -1.52    0.128
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.441
## Multiple R-squared:  0.0641, Adjusted R-squared:  0.0444
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
##  65 weights are ~= 1. The remaining 763 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.247   0.864   0.954   0.899   0.986   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.21e-04          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Sample size for the above analysis:  828"
## [1] ""
## [1] ""
## [1] "####################################"
## [1] "Analysis of AJSC 2000"
## [1] "####################################"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  212  146  173  144  193  249  239  274  272  310  420  458  548  685  649
## 2011 2012
##  669  641
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  183  121  131  106  152  185  188  226  215  252  342  364  426  543  512
## 2011 2012
```

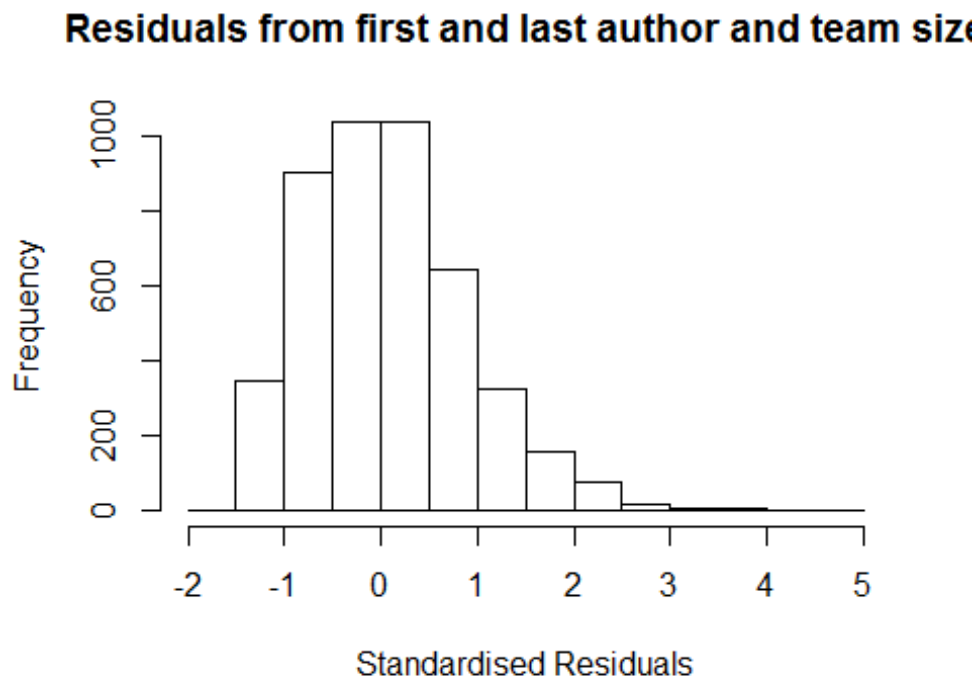
```
##  515  503
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  168  117  123  104  146  172  174  206  197  222  322  339  379  501  455
## 2011 2012
##  472  442
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
##  Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 220, df = 16, p-value <2e-16
```




```
##
##  Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 1.8, df = 1, p-value = 0.2
```



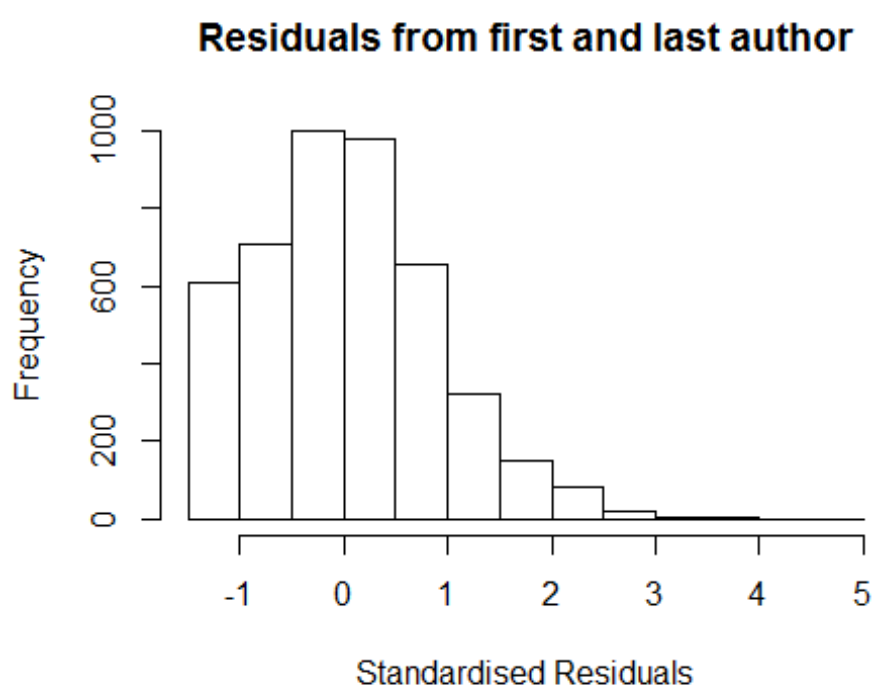

```
## [1] "Female first author team size 2018 geometric mean: 1.56629260867482"
## [1] "Male first author team size 2018 geometric mean: 1.49897708219105"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 15000, p-value = 0.5
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 1.56084598808659"
## [1] "Male last author team size 2018 geometric mean: 1.50039757748293"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 15000, p-value = 0.5
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.683  1           1.297
## LastAuthorFemale  1.690  1           1.300
## UniqueAuthors     1.067  4           1.008
## Year              1.071 16           1.002
```




```
## [1] "List of  28 outliers with residuals above  2.5"
##         ScopusId  NLCS Year OneField Fields residuals
## 717   0042846710 4.822 2000     2000      1     3.826
## 718   0347037232 4.888 2000     2000      1     3.623
## 724  84891802579 3.945 2000     2000      1     2.949
## 772   0001790323 5.594 2000     2000      1     4.598
## 777   0034374265 4.445 2000     2000      1     3.099
## 780   0034403543 3.822 2000     2000      1     2.686
## 782   0039024028 4.888 2000     2000      1     3.542
## 788  16244399574 5.423 2000     2000      1     4.216
## 789  33644528943 4.682 2000     2000      1     3.686
## 809   6344283665 5.673 2000     2000      1     4.677
## 1244  0001378820 3.701 2002     1800      4     2.575
## 3912 40249110107 3.391 2008     2000      1     2.518
## 4419 77953083241 3.587 2009     2000      1     2.653
## 4420 77954605495 3.743 2009     2000      1     2.599
## 4636 85017105478 3.948 2009     2000      1     2.664
## 4647 85017167724 3.740 2009     2000      1     2.596
## 5428 85017138638 4.056 2010     2000      1     2.849
## 5443 74349104010 3.947 2010     2000      1     2.950
## 5515 85017125001 3.600 2010     2000      1     2.603
## 5806 80053983353 3.681 2011     2000      1     2.500
## 5978 79960792628 3.790 2011     2000      1     2.609
## 6299 79958135412 4.109 2011     2000      1     3.139
## 6702 84868030545 3.603 2012     2000      1     2.646
## 6703 84868030921 3.692 2012     2000      1     2.558
## 6749 84865191893 4.782 2012     2000      1     3.428
```

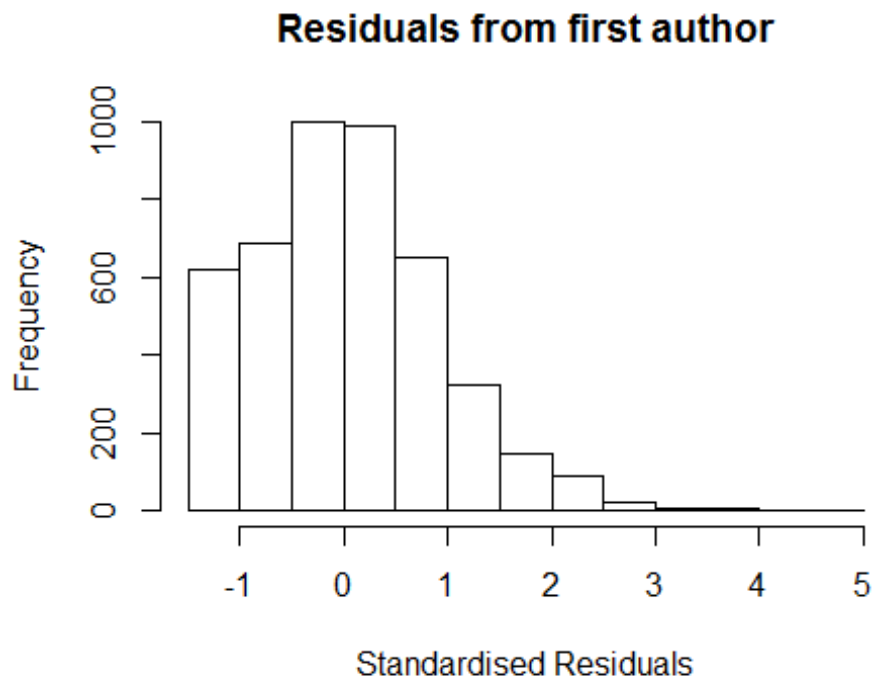
```
## 6839 84861658535 5.103 2012     2000      1     3.935
## 7066 84857150817 4.544 2012     2000      1     3.376
## 7068 84857174901 3.678 2012     2000      1     2.510
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
UniqueAuthors +
##     Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##     k.max = 1000))
##  \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.69539 -0.56652 -0.00447  0.54483  4.67651
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         0.77353    0.06567   11.78  < 2e-16 ***
## FirstAuthorFemale1  0.05798    0.04232    1.37  0.17076
## LastAuthorFemale1   0.08118    0.04114    1.97  0.04852 *
## UniqueAuthors2      0.21055    0.02986    7.05  2.1e-12 ***
## UniqueAuthors3      0.33850    0.04452    7.60  3.5e-14 ***
## UniqueAuthors4      0.17692    0.10106    1.75  0.08007 .
## UniqueAuthors5      0.88813    0.15479    5.74  1.0e-08 ***
## Year1997            0.03372    0.09373    0.36  0.71900
## Year1998            0.00543    0.09327    0.06  0.95354
## Year1999            0.09943    0.09648    1.03  0.30278
## Year2000            0.22295    0.12373    1.80  0.07162 .
## Year2001            0.18791    0.08864    2.12  0.03406 *
## Year2002            0.17542    0.08714    2.01  0.04416 *
## Year2003            0.28355    0.08955    3.17  0.00155 **
## Year2004            0.14329    0.08313    1.72  0.08484 .
## Year2005            0.28885    0.07794    3.71  0.00021 ***
## Year2006            0.09319    0.07610    1.22  0.22080
## Year2007            0.14979    0.07635    1.96  0.04983 *
## Year2008            0.09978    0.07496    1.33  0.18325
## Year2009            0.16041    0.07437    2.16  0.03108 *
## Year2010            0.22307    0.07815    2.85  0.00433 **
## Year2011            0.19648    0.07585    2.59  0.00962 **
## Year2012            0.18375    0.07975    2.30  0.02126 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.833
## Multiple R-squared:  0.0398, Adjusted R-squared:  0.0351
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
##  4 observations c(548,561,581,4373)
##   are outliers with |weight| = 0 ( < 2.2e-05);
```

```
##  361 weights are ~= 1. The remaining 4174 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0015  0.8790  0.9500  0.9090  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          2.20e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                    psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.638  1           1.280
## LastAuthorFemale  1.633  1           1.278
## Year              1.023 16           1.001
```



```
## [1] "List of  33 outliers with residuals above  2.5"
##         ScopusId  NLCS Year OneField Fields residuals
## 716   0042373524 3.688 2000     2000      1     2.612
## 717   0042846710 4.822 2000     2000      1     3.746
## 718   0347037232 4.888 2000     2000      1     3.757
## 724  84891802579 3.945 2000     2000      1     2.869
## 772   0001790323 5.594 2000     2000      1     4.518
## 777   0034374265 4.445 2000     2000      1     3.207
## 780   0034403543 3.822 2000     2000      1     2.584
## 782   0039024028 4.888 2000     2000      1     3.650
## 783   0039449757 3.822 2000     2000      1     2.691
## 788  16244399574 5.423 2000     2000      1     4.347
## 789  33644528943 4.682 2000     2000      1     3.606
## 809   6344283665 5.673 2000     2000      1     4.597
## 1244  0001378820 3.701 2002     1800      4     2.679
## 4419 77953083241 3.587 2009     2000      1     2.567
## 4420 77954605495 3.743 2009     2000      1     2.723
## 4546 74949137297 3.531 2009     2000      1     2.511
## 4636 85017105478 3.948 2009     2000      1     2.766
## 4647 85017167724 3.740 2009     2000      1     2.720
## 5164 79957454147 3.716 2010     2000      1     2.560
## 5428 85017138638 4.056 2010     2000      1     2.955
## 5443 74349104010 3.947 2010     2000      1     2.846
## 5490 77956947655 3.655 2010     2000      1     2.554
## 5806 80053983353 3.681 2011     2000      1     2.604
## 5978 79960792628 3.790 2011     2000      1     2.713
## 6240 79957827461 3.663 2011     2000      1     2.586
## 6299 79958135412 4.109 2011     2000      1     3.032
## 6702 84868030545 3.603 2012     2000      1     2.534
## 6703 84868030921 3.692 2012     2000      1     2.623
## 6749 84865191893 4.782 2012     2000      1     3.657
## 6839 84861658535 5.103 2012     2000      1     4.034
## 7015 84856199392 3.603 2012     2000      1     2.534
## 7066 84857150817 4.544 2012     2000      1     3.475
## 7068 84857174901 3.678 2012     2000      1     2.609
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.3267 -0.5623 -0.0203  0.5722  4.5973
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         0.89595    0.06508   13.77  < 2e-16 ***
## FirstAuthorFemale1  0.05551    0.04211    1.32  0.18743
## LastAuthorFemale1   0.10714    0.04092    2.62  0.00886 **
```

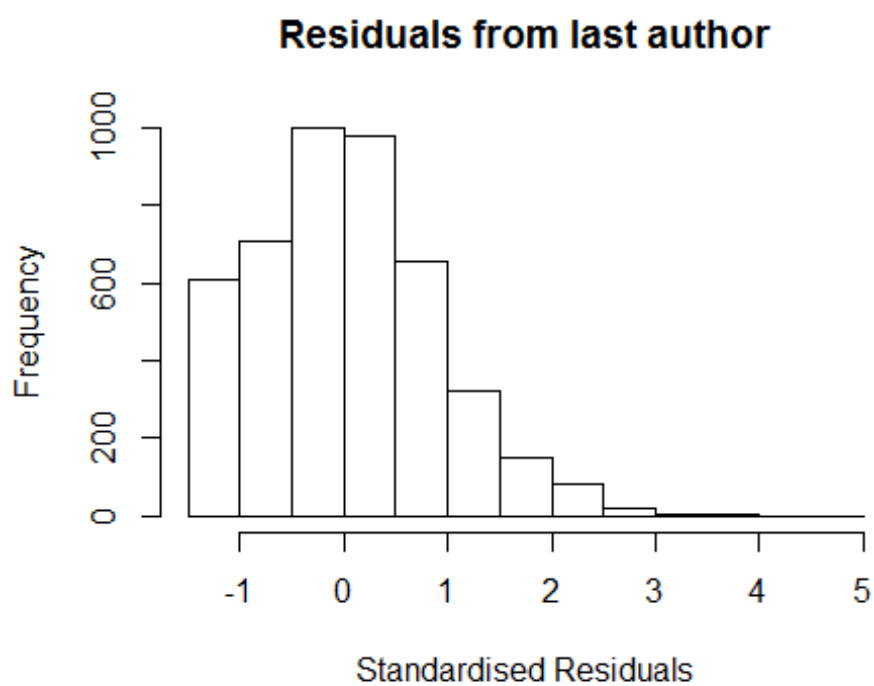
```
## Year1997            0.00237    0.09520    0.02  0.98016
## Year1998           -0.06112    0.09434   -0.65  0.51708
## Year1999            0.06546    0.09855    0.66  0.50660
## Year2000            0.17979    0.12541    1.43  0.15174
## Year2001            0.13972    0.09059    1.54  0.12308
## Year2002            0.12640    0.08788    1.44  0.15039
## Year2003            0.26294    0.09144    2.88  0.00405 **
## Year2004            0.10258    0.08412    1.22  0.22271
## Year2005            0.26811    0.07900    3.39  0.00069 ***
## Year2006            0.06787    0.07683    0.88  0.37705
## Year2007            0.12356    0.07691    1.61  0.10825
## Year2008            0.07681    0.07566    1.02  0.31007
## Year2009            0.12357    0.07493    1.65  0.09920 .
## Year2010            0.20497    0.07955    2.58  0.01001 *
## Year2011            0.18093    0.07648    2.37  0.01804 *
## Year2012            0.17311    0.08061    2.15  0.03182 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.844
## Multiple R-squared:  0.0125, Adjusted R-squared:  0.00858
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
##  4 observations c(548,561,581,4373)
##   are outliers with |weight| = 0 ( < 2.2e-05);
##  375 weights are ~= 1. The remaining 4160 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0095  0.8710  0.9520  0.9090  0.9860  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          2.20e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.016  1           1.008
## Year              1.016 16           1.000
```




```
## [1] "List of  33 outliers with residuals above  2.5"
##         ScopusId  NLCS Year OneField Fields residuals
## 716   0042373524 3.688 2000     2000      1     2.612
## 717   0042846710 4.822 2000     2000      1     3.746
## 718   0347037232 4.888 2000     2000      1     3.757
## 724  84891802579 3.945 2000     2000      1     2.869
## 772   0001790323 5.594 2000     2000      1     4.518
## 777   0034374265 4.445 2000     2000      1     3.207
## 780   0034403543 3.822 2000     2000      1     2.584
## 782   0039024028 4.888 2000     2000      1     3.650
## 783   0039449757 3.822 2000     2000      1     2.691
## 788  16244399574 5.423 2000     2000      1     4.347
## 789  33644528943 4.682 2000     2000      1     3.606
## 809   6344283665 5.673 2000     2000      1     4.597
## 1244  0001378820 3.701 2002     1800      4     2.679
## 4419 77953083241 3.587 2009     2000      1     2.567
## 4420 77954605495 3.743 2009     2000      1     2.723
## 4546 74949137297 3.531 2009     2000      1     2.511
## 4636 85017105478 3.948 2009     2000      1     2.766
## 4647 85017167724 3.740 2009     2000      1     2.720
## 5164 79957454147 3.716 2010     2000      1     2.560
## 5428 85017138638 4.056 2010     2000      1     2.955
## 5443 74349104010 3.947 2010     2000      1     2.846
## 5490 77956947655 3.655 2010     2000      1     2.554
## 5806 80053983353 3.681 2011     2000      1     2.604
## 5978 79960792628 3.790 2011     2000      1     2.713
## 6240 79957827461 3.663 2011     2000      1     2.586
```

```
## 6299 79958135412 4.109 2011     2000      1     3.032
## 6702 84868030545 3.603 2012     2000      1     2.534
## 6703 84868030921 3.692 2012     2000      1     2.623
## 6749 84865191893 4.782 2012     2000      1     3.657
## 6839 84861658535 5.103 2012     2000      1     4.034
## 7015 84856199392 3.603 2012     2000      1     2.534
## 7066 84857150817 4.544 2012     2000      1     3.475
## 7068 84857174901 3.678 2012     2000      1     2.609
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.2943 -0.5644 -0.0184  0.5683  4.5919
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)         0.906495   0.065915   13.75  < 2e-16 ***
## FirstAuthorFemale1  0.122646   0.033143    3.70  0.00022 ***
## Year1997           -0.000705   0.096382   -0.01  0.99417
## Year1998           -0.065508   0.094874   -0.69  0.48993
## Year1999            0.061006   0.099630    0.61  0.54035
## Year2000            0.174616   0.126253    1.38  0.16672
## Year2001            0.134094   0.091260    1.47  0.14180
## Year2002            0.120989   0.088640    1.36  0.17234
## Year2003            0.263252   0.092061    2.86  0.00426 **
## Year2004            0.101551   0.084747    1.20  0.23087
## Year2005            0.265140   0.079683    3.33  0.00088 ***
## Year2006            0.068868   0.077492    0.89  0.37420
## Year2007            0.121613   0.077516    1.57  0.11675
## Year2008            0.074046   0.076387    0.97  0.33243
## Year2009            0.120327   0.075699    1.59  0.11200
## Year2010            0.203985   0.080320    2.54  0.01113 *
## Year2011            0.181911   0.077241    2.36  0.01856 *
## Year2012            0.171071   0.081331    2.10  0.03549 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.846
## Multiple R-squared:  0.011,  Adjusted R-squared:  0.00724
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
##  4 observations c(548,561,581,4373)
##   are outliers with |weight| = 0 ( < 2.2e-05);
##  400 weights are ~= 1. The remaining 4135 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
```

```
##  0.0118  0.8700  0.9520  0.9080  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          2.20e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##                   GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.012  1           1.006
## Year             1.012 16           1.000
```

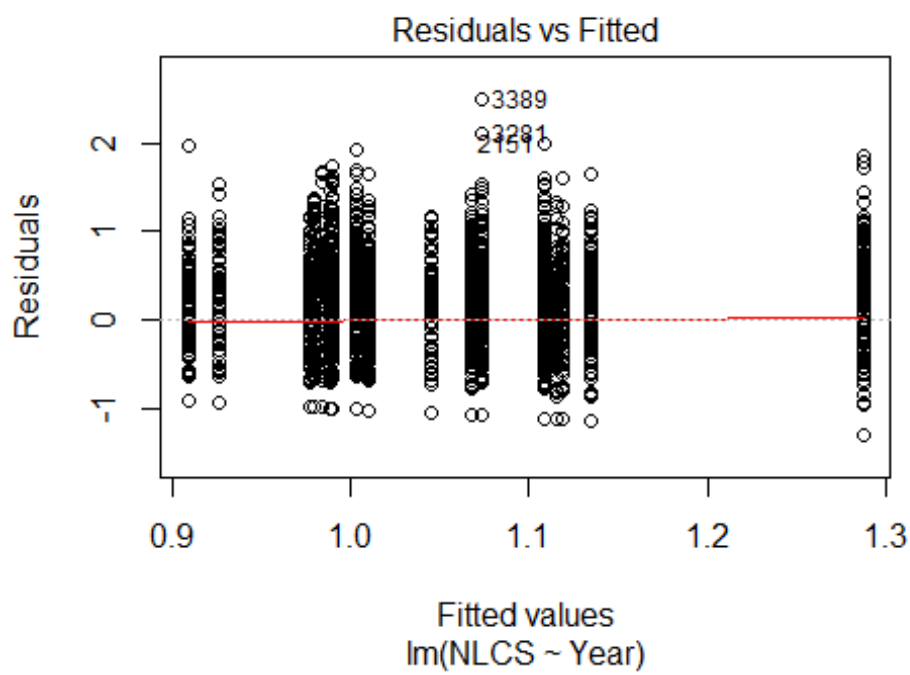



```
## [1] "List of  33 outliers with residuals above  2.5"
##         ScopusId  NLCS Year OneField Fields residuals
## 716   0042373524 3.688 2000     2000      1     2.612
## 717   0042846710 4.822 2000     2000      1     3.746
## 718   0347037232 4.888 2000     2000      1     3.757
```

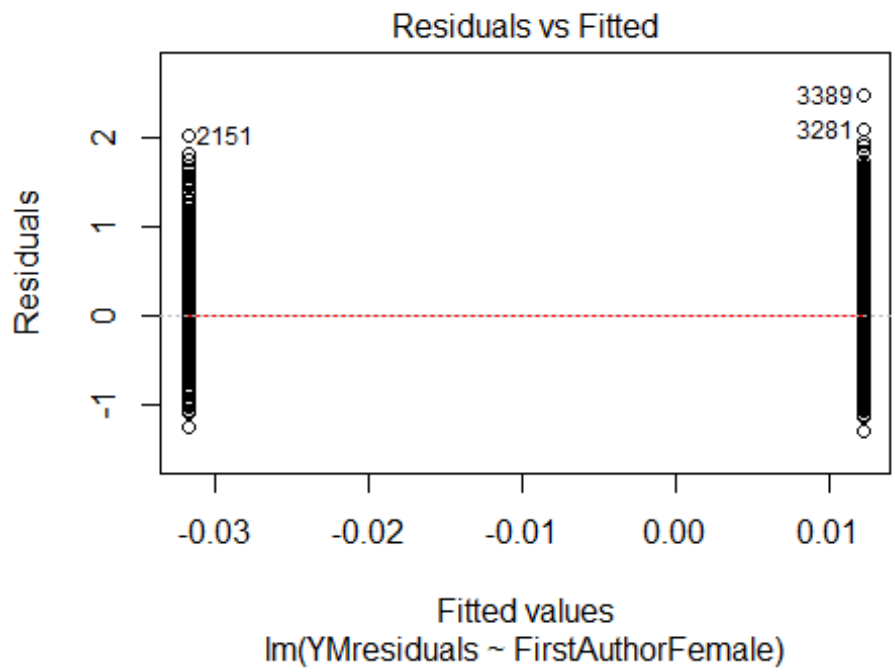
```
## 724  84891802579 3.945 2000     2000      1     2.869
## 772   0001790323 5.594 2000     2000      1     4.518
## 777   0034374265 4.445 2000     2000      1     3.207
## 780   0034403543 3.822 2000     2000      1     2.584
## 782   0039024028 4.888 2000     2000      1     3.650
## 783   0039449757 3.822 2000     2000      1     2.691
## 788  16244399574 5.423 2000     2000      1     4.347
## 789  33644528943 4.682 2000     2000      1     3.606
## 809   6344283665 5.673 2000     2000      1     4.597
## 1244  0001378820 3.701 2002     1800      4     2.679
## 4419 77953083241 3.587 2009     2000      1     2.567
## 4420 77954605495 3.743 2009     2000      1     2.723
## 4546 74949137297 3.531 2009     2000      1     2.511
## 4636 85017105478 3.948 2009     2000      1     2.766
## 4647 85017167724 3.740 2009     2000      1     2.720
## 5164 79957454147 3.716 2010     2000      1     2.560
## 5428 85017138638 4.056 2010     2000      1     2.955
## 5443 74349104010 3.947 2010     2000      1     2.846
## 5490 77956947655 3.655 2010     2000      1     2.554
## 5806 80053983353 3.681 2011     2000      1     2.604
## 5978 79960792628 3.790 2011     2000      1     2.713
## 6240 79957827461 3.663 2011     2000      1     2.586
## 6299 79958135412 4.109 2011     2000      1     3.032
## 6702 84868030545 3.603 2012     2000      1     2.534
## 6703 84868030921 3.692 2012     2000      1     2.623
## 6749 84865191893 4.782 2012     2000      1     3.657
## 6839 84861658535 5.103 2012     2000      1     4.034
## 7015 84856199392 3.603 2012     2000      1     2.534
## 7066 84857150817 4.544 2012     2000      1     3.475
## 7068 84857174901 3.678 2012     2000      1     2.609
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.3087 -0.5664 -0.0216  0.5734  4.5931
##
## Coefficients:
##                   Estimate Std. Error t value Pr(>|t|)
## (Intercept)        0.89831    0.06475   13.87  < 2e-16 ***
## LastAuthorFemale1  0.14032    0.03217    4.36  1.3e-05 ***
## Year1997           0.00384    0.09478    0.04  0.96771
## Year1998          -0.05797    0.09431   -0.61  0.53877
## Year1999           0.06668    0.09824    0.68  0.49733
## Year2000           0.18161    0.12497    1.45  0.14623
## Year2001           0.14249    0.09038    1.58  0.11497
## Year2002           0.12800    0.08771    1.46  0.14451
```

```
## Year2003           0.26185    0.09114    2.87  0.00409 **
## Year2004           0.10305    0.08391    1.23  0.21949
## Year2005           0.27009    0.07891    3.42  0.00063 ***
## Year2006           0.06961    0.07666    0.91  0.36392
## Year2007           0.12431    0.07679    1.62  0.10555
## Year2008           0.08002    0.07540    1.06  0.28862
## Year2009           0.12524    0.07476    1.68  0.09395 .
## Year2010           0.20530    0.07931    2.59  0.00967 **
## Year2011           0.18162    0.07624    2.38  0.01726 *
## Year2012           0.17530    0.08048    2.18  0.02945 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.845
## Multiple R-squared:  0.0121, Adjusted R-squared:  0.00837
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
##  4 observations c(548,561,581,4373)
##   are outliers with |weight| = 0 ( < 2.2e-05);
##  356 weights are ~= 1. The remaining 4179 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0055  0.8710  0.9520  0.9090  0.9860  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          2.20e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Sample size for the above analysis:  4539"
## [1] ""
## [1] ""
## [1] "##################################"
## [1] "Analysis of AJSC 2001"
## [1] "##################################"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  146  125  124  124  146  129  142  177  148  186  189  260  316  383  355
```

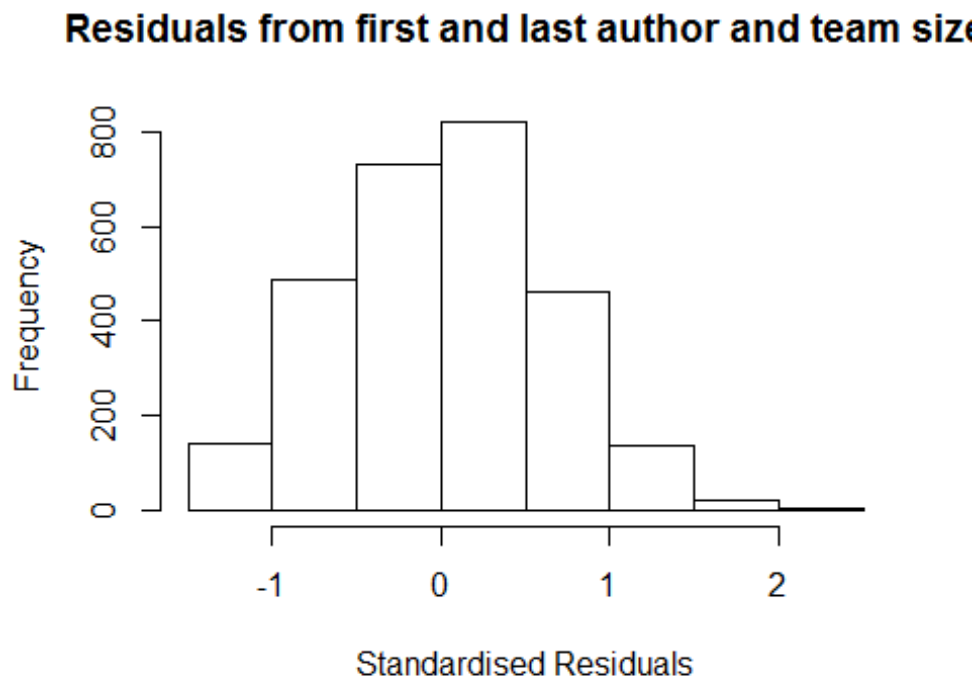
```
## 2011 2012
##  359  307
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  126  105  108  106  125  103  128  159  127  157  153  221  268  330  279
## 2011 2012
##  286  256
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  116   99  103   99  121   97  120  149  120  140  144  202  250  300  258
## 2011 2012
##  256  227
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
##  Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 42, df = 16, p-value = 4e-04
```




```
##
##  Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 1.2, df = 1, p-value = 0.3
```

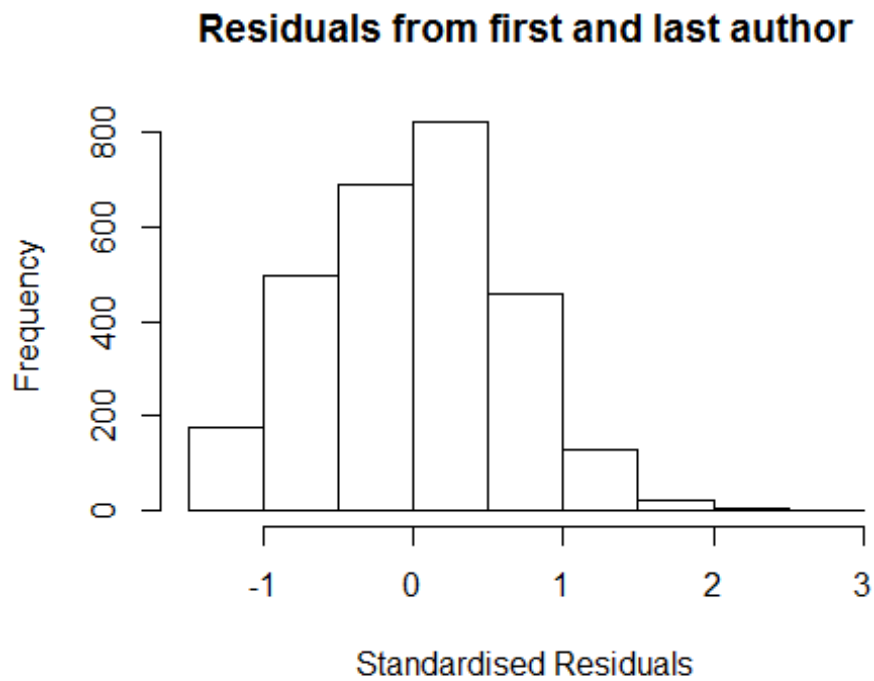



```
## [1] "Female first author team size 2018 geometric mean: 1.67619243545721"
## [1] "Male first author team size 2018 geometric mean: 1.6619779624367"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 9300, p-value = 1
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 1.57391888176074"
## [1] "Male last author team size 2018 geometric mean: 1.70326077386946"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 7800, p-value = 0.2
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.662  1           1.289
## LastAuthorFemale  1.643  1           1.282
## UniqueAuthors     1.098  4           1.012
## Year              1.179 16           1.005
```



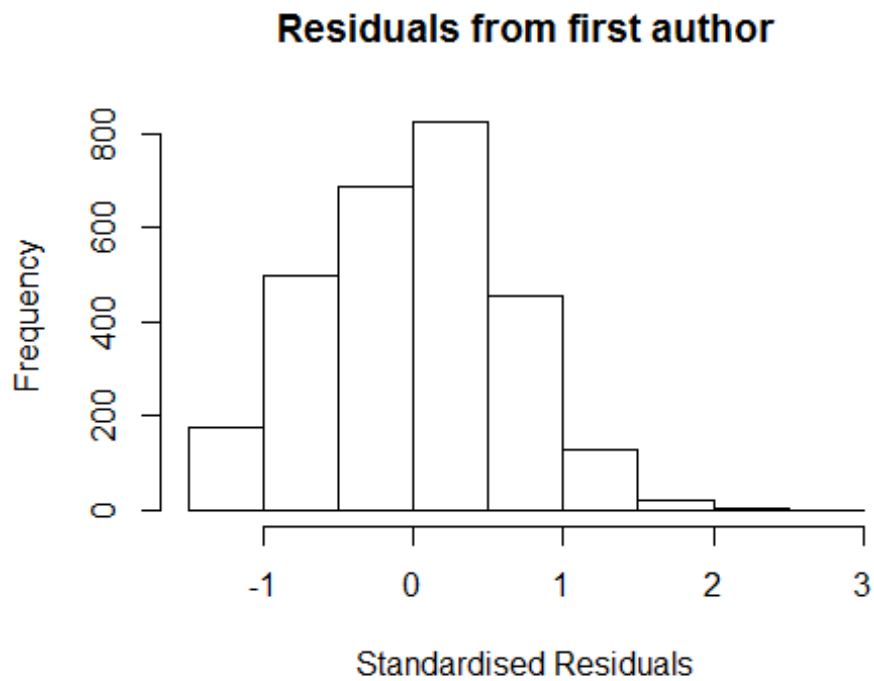

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
UniqueAuthors +
##     Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##     k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.4353 -0.4567  0.0249  0.4499  2.3732
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         0.91272    0.06911   13.21  < 2e-16 ***
## FirstAuthorFemale1 -0.03971    0.03628   -1.09   0.2739
## LastAuthorFemale1  -0.00593    0.03659   -0.16   0.8713
## UniqueAuthors2      0.22265    0.02871    7.76  1.2e-14 ***
## UniqueAuthors3      0.27342    0.03671    7.45  1.3e-13 ***
## UniqueAuthors4      0.34928    0.06939    5.03  5.1e-07 ***
## UniqueAuthors5      0.25822    0.13297    1.94   0.0522 .
## Year1997           -0.10102    0.09335   -1.08   0.2793
## Year1998           -0.00714    0.09252   -0.08   0.9385
## Year1999            0.01476    0.08825    0.17   0.8672
```

```
## Year2000            0.00548    0.08350    0.07   0.9477
## Year2001            0.10702    0.09695    1.10   0.2697
## Year2002            0.13550    0.09070    1.49   0.1353
## Year2003           -0.05286    0.08527   -0.62   0.5354
## Year2004            0.11913    0.08188    1.45   0.1458
## Year2005            0.29991    0.09852    3.04   0.0024 **
## Year2006           -0.11174    0.08482   -1.32   0.1878
## Year2007            0.01318    0.08294    0.16   0.8738
## Year2008            0.06450    0.08068    0.80   0.4241
## Year2009           -0.05252    0.07545   -0.70   0.4865
## Year2010           -0.00628    0.07564   -0.08   0.9338
## Year2011            0.00670    0.07755    0.09   0.9312
## Year2012           -0.05756    0.08225   -0.70   0.4841
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.659
## Multiple R-squared:  0.0549, Adjusted R-squared:  0.0474
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  237 weights are ~= 1. The remaining 2564 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.167   0.866   0.950   0.915   0.985   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          3.57e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.708  1           1.307
## LastAuthorFemale  1.691  1           1.301
## Year              1.100 16           1.003
```



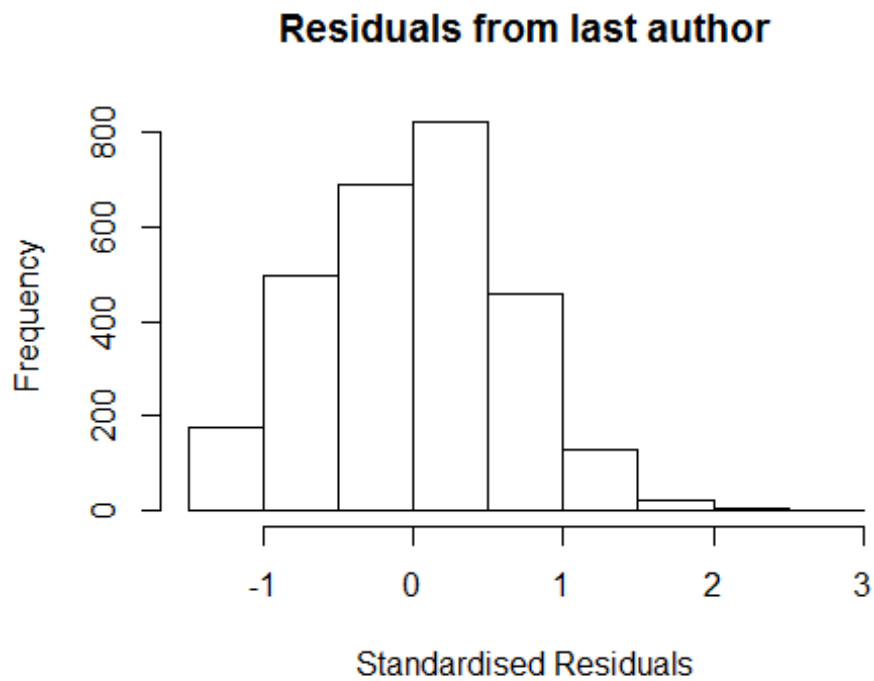

```
## [1] "List of  1 outliers with residuals above  2.5"
##         ScopusId  NLCS Year OneField Fields residuals
## 3389 79960844583 3.566 2011     2001      1      2.53
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.3222 -0.4634  0.0292  0.4465  2.5303
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)          1.01033    0.06810   14.84   <2e-16 ***
## FirstAuthorFemale1 -0.03633    0.03757   -0.97   0.3336
## LastAuthorFemale1  -0.01363    0.03798   -0.36   0.7197
## Year1997           -0.11143    0.09369   -1.19   0.2344
## Year1998           -0.00537    0.09226   -0.06   0.9536
## Year1999            0.03380    0.09030    0.37   0.7082
## Year2000           -0.02476    0.08513   -0.29   0.7712
## Year2001            0.12608    0.09721    1.30   0.1947
## Year2002            0.14196    0.09240    1.54   0.1246
## Year2003           -0.03157    0.08733   -0.36   0.7177
## Year2004            0.12793    0.08331    1.54   0.1247
## Year2005            0.31192    0.09943    3.14   0.0017 **
```

```
## Year2006           -0.08683    0.08586   -1.01   0.3119
## Year2007            0.04399    0.08316    0.53   0.5969
## Year2008            0.09282    0.08177    1.14   0.2564
## Year2009           -0.04623    0.07634   -0.61   0.5449
## Year2010            0.00456    0.07638    0.06   0.9524
## Year2011            0.02538    0.07850    0.32   0.7465
## Year2012           -0.01686    0.08396   -0.20   0.8409
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.671
## Multiple R-squared:  0.0199, Adjusted R-squared:  0.0136
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  224 weights are ~= 1. The remaining 2577 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.123   0.861   0.950   0.916   0.984   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          3.57e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.071  1           1.035
## Year              1.071 16           1.002
```




```
## [1] "List of  1 outliers with residuals above  2.5"
##         ScopusId  NLCS Year OneField Fields residuals
## 3389 79960844583 3.566 2011     2001      1      2.53
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##    Min     1Q Median     3Q    Max
## -1.321 -0.462  0.031  0.446  2.531
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)          1.00918    0.06810   14.82   <2e-16 ***
## FirstAuthorFemale1 -0.04508    0.02982   -1.51   0.1307
## Year1997           -0.11306    0.09342   -1.21   0.2263
## Year1998           -0.00582    0.09229   -0.06   0.9497
## Year1999            0.03291    0.09029    0.36   0.7155
## Year2000           -0.02475    0.08518   -0.29   0.7714
## Year2001            0.12562    0.09717    1.29   0.1962
## Year2002            0.14154    0.09243    1.53   0.1258
## Year2003           -0.03217    0.08732   -0.37   0.7126
## Year2004            0.12809    0.08331    1.54   0.1243
## Year2005            0.31136    0.09939    3.13   0.0017 **
## Year2006           -0.08620    0.08586   -1.00   0.3155
```

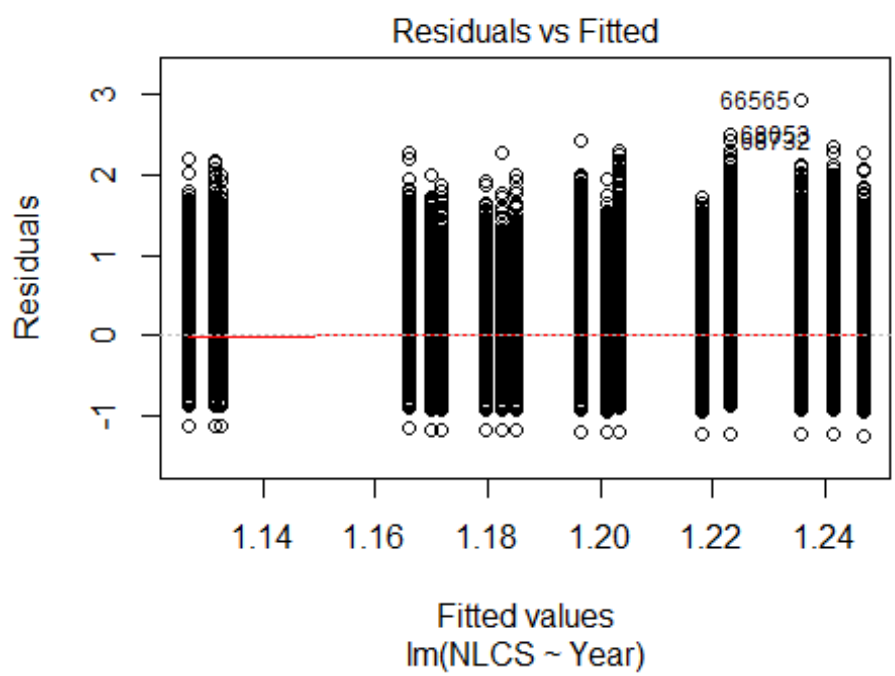
```
## Year2007            0.04414    0.08318    0.53   0.5957
## Year2008            0.09266    0.08179    1.13   0.2574
## Year2009           -0.04636    0.07636   -0.61   0.5438
## Year2010            0.00484    0.07642    0.06   0.9495
## Year2011            0.02586    0.07855    0.33   0.7420
## Year2012           -0.01746    0.08388   -0.21   0.8351
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.67
## Multiple R-squared:  0.0199, Adjusted R-squared:  0.0139
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  229 weights are ~= 1. The remaining 2572 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.123   0.861   0.950   0.916   0.984   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          3.57e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##                  GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.06  1           1.030
## Year             1.06 16           1.002
```



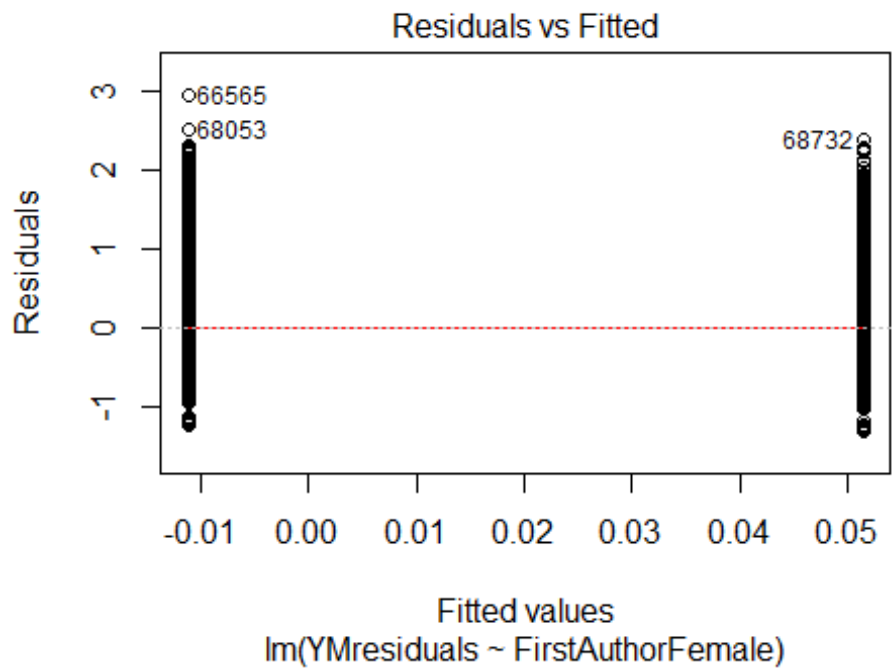

```
## [1] "List of  1 outliers with residuals above  2.5"
##         ScopusId  NLCS Year OneField Fields residuals
## 3389 79960844583 3.566 2011     2001      1      2.53
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.3215 -0.4640  0.0274  0.4476  2.5345
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.00237    0.06717   14.92   <2e-16 ***
## LastAuthorFemale1 -0.03703    0.03018   -1.23   0.2199
## Year1997          -0.10770    0.09382   -1.15   0.2511
## Year1998          -0.00492    0.09223   -0.05   0.9575
## Year1999           0.03641    0.09010    0.40   0.6862
## Year2000          -0.02101    0.08482   -0.25   0.8044
## Year2001           0.12889    0.09698    1.33   0.1839
## Year2002           0.14650    0.09219    1.59   0.1121
## Year2003          -0.02640    0.08718   -0.30   0.7621
## Year2004           0.13224    0.08299    1.59   0.1112
## Year2005           0.31913    0.09891    3.23   0.0013 **
## Year2006          -0.08294    0.08560   -0.97   0.3327
```

```
## Year2007           0.04968    0.08275    0.60   0.5483
## Year2008           0.09821    0.08106    1.21   0.2258
## Year2009          -0.04065    0.07587   -0.54   0.5922
## Year2010           0.00902    0.07598    0.12   0.9055
## Year2011           0.02909    0.07819    0.37   0.7099
## Year2012          -0.01207    0.08353   -0.14   0.8851
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.671
## Multiple R-squared:  0.0196, Adjusted R-squared:  0.0136
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  233 weights are ~= 1. The remaining 2568 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.122   0.861   0.950   0.915   0.984   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          3.57e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Sample size for the above analysis:  2801"
## [1] ""
## [1] ""
## [1] "####################################"
## [1] "Analysis of AJSC 2002"
## [1] "####################################"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 3388 3217 3081 3025 3296 2330 3431 2747 3075 3277 3563 3839 4458 4924 4941
## 2011 2012
## 4871 4628
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 2831 2621 2490 2507 2615 1706 2841 2215 2493 2684 2892 3103 3586 3930 3917
## 2011 2012
```

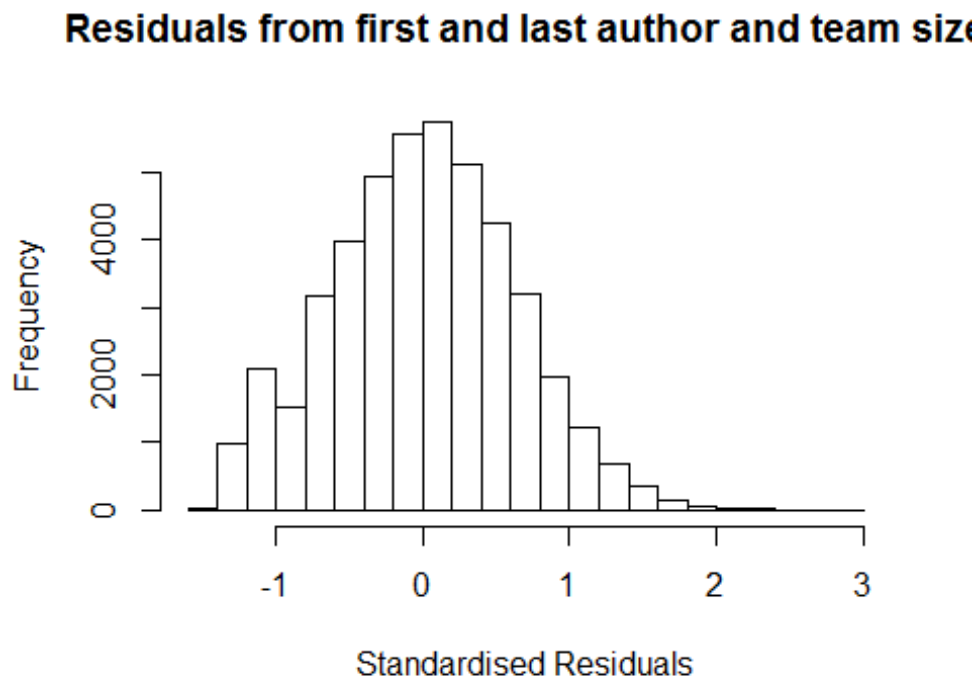
```
## 3868 3681
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 2616 2431 2307 2353 2437 1553 2601 2028 2241 2432 2589 2754 3154 3488 3465
## 2011 2012
## 3405 3252
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
##  Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 680, df = 16, p-value <2e-16
```




```
##
##  Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 0.66, df = 1, p-value = 0.4
```

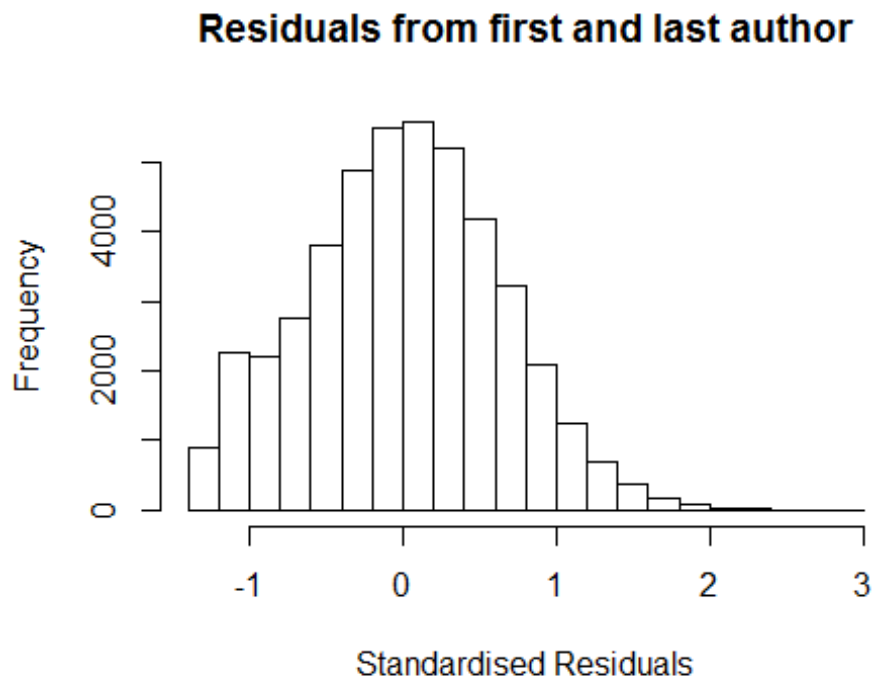



```
## [1] "Female first author team size 2018 geometric mean: 1.8898039485991"
## [1] "Male first author team size 2018 geometric mean: 1.73770336268746"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 880000, p-value = 0.001
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 1.82699090197659"
## [1] "Male last author team size 2018 geometric mean: 1.75761786004683"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 830000, p-value = 0.2
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##                     GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.347  1           1.161
## LastAuthorFemale  1.349  1           1.161
## UniqueAuthors     1.025  4           1.003
## Year              1.027 16           1.001
```



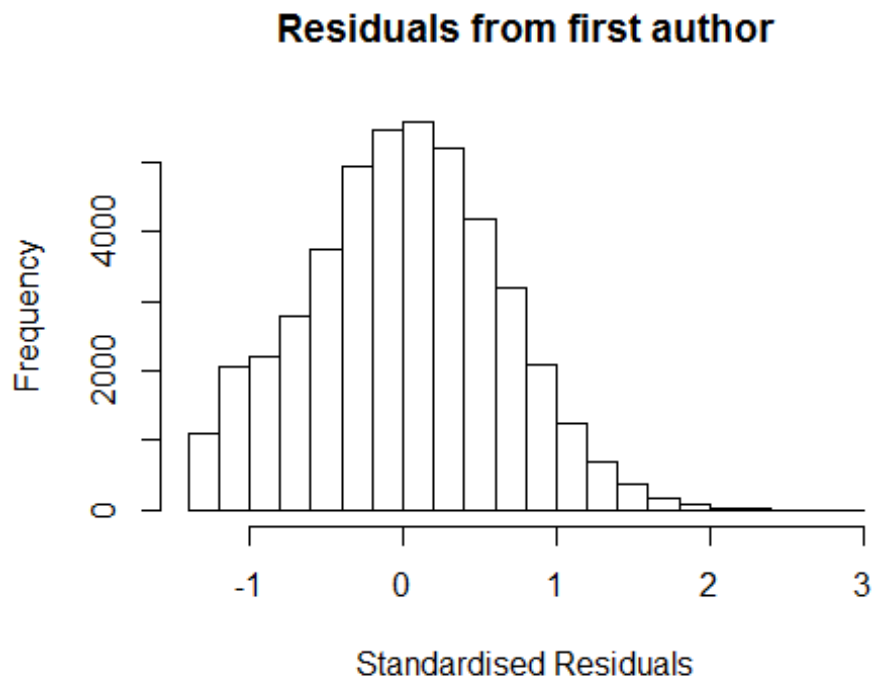

```
## [1] "List of  1 outliers with residuals above  2.5"
##          ScopusId  NLCS Year OneField Fields residuals
## 66565 84906331429 4.168 2011     1400      5     2.882
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
UniqueAuthors +
##     Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##     k.max = 1000))
##  \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.50289 -0.42574  0.00914  0.43000  2.88152
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.06343    0.01373   77.45  < 2e-16 ***
## FirstAuthorFemale1  0.05188    0.00935    5.55  2.9e-08 ***
## LastAuthorFemale1   0.02028    0.00940    2.16  0.03103 *
## UniqueAuthors2      0.21825    0.00684   31.89  < 2e-16 ***
## UniqueAuthors3      0.29187    0.00916   31.87  < 2e-16 ***
## UniqueAuthors4      0.35098    0.01955   17.96  < 2e-16 ***
## UniqueAuthors5      0.24943    0.03183    7.84  4.7e-15 ***
## Year1997           -0.03095    0.01866   -1.66  0.09729 .
## Year1998           -0.07000    0.01914   -3.66  0.00025 ***
## Year1999           -0.06445    0.01841   -3.50  0.00047 ***
```

```
## Year2000           -0.06641    0.01846   -3.60  0.00032 ***
## Year2001           -0.01309    0.02098   -0.62  0.53283
## Year2002           -0.03792    0.01780   -2.13  0.03316 *
## Year2003           -0.02412    0.01842   -1.31  0.19039
## Year2004           -0.01797    0.01815   -0.99  0.32208
## Year2005           -0.04166    0.01788   -2.33  0.01979 *
## Year2006           -0.03698    0.01728   -2.14  0.03236 *
## Year2007           -0.01130    0.01772   -0.64  0.52366
## Year2008            0.01631    0.01781    0.92  0.35990
## Year2009           -0.02625    0.01774   -1.48  0.13882
## Year2010            0.00546    0.01817    0.30  0.76367
## Year2011            0.00480    0.01825    0.26  0.79257
## Year2012           -0.02988    0.01924   -1.55  0.12045
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.632
## Multiple R-squared:  0.0431, Adjusted R-squared:  0.0426
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  3816 weights are ~= 1. The remaining 41290 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0027  0.8740  0.9500  0.9090  0.9860  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          2.22e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.346  1           1.160
## LastAuthorFemale  1.347  1           1.161
## Year              1.006 16           1.000
```



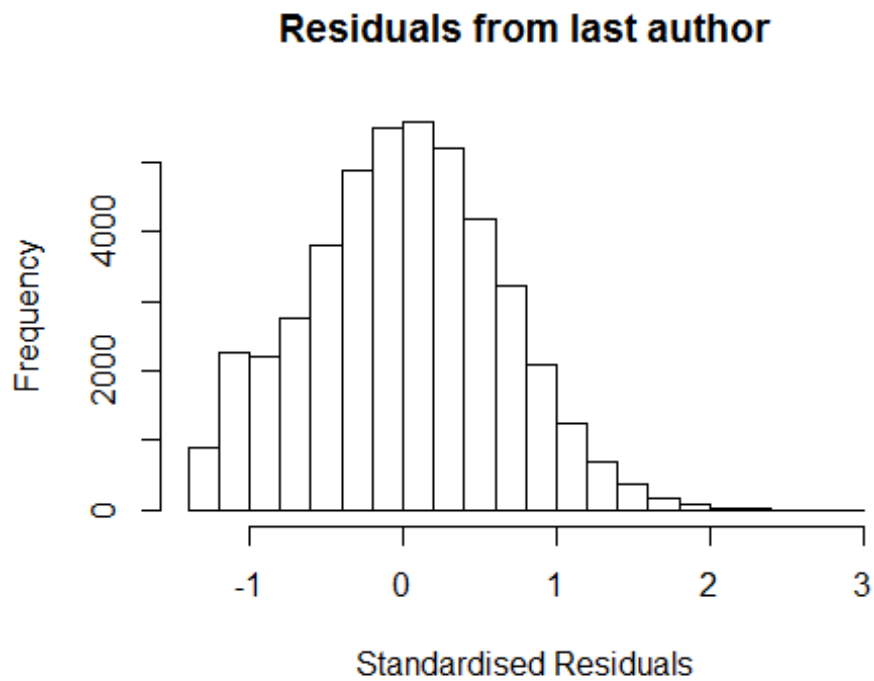

```
## [1] "List of  2 outliers with residuals above  2.5"
##          ScopusId  NLCS Year OneField Fields residuals
## 66565 84906331429 4.168 2011     1400      5     2.962
## 68053 84866640011 3.733 2012     1400      5     2.558
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.30222 -0.43036  0.00951  0.43541  2.96229
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)          1.16762    0.01359   85.93  < 2e-16 ***
## FirstAuthorFemale1   0.06185    0.00948    6.52  7.1e-11 ***
## LastAuthorFemale1    0.02681    0.00952    2.81  0.00488 **
## Year1997            -0.02895    0.01892   -1.53  0.12592
## Year1998            -0.07207    0.01932   -3.73  0.00019 ***
## Year1999            -0.06477    0.01873   -3.46  0.00055 ***
## Year2000            -0.06214    0.01869   -3.32  0.00089 ***
## Year2001            -0.00915    0.02139   -0.43  0.66861
## Year2002            -0.02068    0.01805   -1.15  0.25185
## Year2003            -0.01109    0.01869   -0.59  0.55288
## Year2004             0.00981    0.01836    0.53  0.59325
```

```
## Year2005           -0.02063    0.01811   -1.14  0.25466
## Year2006           -0.01125    0.01752   -0.64  0.52063
## Year2007            0.01552    0.01787    0.87  0.38528
## Year2008            0.04594    0.01807    2.54  0.01102 *
## Year2009            0.00119    0.01799    0.07  0.94725
## Year2010            0.03210    0.01848    1.74  0.08232 .
## Year2011            0.03809    0.01852    2.06  0.03970 *
## Year2012            0.00697    0.01955    0.36  0.72166
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.644
## Multiple R-squared:  0.00513,    Adjusted R-squared:  0.00473
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  3793 weights are ~= 1. The remaining 41313 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0012  0.8680  0.9500  0.9090  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          2.22e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.004  1           1.002
## Year              1.004 16           1.000
```




```
## [1] "List of  2 outliers with residuals above  2.5"
##          ScopusId  NLCS Year OneField Fields residuals
## 66565 84906331429 4.168 2011     1400      5     2.962
## 68053 84866640011 3.733 2012     1400      5     2.558
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.29180 -0.43120  0.00987  0.43594  2.95981
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.16959    0.01356   86.28  < 2e-16 ***
## FirstAuthorFemale1  0.07583    0.00818    9.27  < 2e-16 ***
## Year1997           -0.02895    0.01891   -1.53  0.12581
## Year1998           -0.07220    0.01931   -3.74  0.00019 ***
## Year1999           -0.06486    0.01872   -3.46  0.00053 ***
## Year2000           -0.06224    0.01868   -3.33  0.00087 ***
## Year2001           -0.00924    0.02137   -0.43  0.66542
## Year2002           -0.02056    0.01805   -1.14  0.25456
## Year2003           -0.01129    0.01868   -0.60  0.54547
## Year2004            0.01026    0.01836    0.56  0.57636
## Year2005           -0.02027    0.01810   -1.12  0.26276
```

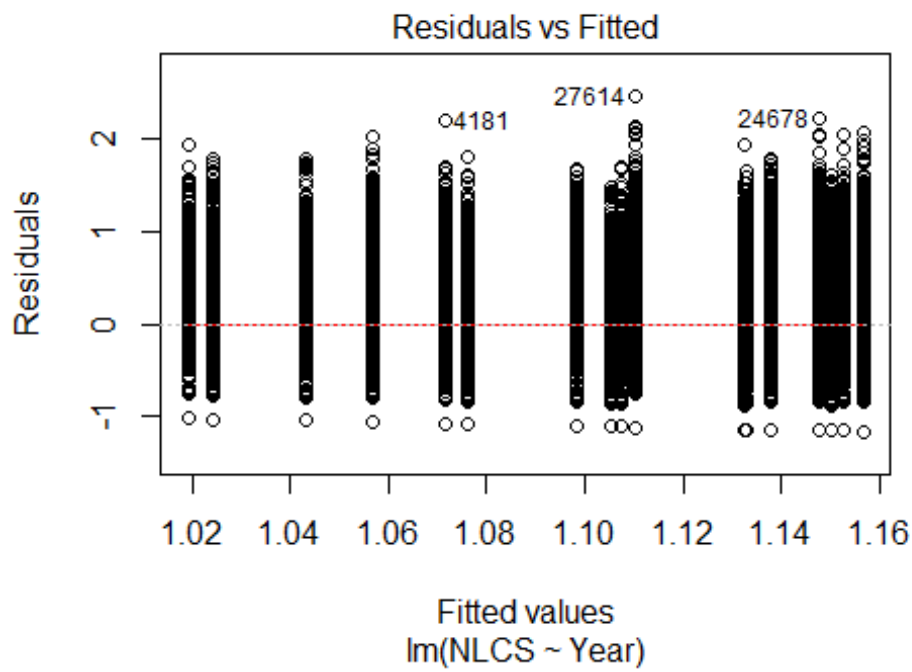
```
## Year2006           -0.01077    0.01751   -0.61  0.53868
## Year2007            0.01595    0.01787    0.89  0.37218
## Year2008            0.04638    0.01807    2.57  0.01027 *
## Year2009            0.00175    0.01799    0.10  0.92270
## Year2010            0.03254    0.01847    1.76  0.07814 .
## Year2011            0.03860    0.01852    2.08  0.03710 *
## Year2012            0.00761    0.01955    0.39  0.69726
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.644
## Multiple R-squared:  0.00494,    Adjusted R-squared:  0.00457
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  3787 weights are ~= 1. The remaining 41319 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0015  0.8680  0.9500  0.9100  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          2.22e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##                   GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.004  1           1.002
## Year             1.004 16           1.000
```



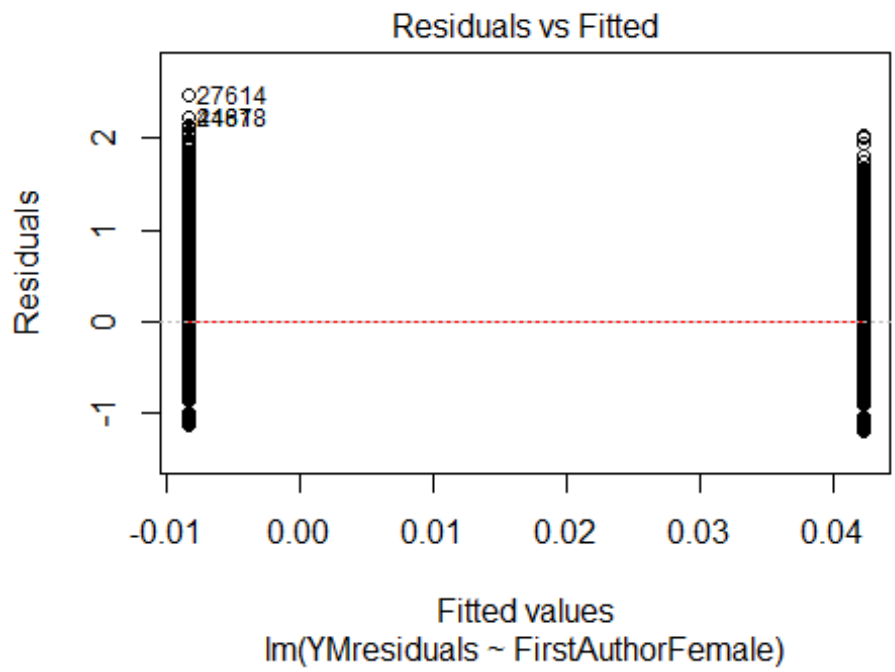

```
## [1] "List of  2 outliers with residuals above  2.5"
##          ScopusId  NLCS Year OneField Fields residuals
## 66565 84906331429 4.168 2011     1400      5     2.962
## 68053 84866640011 3.733 2012     1400      5     2.558
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.27880 -0.43232  0.00963  0.43506  2.95666
##
## Coefficients:
##                   Estimate Std. Error t value Pr(>|t|)
## (Intercept)        1.17194    0.01359   86.26  < 2e-16 ***
## LastAuthorFemale1  0.05938    0.00820    7.24  4.4e-13 ***
## Year1997          -0.02870    0.01894   -1.52  0.12972
## Year1998          -0.07117    0.01933   -3.68  0.00023 ***
## Year1999          -0.06422    0.01875   -3.43  0.00061 ***
## Year2000          -0.06152    0.01869   -3.29  0.00100 ***
## Year2001          -0.00855    0.02141   -0.40  0.68968
## Year2002          -0.01988    0.01805   -1.10  0.27076
## Year2003          -0.00964    0.01869   -0.52  0.60583
## Year2004           0.01007    0.01837    0.55  0.58357
## Year2005          -0.01992    0.01813   -1.10  0.27175
```

```
## Year2006          -0.01085    0.01753   -0.62  0.53586
## Year2007           0.01600    0.01789    0.89  0.37097
## Year2008           0.04749    0.01810    2.62  0.00870 **
## Year2009           0.00246    0.01800    0.14  0.89119
## Year2010           0.03401    0.01848    1.84  0.06579 .
## Year2011           0.03941    0.01853    2.13  0.03341 *
## Year2012           0.00922    0.01956    0.47  0.63745
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.644
## Multiple R-squared:  0.00414,    Adjusted R-squared:  0.00376
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  3771 weights are ~= 1. The remaining 41335 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0016  0.8680  0.9500  0.9100  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          2.22e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Sample size for the above analysis:  45106"
## [1] ""
## [1] ""
## [1] "####################################"
## [1] "Analysis of AJSC 2003"
## [1] "####################################"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 1285 1160 1168 1238 1222 1296 1357 1075 1195 1256 1433 1500 1753 1914 1887
## 2011 2012
## 1997 1954
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 1060  950  957 1017  993 1042 1146  856  943 1001 1142 1191 1365 1489 1466
```

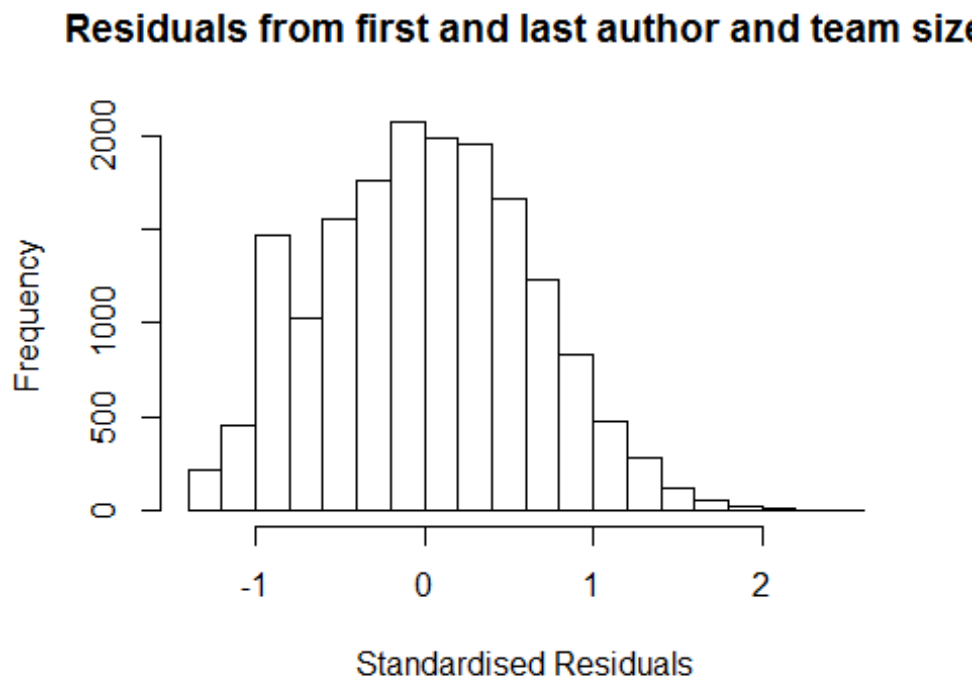
```
## 2011 2012
## 1543 1501
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  944  864  876  937  923  933 1023  776  827  878  999  998 1160 1261 1236
## 2011 2012
## 1309 1267
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
##  Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 230, df = 16, p-value <2e-16
```




```
##
##  Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 0.011, df = 1, p-value = 0.9
```

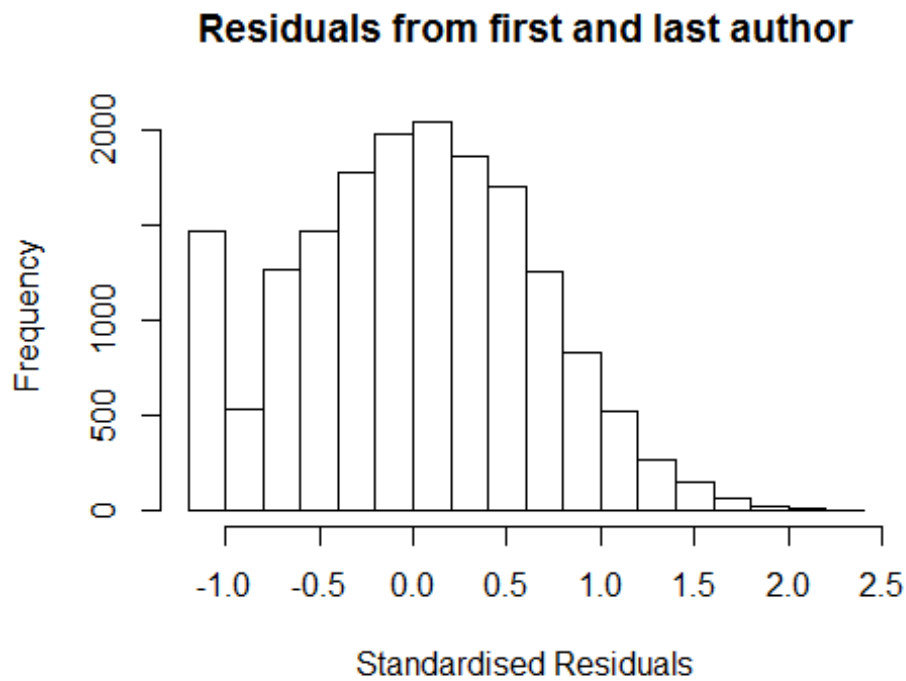



```
## [1] "Female first author team size 2018 geometric mean: 1.94241642711268"
## [1] "Male first author team size 2018 geometric mean: 1.7755665599566"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 92000, p-value = 0.03
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 1.94618528925534"
## [1] "Male last author team size 2018 geometric mean: 1.7755272795136"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 91000, p-value = 0.02
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.268  1           1.126
## LastAuthorFemale  1.270  1           1.127
## UniqueAuthors     1.027  4           1.003
## Year              1.028 16           1.001
```



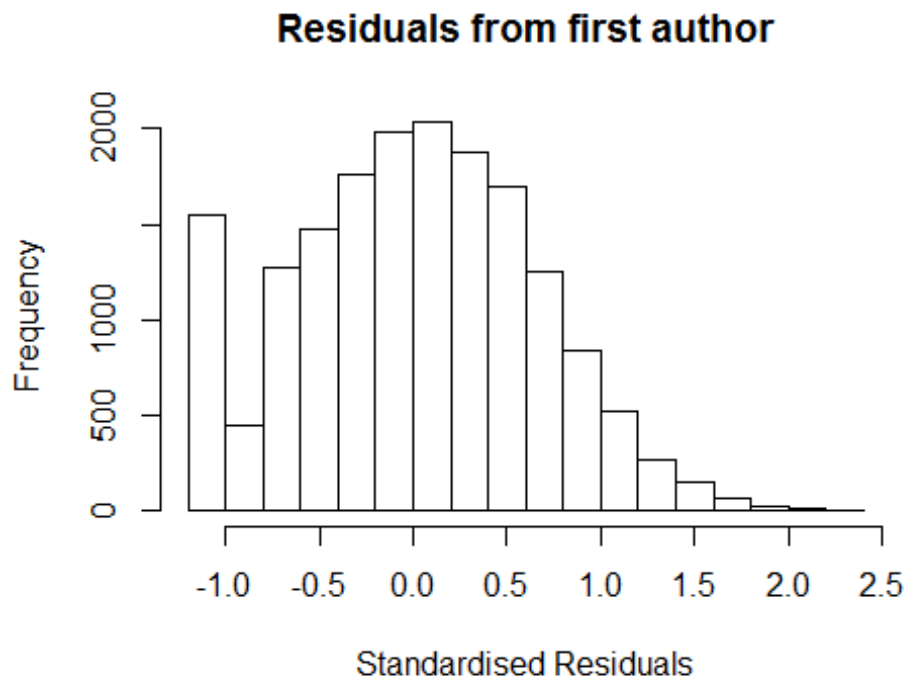

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
UniqueAuthors +
##     Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##     k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.3868 -0.4520  0.0056  0.4416  2.4474
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         0.89598    0.02237   40.05  < 2e-16 ***
## FirstAuthorFemale1  0.05040    0.01512    3.33  0.00086 ***
## LastAuthorFemale1   0.00390    0.01521    0.26  0.79773
## UniqueAuthors2      0.23293    0.01126   20.69  < 2e-16 ***
## UniqueAuthors3      0.31658    0.01408   22.49  < 2e-16 ***
## UniqueAuthors4      0.41339    0.03171   13.04  < 2e-16 ***
## UniqueAuthors5      0.38954    0.07443    5.23  1.7e-07 ***
## Year1997           -0.03395    0.03048   -1.11  0.26535
## Year1998           -0.02524    0.03099   -0.81  0.41544
## Year1999            0.01279    0.02955    0.43  0.66517
```

```
## Year2000           -0.02914    0.03038   -0.96  0.33747
## Year2001            0.02385    0.03079    0.77  0.43846
## Year2002            0.00402    0.02869    0.14  0.88847
## Year2003            0.04152    0.02958    1.40  0.16041
## Year2004            0.03449    0.02961    1.16  0.24408
## Year2005            0.06124    0.02946    2.08  0.03764 *
## Year2006            0.05312    0.02854    1.86  0.06273 .
## Year2007            0.04319    0.02919    1.48  0.13896
## Year2008            0.04128    0.02906    1.42  0.15550
## Year2009            0.03898    0.02923    1.33  0.18237
## Year2010            0.05084    0.03034    1.68  0.09380 .
## Year2011            0.03058    0.02917    1.05  0.29460
## Year2012            0.00184    0.03060    0.06  0.95203
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.657
## Multiple R-squared:  0.0495, Adjusted R-squared:  0.0483
## Convergence in 10 IRWLS iterations
##
## Robustness weights:
##  1427 weights are ~= 1. The remaining 15784 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.135   0.870   0.950   0.916   0.985   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          5.81e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.246  1           1.116
## LastAuthorFemale  1.250  1           1.118
## Year              1.008 16           1.000
```



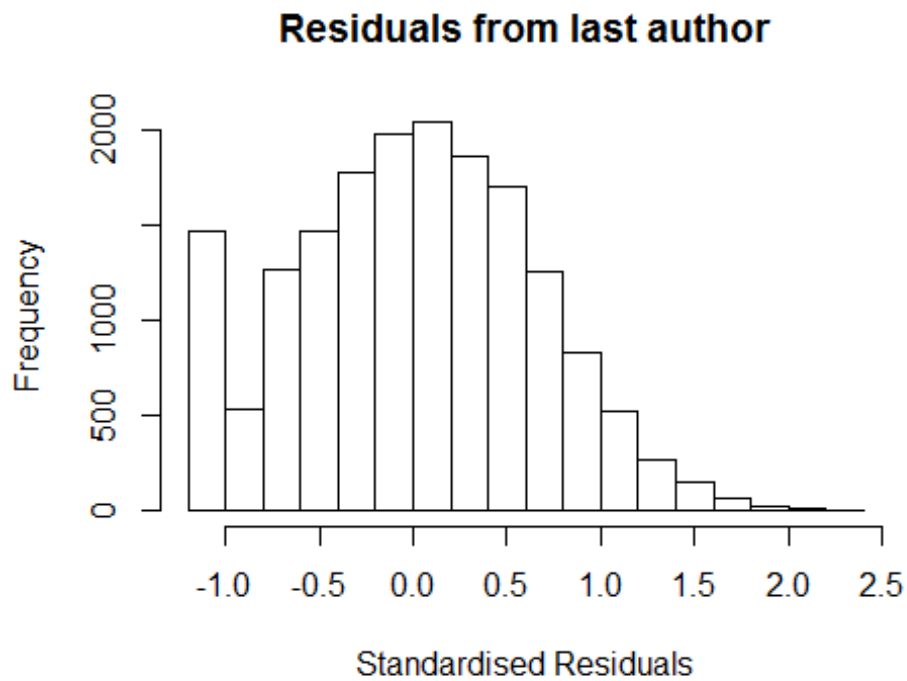

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.1886 -0.4538  0.0121  0.4564  2.2756
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)           1.0283     0.0218   47.06  < 2e-16 ***
## FirstAuthorFemale1    0.0604     0.0153    3.94  8.2e-05 ***
## LastAuthorFemale1     0.0151     0.0154    0.98   0.3278
## Year1997             -0.0364     0.0310   -1.18   0.2396
## Year1998             -0.0287     0.0314   -0.92   0.3602
## Year1999              0.0120     0.0299    0.40   0.6876
## Year2000             -0.0170     0.0307   -0.55   0.5795
## Year2001              0.0385     0.0310    1.24   0.2136
## Year2002              0.0198     0.0292    0.68   0.4983
## Year2003              0.0550     0.0299    1.84   0.0660 .
## Year2004              0.0597     0.0297    2.01   0.0442 *
## Year2005              0.0848     0.0297    2.85   0.0043 **
```

```
## Year2006             0.0735     0.0291    2.52   0.0116 *
## Year2007             0.0747     0.0297    2.51   0.0120 *
## Year2008             0.0655     0.0297    2.20   0.0276 *
## Year2009             0.0705     0.0297    2.37   0.0177 *
## Year2010             0.0736     0.0308    2.39   0.0171 *
## Year2011             0.0701     0.0297    2.36   0.0181 *
## Year2012             0.0348     0.0313    1.11   0.2661
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.671
## Multiple R-squared:  0.00498,    Adjusted R-squared:  0.00393
## Convergence in 10 IRWLS iterations
##
## Robustness weights:
##  1492 weights are ~= 1. The remaining 15719 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.226   0.877   0.949   0.915   0.985   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          5.81e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.003  1           1.001
## Year              1.003 16           1.000
```




```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.1818 -0.4550  0.0115  0.4579  2.2739
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)           1.0294     0.0218   47.22  < 2e-16 ***
## FirstAuthorFemale1    0.0674     0.0137    4.90  9.6e-07 ***
## Year1997             -0.0364     0.0310   -1.17   0.2402
## Year1998             -0.0286     0.0314   -0.91   0.3617
## Year1999              0.0120     0.0299    0.40   0.6874
## Year2000             -0.0171     0.0307   -0.56   0.5781
## Year2001              0.0386     0.0310    1.25   0.2128
## Year2002              0.0197     0.0292    0.68   0.4992
## Year2003              0.0551     0.0299    1.84   0.0657 .
## Year2004              0.0599     0.0297    2.02   0.0433 *
## Year2005              0.0850     0.0297    2.86   0.0042 **
## Year2006              0.0741     0.0291    2.54   0.0110 *
```

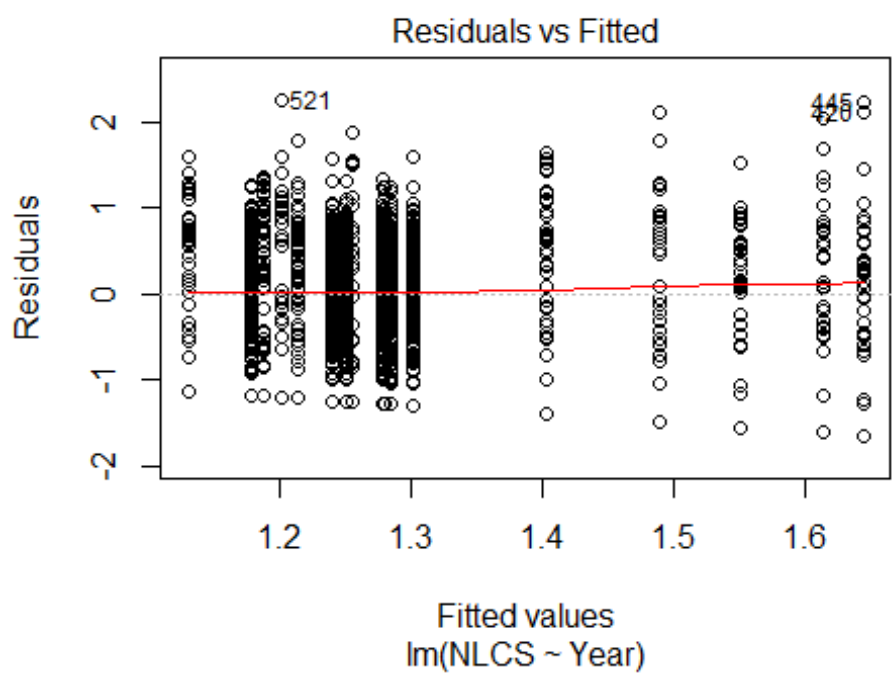
```
## Year2007             0.0752     0.0297    2.53   0.0115 *
## Year2008             0.0658     0.0297    2.21   0.0269 *
## Year2009             0.0710     0.0297    2.39   0.0169 *
## Year2010             0.0741     0.0308    2.40   0.0162 *
## Year2011             0.0706     0.0297    2.38   0.0173 *
## Year2012             0.0351     0.0312    1.12   0.2616
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.671
## Multiple R-squared:  0.00492,    Adjusted R-squared:  0.00393
## Convergence in 10 IRWLS iterations
##
## Robustness weights:
##  1490 weights are ~= 1. The remaining 15721 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.227   0.877   0.949   0.915   0.985   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          5.81e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##                   GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.005  1           1.003
## Year             1.005 16           1.000
```



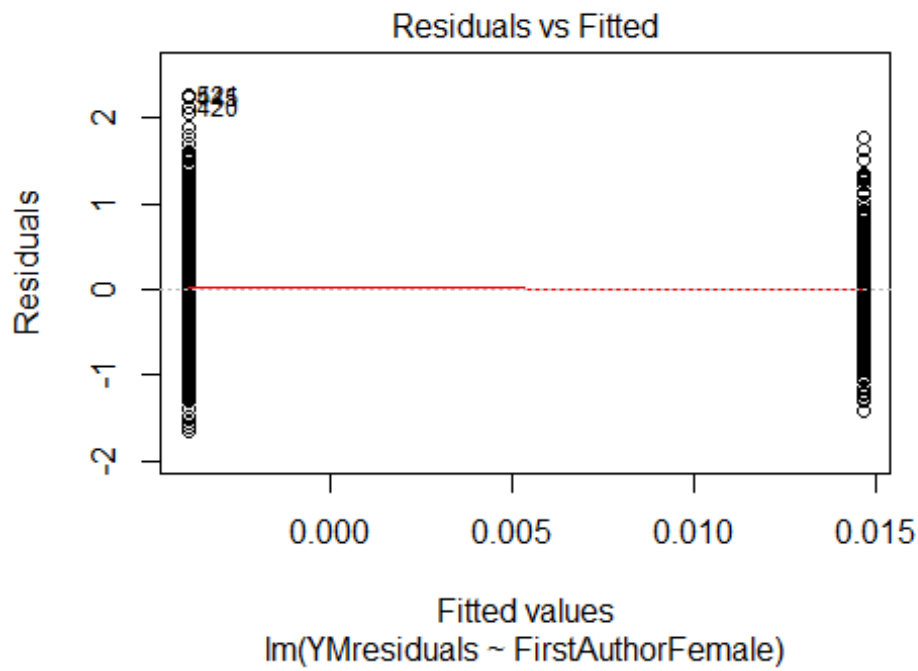

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.1620 -0.4564  0.0105  0.4545  2.2691
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)          1.0322     0.0219   47.22   <2e-16 ***
## LastAuthorFemale1    0.0430     0.0138    3.12   0.0018 **
## Year1997            -0.0355     0.0310   -1.15   0.2514
## Year1998            -0.0274     0.0314   -0.87   0.3838
## Year1999             0.0137     0.0299    0.46   0.6472
## Year2000            -0.0157     0.0307   -0.51   0.6087
## Year2001             0.0403     0.0310    1.30   0.1933
## Year2002             0.0209     0.0292    0.72   0.4738
## Year2003             0.0573     0.0300    1.91   0.0559 .
## Year2004             0.0608     0.0297    2.05   0.0406 *
## Year2005             0.0868     0.0298    2.92   0.0035 **
## Year2006             0.0745     0.0292    2.55   0.0107 *
```

```
## Year2007            0.0760     0.0297    2.56   0.0105 *
## Year2008            0.0674     0.0298    2.26   0.0236 *
## Year2009            0.0726     0.0298    2.44   0.0146 *
## Year2010            0.0753     0.0308    2.44   0.0147 *
## Year2011            0.0727     0.0297    2.45   0.0143 *
## Year2012            0.0363     0.0313    1.16   0.2467
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.671
## Multiple R-squared:  0.00404,    Adjusted R-squared:  0.00306
## Convergence in 10 IRWLS iterations
##
## Robustness weights:
##  1473 weights are ~= 1. The remaining 15738 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.229   0.877   0.949   0.915   0.985   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          5.81e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Sample size for the above analysis:  17211"
## [1] ""
## [1] ""
## [1] "####################################"
## [1] "Analysis of AJSC 2100"
## [1] "####################################"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##   66   51   75   68   76   83   70   71   97  116  148  784  778  785  918
## 2011 2012
##  951 1077
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##   41   33   42   40   51   41   37   46   67   79   95  461  455  457  506
## 2011 2012
```

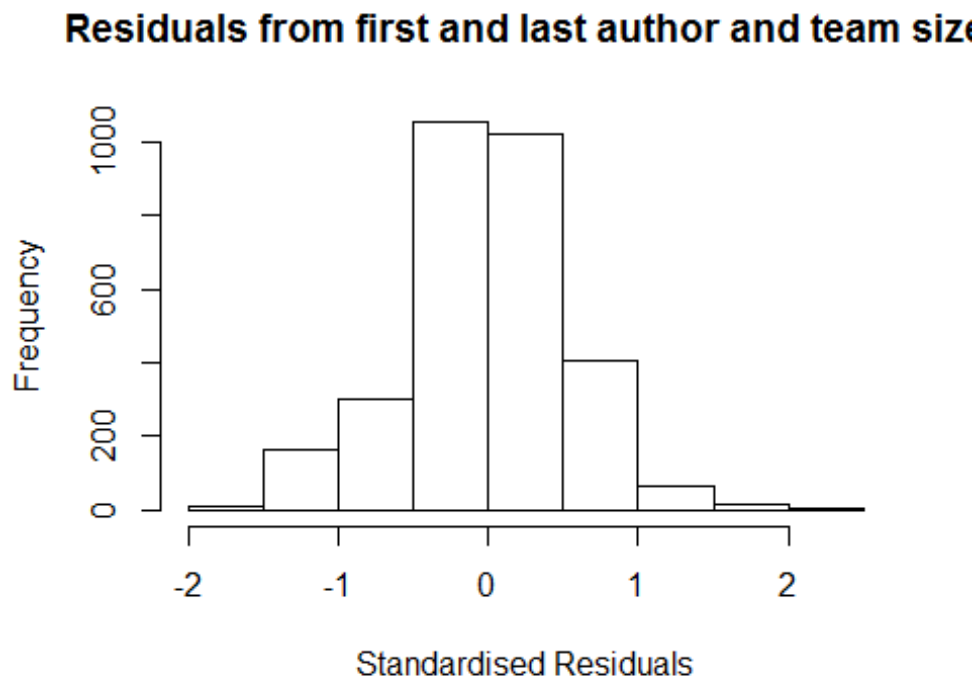
```
##  572  652
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##   38   27   38   35   47   36   34   42   63   69   83  382  363  358  419
## 2011 2012
##  466  539
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
##  Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 530, df = 16, p-value <2e-16
```




```
##
##  Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 40, df = 1, p-value = 3e-10
```




```
## [1] "Female first author team size 2018 geometric mean: 3.41378130061964"
## [1] "Male first author team size 2018 geometric mean: 3.5372741679497"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 26000, p-value = 0.5
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 3.23558142146045"
## [1] "Male last author team size 2018 geometric mean: 3.57101067128383"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 20000, p-value = 0.2
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.033  1           1.017
## LastAuthorFemale  1.046  1           1.023
## UniqueAuthors     1.281  4           1.031
## Year              1.321 16           1.009
```

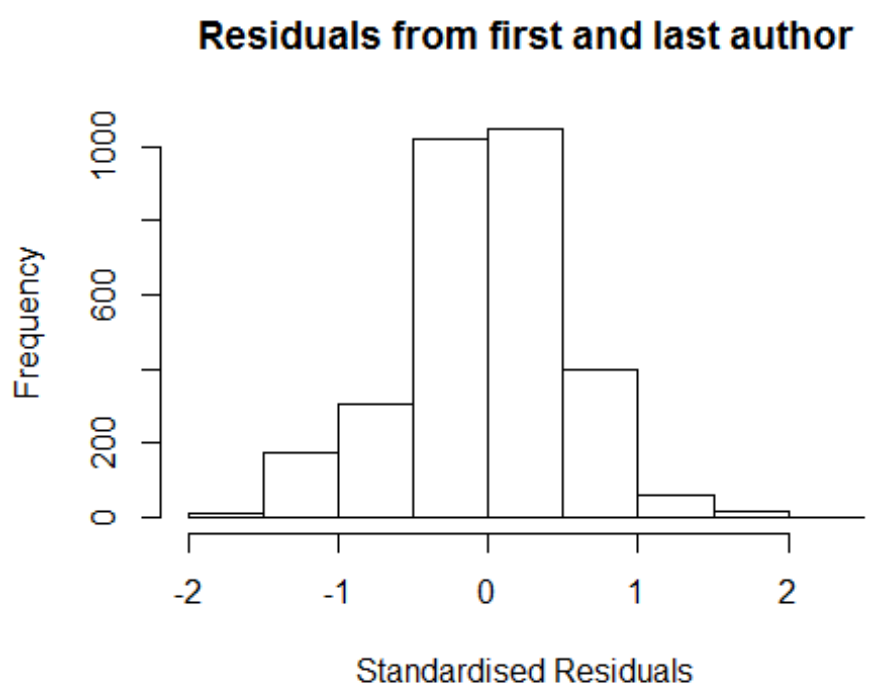



```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
UniqueAuthors +
##     Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##     k.max = 1000))
##  \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.83693 -0.31113 -0.00457  0.33101  2.33247
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)          1.6083     0.1610    9.99  < 2e-16 ***
## FirstAuthorFemale1  -0.0140     0.0214   -0.66  0.51244
## LastAuthorFemale1   -0.0110     0.0268   -0.41  0.68201
## UniqueAuthors2       0.2287     0.0426    5.37  8.7e-08 ***
## UniqueAuthors3       0.2580     0.0420    6.14  9.2e-10 ***
## UniqueAuthors4       0.2615     0.0420    6.22  5.5e-10 ***
## UniqueAuthors5       0.3219     0.0400    8.04  1.2e-15 ***
## Year1997            -0.8349     0.4355   -1.92  0.05533 .
## Year1998            -0.3365     0.2636   -1.28  0.20189
## Year1999            -0.2145     0.3279   -0.65  0.51300
```

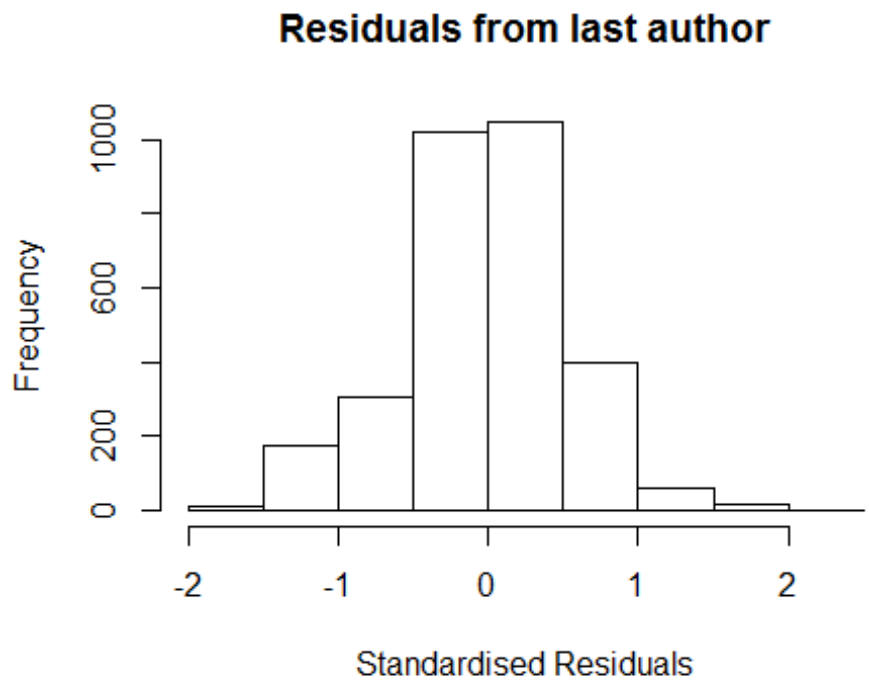
```
## Year2000            -0.5196     0.4157   -1.25  0.21136
## Year2001            -0.1686     0.2280   -0.74  0.45974
## Year2002            -0.7104     0.4227   -1.68  0.09295 .
## Year2003            -0.5320     0.3113   -1.71  0.08753 .
## Year2004            -0.4901     0.2407   -2.04  0.04184 *
## Year2005            -0.4783     0.2042   -2.34  0.01924 *
## Year2006            -0.3972     0.1784   -2.23  0.02603 *
## Year2007            -0.5985     0.1606   -3.73  0.00020 ***
## Year2008            -0.5499     0.1604   -3.43  0.00062 ***
## Year2009            -0.5497     0.1603   -3.43  0.00061 ***
## Year2010            -0.6576     0.1601   -4.11  4.1e-05 ***
## Year2011            -0.5927     0.1601   -3.70  0.00022 ***
## Year2012            -0.5387     0.1599   -3.37  0.00076 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.477
## Multiple R-squared:  0.0654, Adjusted R-squared:  0.0586
## Convergence in 42 IRWLS iterations
##
## Robustness weights:
##  2 observations c(87,240) are outliers with |weight| = 0 ( < 3.3e-05);
##  250 weights are ~= 1. The remaining 2787 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0296  0.8500  0.9510  0.8850  0.9870  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          3.29e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.023  1           1.012
## LastAuthorFemale  1.045  1           1.022
## Year              1.059 16           1.002

## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
```

```
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.73678 -0.31297  0.00255  0.33272  2.39149
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.73678    0.15181   11.44  < 2e-16 ***
## FirstAuthorFemale1  0.00578    0.02146    0.27  0.78783
## LastAuthorFemale1  -0.01100    0.02638   -0.42  0.67661
## Year1997           -0.88238    0.39630   -2.23  0.02605 *
## Year1998           -0.27047    0.27134   -1.00  0.31894
## Year1999           -0.14944    0.36208   -0.41  0.67983
## Year2000           -0.41850    0.40276   -1.04  0.29885
## Year2001           -0.16866    0.23019   -0.73  0.46379
## Year2002           -0.66927    0.61719   -1.08  0.27828
## Year2003           -0.58105    0.30626   -1.90  0.05789 .
## Year2004           -0.50087    0.25050   -2.00  0.04565 *
## Year2005           -0.46297    0.19751   -2.34  0.01914 *
## Year2006           -0.36264    0.17197   -2.11  0.03505 *
## Year2007           -0.49220    0.15311   -3.21  0.00132 **
## Year2008           -0.44433    0.15272   -2.91  0.00365 **
## Year2009           -0.44381    0.15273   -2.91  0.00369 **
## Year2010           -0.55927    0.15267   -3.66  0.00025 ***
## Year2011           -0.48798    0.15259   -3.20  0.00140 **
## Year2012           -0.42880    0.15235   -2.81  0.00492 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.48
## Multiple R-squared:  0.0306, Adjusted R-squared:  0.0249
## Convergence in 48 IRWLS iterations
##
## Robustness weights:
##  observation 240 is an outlier with |weight| = 0 ( < 3.3e-05);
##  269 weights are ~= 1. The remaining 2769 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0098  0.8470  0.9480  0.8820  0.9860  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          3.29e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
```

```
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                    psi           subsampling                   cov
##             "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                   "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"

## Warning in lmrob.fit(x, y, control, init = init, mf = mf): M-step did NOT
## converge. Returning unconverged SM-estimate
```



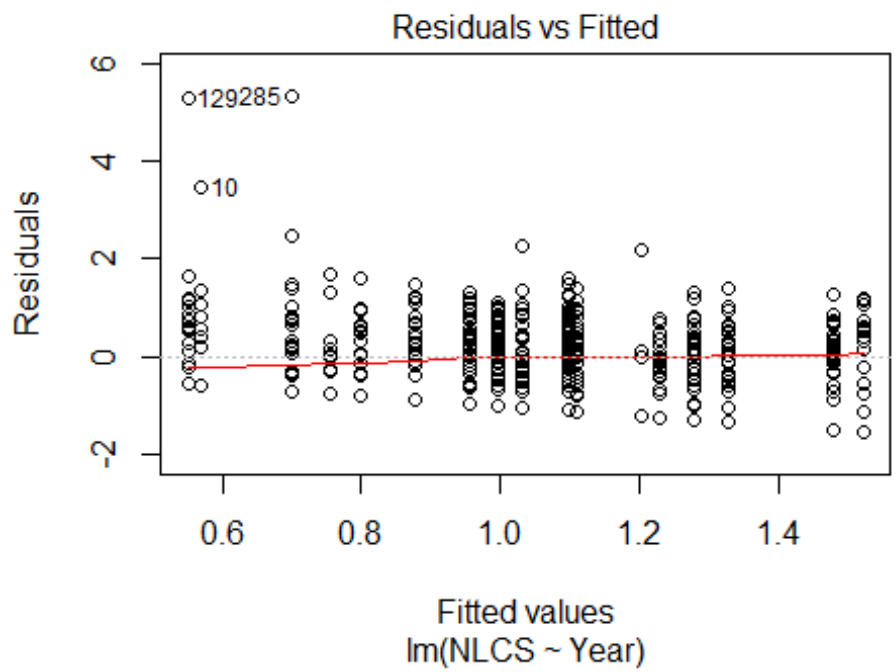

```
## [1] "Regression 4: Last author gender, Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.043  1           1.021
## Year             1.043 16           1.001
```



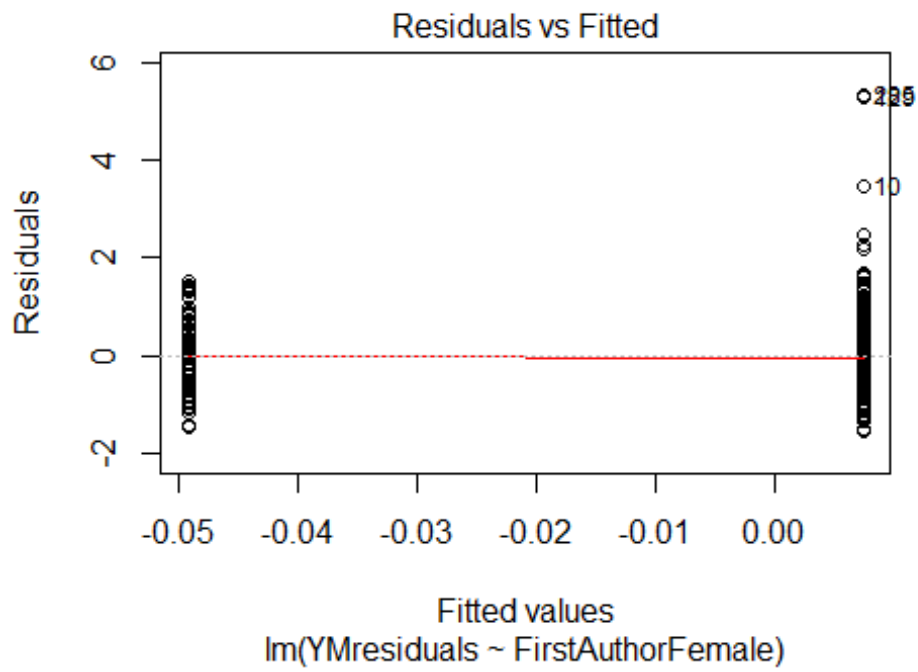

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.73764 -0.31413  0.00122  0.33365  2.39071
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.73764    0.15155   11.47  < 2e-16 ***
## LastAuthorFemale1 -0.00999    0.02643   -0.38  0.70545
## Year1997          -0.88220    0.39648   -2.23  0.02615 *
## Year1998          -0.27034    0.27235   -0.99  0.32097
## Year1999          -0.14954    0.36207   -0.41  0.67962
## Year2000          -0.41901    0.40210   -1.04  0.29747
## Year2001          -0.16901    0.23026   -0.73  0.46302
## Year2002          -0.66935    0.61652   -1.09  0.27770
## Year2003          -0.58114    0.30495   -1.91  0.05679 .
## Year2004          -0.50157    0.25051   -2.00  0.04535 *
## Year2005          -0.46307    0.19754   -2.34  0.01913 *
## Year2006          -0.36293    0.17193   -2.11  0.03486 *
```

```
## Year2007          -0.49176    0.15314   -3.21  0.00134 **
## Year2008          -0.44387    0.15275   -2.91  0.00369 **
## Year2009          -0.44352    0.15275   -2.90  0.00372 **
## Year2010          -0.55886    0.15270   -3.66  0.00026 ***
## Year2011          -0.48767    0.15261   -3.20  0.00141 **
## Year2012          -0.42833    0.15239   -2.81  0.00497 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.48
## Multiple R-squared:  0.0306, Adjusted R-squared:  0.0252
## Convergence in 41 IRWLS iterations
##
## Robustness weights:
##  observation 240 is an outlier with |weight| = 0 ( < 3.3e-05);
##  271 weights are ~= 1. The remaining 2767 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0099  0.8470  0.9490  0.8820  0.9860  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          3.29e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                    psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Sample size for the above analysis:  3039"
## [1] ""
## [1] ""
## [1] "####################################"
## [1] "Analysis of AJSC 2101"
## [1] "####################################"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##   37    8   22   32   61   66   66   38   52   54   76   82   54   42   85
## 2011 2012
##   74   98
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##   23    6   17   25   53   39   45   17   31   33   33   38   22   29   70
```

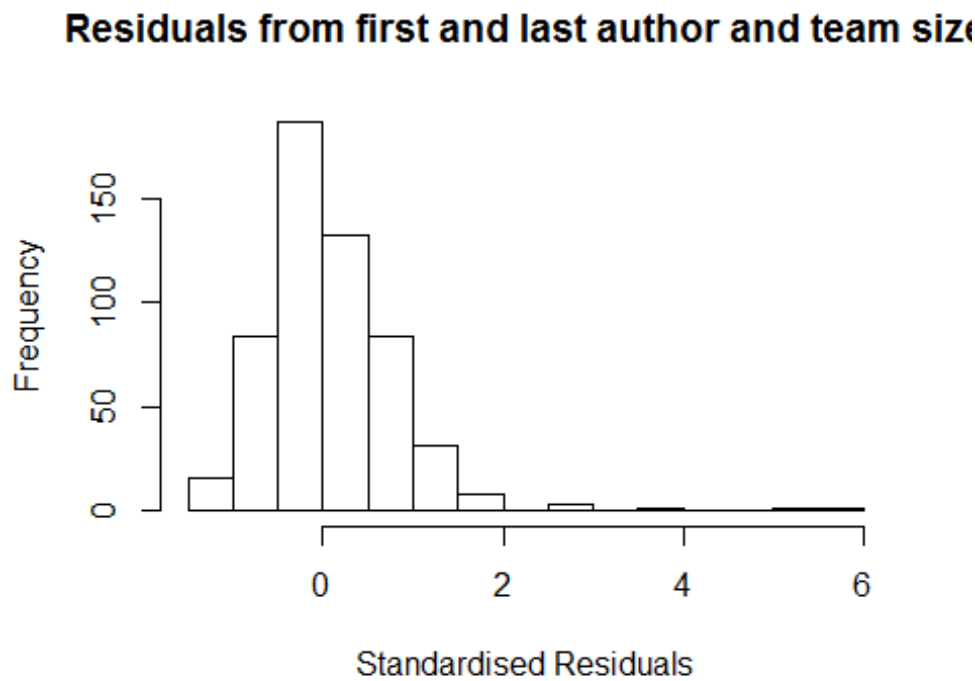
```
## 2011 2012
##   56   76
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##   22    6   17   21   53   39   43   14   28   23   30   28   19   23   63
## 2011 2012
##   50   68
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
##  Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 53, df = 16, p-value = 7e-06
```




```
##
##  Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 0.12, df = 1, p-value = 0.7
```



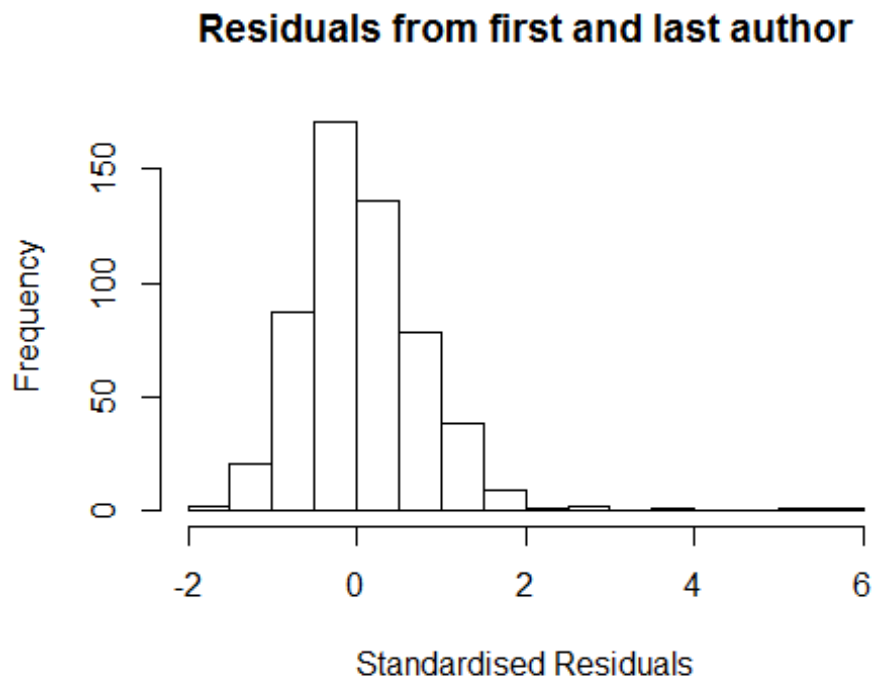

```
## [1] "Female first author team size 2018 geometric mean: 2.29635907184762"
## [1] "Male first author team size 2018 geometric mean: 2.2033904728449"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 1200, p-value = 0.8
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 2.19802298242095"
## [1] "Male last author team size 2018 geometric mean: 2.2302227762497"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 1300, p-value = 0.9
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.854  1           1.362
## LastAuthorFemale  1.877  1           1.370
## UniqueAuthors     2.401  4           1.116
## Year              2.962 16           1.035
```




```
## [1] "List of  6 outliers with residuals above  2.5"
##        ScopusId  NLCS Year OneField Fields residuals
## 10   0030406862 4.040 1996     1909      2     3.763
## 39   0031362657 3.396 1997     2101      1     2.749
## 129  0034197259 5.817 2000     2101      1     5.031
## 285 24944541480 6.021 2002     2101      1     5.682
## 286 25144462744 3.163 2002     2101      1     2.544
## 393  3142752164 3.312 2004     1603      4     2.590
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
UniqueAuthors +
##     Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##     k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.4101 -0.3539 -0.0408  0.4843  5.6818
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)          0.2773     0.1370    2.02   0.0435 *
## FirstAuthorFemale1  -0.0920     0.1072   -0.86   0.3914
## LastAuthorFemale1   -0.0461     0.1053   -0.44   0.6616
## UniqueAuthors2       0.2799     0.0886    3.16   0.0017 **
## UniqueAuthors3       0.4552     0.1040    4.38  1.5e-05 ***
```

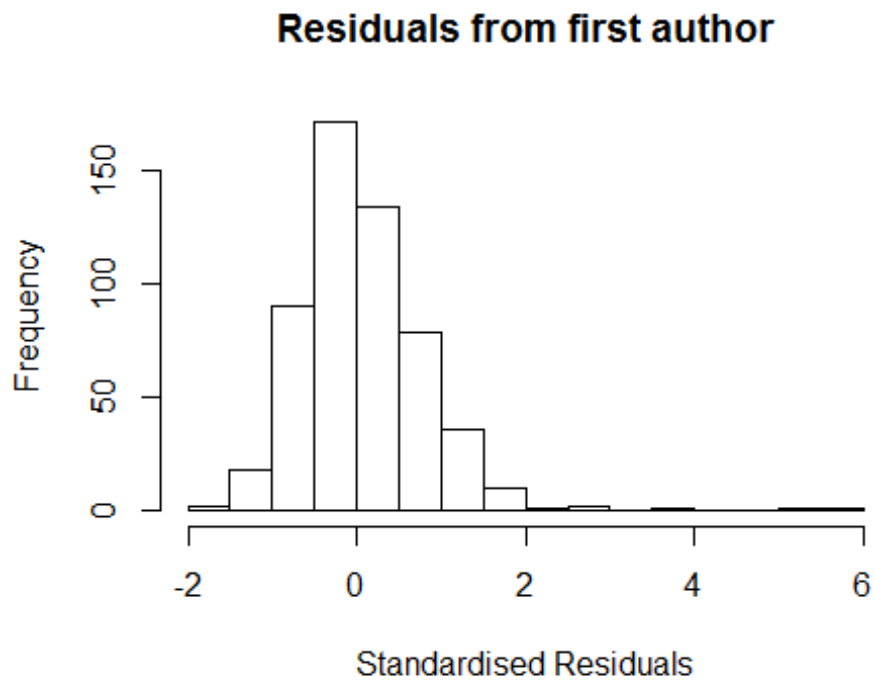
```
## UniqueAuthors4       0.7868     0.1084    7.26  1.4e-12 ***
## UniqueAuthors5       0.7634     0.0891    8.57  < 2e-16 ***
## Year1997             0.3696     0.3719    0.99   0.3208
## Year1998             0.1866     0.2252    0.83   0.4077
## Year1999             0.4338     0.2046    2.12   0.0344 *
## Year2000             0.0535     0.1587    0.34   0.7359
## Year2001             0.3915     0.1906    2.05   0.0405 *
## Year2002             0.0619     0.1679    0.37   0.7126
## Year2003             0.6706     0.2878    2.33   0.0202 *
## Year2004             0.4443     0.2111    2.10   0.0358 *
## Year2005             0.4886     0.1945    2.51   0.0123 *
## Year2006             0.5431     0.1891    2.87   0.0042 **
## Year2007             0.9319     0.1935    4.82  1.9e-06 ***
## Year2008             0.5270     0.1814    2.91   0.0038 **
## Year2009             0.5880     0.1814    3.24   0.0013 **
## Year2010             0.4881     0.1628    3.00   0.0028 **
## Year2011             0.3321     0.1675    1.98   0.0479 *
## Year2012             0.3461     0.1568    2.21   0.0278 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.601
## Multiple R-squared:  0.323,  Adjusted R-squared:  0.294
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
##  3 observations c(4,76,167) are outliers with |weight| = 0 ( < 0.00018);
##  35 weights are ~= 1. The remaining 509 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0022  0.8730  0.9500  0.8990  0.9810  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.83e-04          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##                     GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.631  1           1.277
```

```
## LastAuthorFemale  1.725  1           1.313
## Year              1.367 16           1.010
```



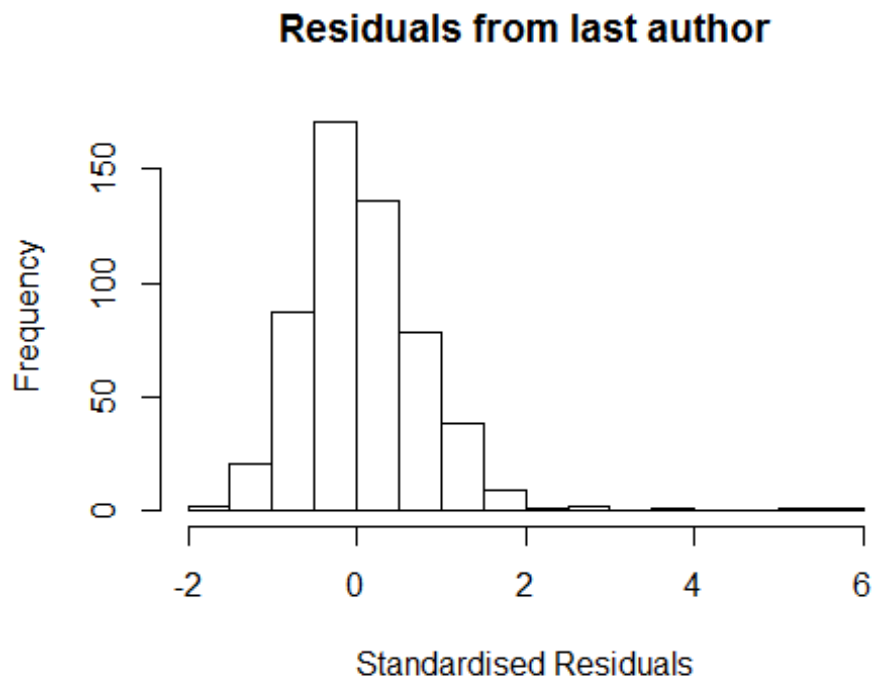

```
## [1] "List of  5 outliers with residuals above  2.5"
##        ScopusId  NLCS Year OneField Fields residuals
## 10   0030406862 4.040 1996     1909      2     3.679
## 39   0031362657 3.396 1997     2101      1     2.517
## 129  0034197259 5.817 2000     2101      1     5.437
## 285 24944541480 6.021 2002     2101      1     5.594
## 286 25144462744 3.163 2002     2101      1     2.736
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##    Min     1Q Median     3Q    Max
## -1.595 -0.384 -0.032  0.482  5.594
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         0.36088    0.14603    2.47   0.0138 *
## FirstAuthorFemale1  0.00865    0.11608    0.07   0.9406
## LastAuthorFemale1  -0.06473    0.11340   -0.57   0.5684
## Year1997            0.51776    0.45795    1.13   0.2587
## Year1998            0.25113    0.23475    1.07   0.2852
```

```
## Year1999            0.43890    0.20337    2.16   0.0314 *
## Year2000            0.01879    0.16719    0.11   0.9106
## Year2001            0.50246    0.18811    2.67   0.0078 **
## Year2002            0.06615    0.17816    0.37   0.7105
## Year2003            1.01917    0.35044    2.91   0.0038 **
## Year2004            0.62927    0.24750    2.54   0.0113 *
## Year2005            0.67675    0.21081    3.21   0.0014 **
## Year2006            0.91815    0.19588    4.69  3.5e-06 ***
## Year2007            1.22556    0.17465    7.02  7.0e-12 ***
## Year2008            0.87189    0.18757    4.65  4.2e-06 ***
## Year2009            0.88898    0.21072    4.22  2.9e-05 ***
## Year2010            0.71360    0.17532    4.07  5.4e-05 ***
## Year2011            0.56456    0.18573    3.04   0.0025 **
## Year2012            0.59780    0.16784    3.56   0.0004 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.661
## Multiple R-squared:  0.194,  Adjusted R-squared:  0.167
## Convergence in 15 IRWLS iterations
##
## Robustness weights:
##  3 observations c(4,76,167) are outliers with |weight| = 0 ( < 0.00018);
##  41 weights are ~= 1. The remaining 503 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0482  0.8500  0.9540  0.9010  0.9790  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.83e-04          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##                     GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.147  1           1.071
## Year              1.147 16           1.004
```




```
## [1] "List of  5 outliers with residuals above  2.5"
##        ScopusId  NLCS Year OneField Fields residuals
## 10   0030406862 4.040 1996     1909      2     3.679
## 39   0031362657 3.396 1997     2101      1     2.517
## 129  0034197259 5.817 2000     2101      1     5.437
## 285 24944541480 6.021 2002     2101      1     5.594
## 286 25144462744 3.163 2002     2101      1     2.736
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##    Min     1Q Median     3Q    Max
## -1.569 -0.399 -0.024  0.484  5.602
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)           0.3611     0.1467    2.46  0.01415 *
## FirstAuthorFemale1   -0.0199     0.0992   -0.20  0.84082
## Year1997              0.5026     0.4367    1.15  0.25034
## Year1998              0.2499     0.2350    1.06  0.28800
## Year1999              0.4333     0.2050    2.11  0.03503 *
## Year2000              0.0148     0.1676    0.09  0.92953
## Year2001              0.5025     0.1885    2.67  0.00793 **
## Year2002              0.0579     0.1785    0.32  0.74593
```

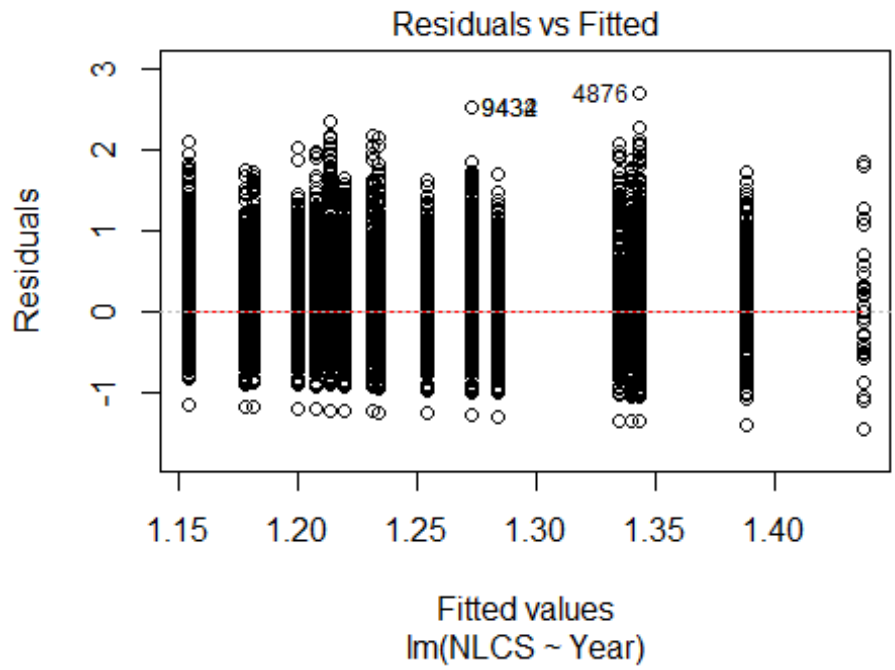
```
## Year2003             1.0209     0.3485    2.93  0.00354 **
## Year2004             0.6310     0.2485    2.54  0.01138 *
## Year2005             0.6629     0.2082    3.18  0.00154 **
## Year2006             0.9158     0.1970    4.65  4.2e-06 ***
## Year2007             1.2280     0.1750    7.02  7.0e-12 ***
## Year2008             0.8718     0.1881    4.63  4.5e-06 ***
## Year2009             0.8897     0.2115    4.21  3.0e-05 ***
## Year2010             0.7074     0.1768    4.00  7.2e-05 ***
## Year2011             0.5542     0.1839    3.01  0.00271 **
## Year2012             0.5942     0.1686    3.53  0.00046 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.659
## Multiple R-squared:  0.194,  Adjusted R-squared:  0.168
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
##  3 observations c(4,76,167) are outliers with |weight| = 0 ( < 0.00018);
##  42 weights are ~= 1. The remaining 502 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0443  0.8500  0.9540  0.9000  0.9790  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.83e-04          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##                   GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.228  1           1.108
## Year             1.228 16           1.006
```



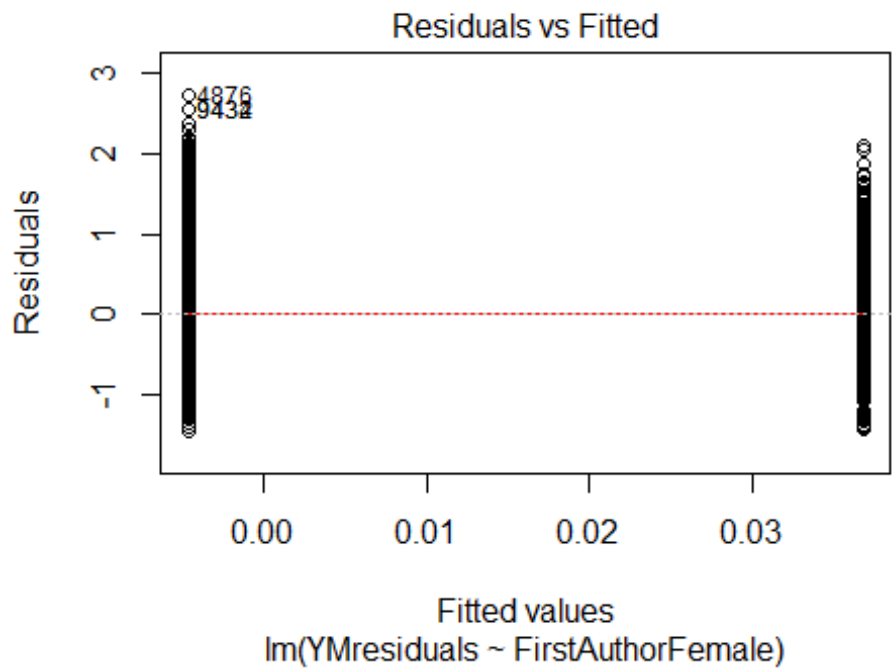

```
## [1] "List of  5 outliers with residuals above  2.5"
##        ScopusId  NLCS Year OneField Fields residuals
## 10   0030406862 4.040 1996     1909      2     3.679
## 39   0031362657 3.396 1997     2101      1     2.517
## 129  0034197259 5.817 2000     2101      1     5.437
## 285 24944541480 6.021 2002     2101      1     5.594
## 286 25144462744 3.163 2002     2101      1     2.736
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.5881 -0.3796 -0.0319  0.4824  5.5941
##
## Coefficients:
##                   Estimate Std. Error t value Pr(>|t|)
## (Intercept)         0.3604     0.1465    2.46  0.01417 *
## LastAuthorFemale1  -0.0599     0.0956   -0.63  0.53137
## Year1997            0.5159     0.4547    1.13  0.25708
## Year1998            0.2502     0.2348    1.07  0.28706
## Year1999            0.4390     0.2038    2.15  0.03168 *
## Year2000            0.0192     0.1669    0.11  0.90854
## Year2001            0.5035     0.1879    2.68  0.00760 **
## Year2002            0.0665     0.1774    0.37  0.70801
```

```
## Year2003            1.0206     0.3510    2.91  0.00379 **
## Year2004            0.6308     0.2469    2.55  0.01092 *
## Year2005            0.6768     0.2108    3.21  0.00140 **
## Year2006            0.9185     0.1964    4.68  3.7e-06 ***
## Year2007            1.2277     0.1751    7.01  7.3e-12 ***
## Year2008            0.8725     0.1879    4.64  4.4e-06 ***
## Year2009            0.8905     0.2094    4.25  2.5e-05 ***
## Year2010            0.7150     0.1747    4.09  4.9e-05 ***
## Year2011            0.5654     0.1859    3.04  0.00247 **
## Year2012            0.5988     0.1678    3.57  0.00039 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.658
## Multiple R-squared:  0.195,  Adjusted R-squared:  0.169
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
##  3 observations c(4,76,167) are outliers with |weight| = 0 ( < 0.00018);
##  41 weights are ~= 1. The remaining 503 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0454  0.8490  0.9540  0.9000  0.9790  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.83e-04          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Sample size for the above analysis:  547"
## [1] ""
## [1] ""
## [1] "####################################"
## [1] "Analysis of AJSC 2102"
## [1] "####################################"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 1330 1185 1020   88 1100  939  782  703  784  811 1002 1036 1085 1337 1359
## 2011 2012
```

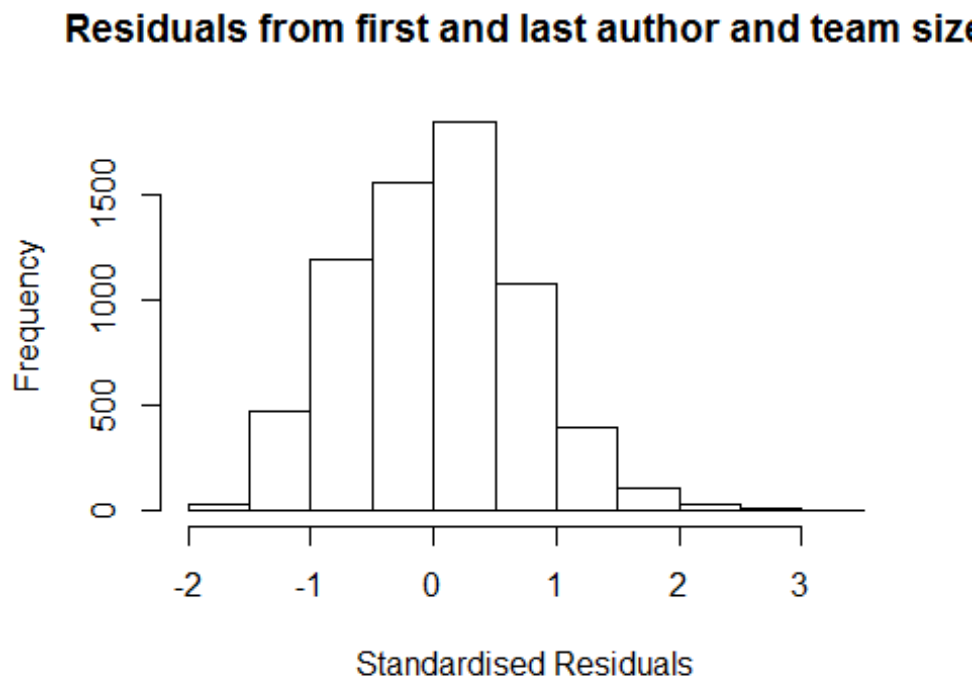
```
## 1446 1273
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  507  479  364   34  428  286  376  321  406  410  503  562  552  741  740
## 2011 2012
##  776  685
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  422  408  319   29  347  237  314  270  330  355  415  466  449  613  595
## 2011 2012
##  610  553
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
##  Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 210, df = 16, p-value <2e-16
```




```
##
##  Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 4.1, df = 1, p-value = 0.04
```

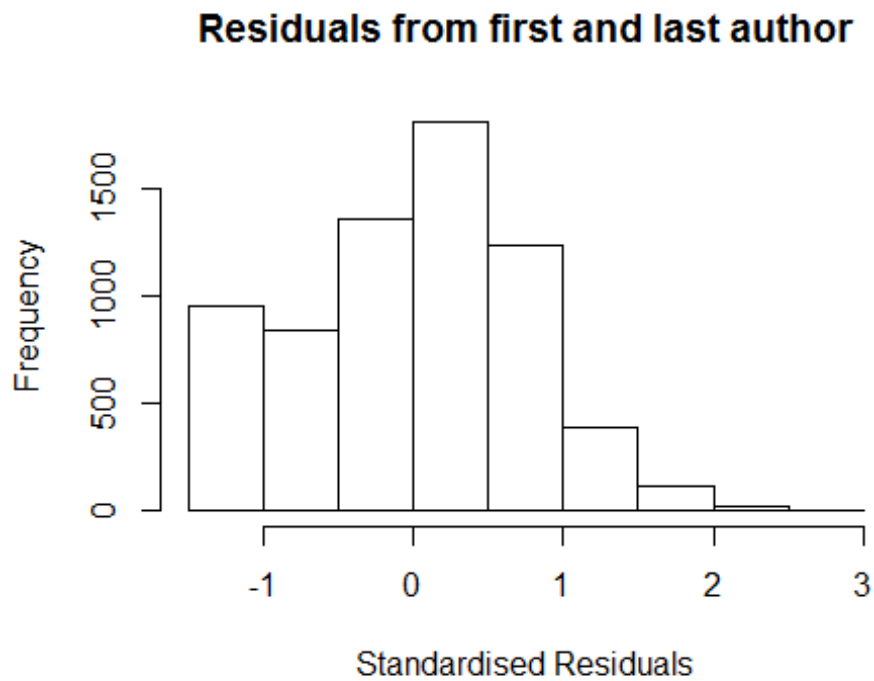



```
## [1] "Female first author team size 2018 geometric mean: 3.30208969448841"
## [1] "Male first author team size 2018 geometric mean: 3.12608924788705"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 36000, p-value = 0.2
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 3.29306124145018"
## [1] "Male last author team size 2018 geometric mean: 3.13144245208258"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 31000, p-value = 0.4
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.034  1           1.017
## LastAuthorFemale  1.030  1           1.015
## UniqueAuthors     1.076  4           1.009
## Year              1.083 16           1.002
```



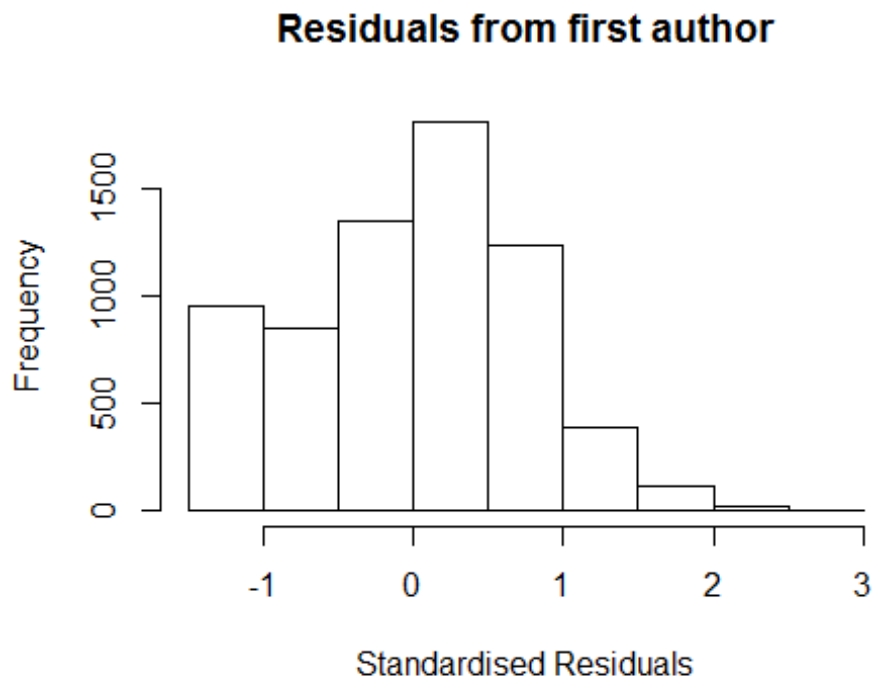

```
## [1] "List of  5 outliers with residuals above  2.5"
##         ScopusId  NLCS Year OneField Fields residuals
## 2337  0030720685 3.368 1997     1500      3     2.587
## 4876  0034193710 4.050 2000     1500      3     2.645
## 5880  0035340070 3.329 2001     2102      2     2.511
## 9432 18144385855 3.797 2005     1500      3     2.999
## 9434 18144418629 3.800 2005     1500      3     3.002
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
UniqueAuthors +
##     Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##     k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.6693 -0.5073  0.0342  0.4792  3.0016
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)          0.80445    0.04180   19.24  < 2e-16 ***
## FirstAuthorFemale1   0.01332    0.02701    0.49  0.62192
## LastAuthorFemale1   -0.01789    0.03249   -0.55  0.58201
## UniqueAuthors2       0.51569    0.02864   18.01  < 2e-16 ***
## UniqueAuthors3       0.61433    0.02952   20.81  < 2e-16 ***
## UniqueAuthors4       0.70261    0.03266   21.51  < 2e-16 ***
```

```
## UniqueAuthors5      0.71339    0.03179   22.44  < 2e-16 ***
## Year1997           -0.02312    0.05742   -0.40  0.68723
## Year1998           -0.13638    0.06302   -2.16  0.03049 *
## Year1999            0.12516    0.15302    0.82  0.41342
## Year2000            0.08444    0.06927    1.22  0.22293
## Year2001            0.01404    0.07402    0.19  0.84961
## Year2002            0.09176    0.06470    1.42  0.15621
## Year2003            0.13809    0.05963    2.32  0.02060 *
## Year2004           -0.08277    0.05602   -1.48  0.13959
## Year2005           -0.00609    0.05615   -0.11  0.91359
## Year2006           -0.04851    0.05157   -0.94  0.34693
## Year2007           -0.02109    0.04931   -0.43  0.66880
## Year2008           -0.12836    0.04924   -2.61  0.00915 **
## Year2009           -0.12964    0.04619   -2.81  0.00502 **
## Year2010           -0.10290    0.04492   -2.29  0.02200 *
## Year2011           -0.16075    0.04564   -3.52  0.00043 ***
## Year2012           -0.14403    0.04585   -3.14  0.00169 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.673
## Multiple R-squared:  0.134,  Adjusted R-squared:  0.132
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
##  512 weights are ~= 1. The remaining 6220 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0089  0.8660  0.9420  0.8980  0.9830  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.49e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.011  1           1.005
## LastAuthorFemale  1.009  1           1.005
## Year              1.015 16           1.000
```



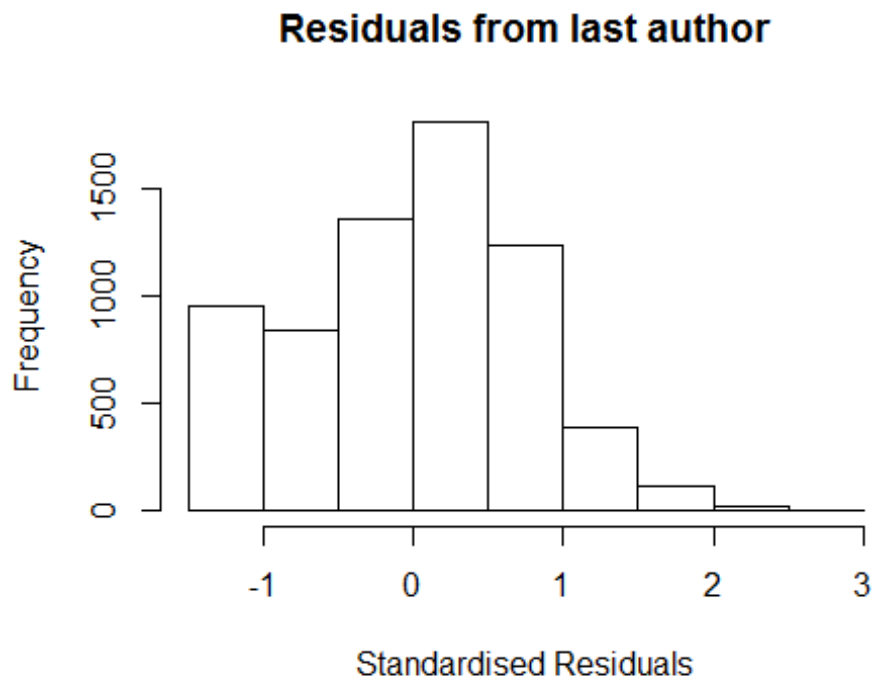

```
## [1] "List of  3 outliers with residuals above  2.5"
##         ScopusId  NLCS Year OneField Fields residuals
## 4876  0034193710 4.050 2000     1500      3     2.744
## 9432 18144385855 3.797 2005     1500      3     2.548
## 9434 18144418629 3.800 2005     1500      3     2.551
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.4598 -0.5657  0.0735  0.5237  2.7443
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.20221    0.03952   30.42   <2e-16 ***
## FirstAuthorFemale1  0.06901    0.02857    2.42   0.0157 *
## LastAuthorFemale1  -0.01483    0.03508   -0.42   0.6725
## Year1997           -0.03371    0.06116   -0.55   0.5815
## Year1998           -0.12117    0.06425   -1.89   0.0594 .
## Year1999            0.19808    0.15398    1.29   0.1984
## Year2000            0.10348    0.06779    1.53   0.1269
## Year2001            0.04170    0.07771    0.54   0.5915
## Year2002            0.14346    0.06765    2.12   0.0340 *
## Year2003            0.18860    0.06249    3.02   0.0026 **
```

```
## Year2004           -0.04173    0.06075   -0.69   0.4922
## Year2005            0.04689    0.06270    0.75   0.4546
## Year2006            0.03196    0.05490    0.58   0.5605
## Year2007            0.04712    0.05341    0.88   0.3777
## Year2008           -0.04469    0.05265   -0.85   0.3960
## Year2009           -0.05007    0.04949   -1.01   0.3117
## Year2010           -0.00791    0.04799   -0.16   0.8691
## Year2011           -0.06550    0.04884   -1.34   0.1799
## Year2012           -0.03798    0.04805   -0.79   0.4293
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.769
## Multiple R-squared:  0.00926,    Adjusted R-squared:  0.0066
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
##  532 weights are ~= 1. The remaining 6200 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.176   0.861   0.946   0.911   0.984   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.49e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.008  1           1.004
## Year              1.008 16           1.000
```




```
## [1] "List of  3 outliers with residuals above  2.5"
##         ScopusId  NLCS Year OneField Fields residuals
## 4876  0034193710 4.050 2000     1500      3     2.744
## 9432 18144385855 3.797 2005     1500      3     2.548
## 9434 18144418629 3.800 2005     1500      3     2.551
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.4563 -0.5690  0.0738  0.5234  2.7455
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.20157    0.03948   30.43   <2e-16 ***
## FirstAuthorFemale1  0.06643    0.02897    2.29   0.0219 *
## Year1997           -0.03411    0.06116   -0.56   0.5770
## Year1998           -0.12090    0.06424   -1.88   0.0599 .
## Year1999            0.19692    0.15354    1.28   0.1997
## Year2000            0.10292    0.06774    1.52   0.1288
## Year2001            0.04118    0.07772    0.53   0.5962
## Year2002            0.14303    0.06762    2.12   0.0345 *
## Year2003            0.18826    0.06245    3.01   0.0026 **
## Year2004           -0.04232    0.06071   -0.70   0.4858
```

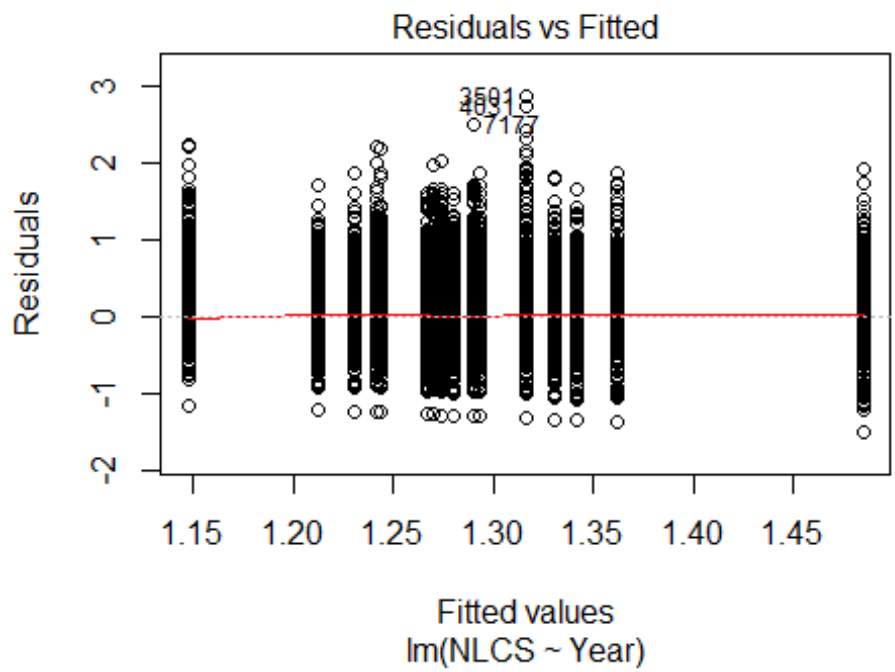
```
## Year2005            0.04643    0.06274    0.74   0.4592
## Year2006            0.03138    0.05490    0.57   0.5676
## Year2007            0.04687    0.05341    0.88   0.3802
## Year2008           -0.04493    0.05265   -0.85   0.3935
## Year2009           -0.05021    0.04948   -1.01   0.3102
## Year2010           -0.00849    0.04794   -0.18   0.8594
## Year2011           -0.06539    0.04884   -1.34   0.1807
## Year2012           -0.03834    0.04804   -0.80   0.4248
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.769
## Multiple R-squared:  0.00924,    Adjusted R-squared:  0.00673
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
##  529 weights are ~= 1. The remaining 6203 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.176   0.861   0.946   0.911   0.984   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.49e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##                   GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.006  1           1.003
## Year             1.006 16           1.000
```



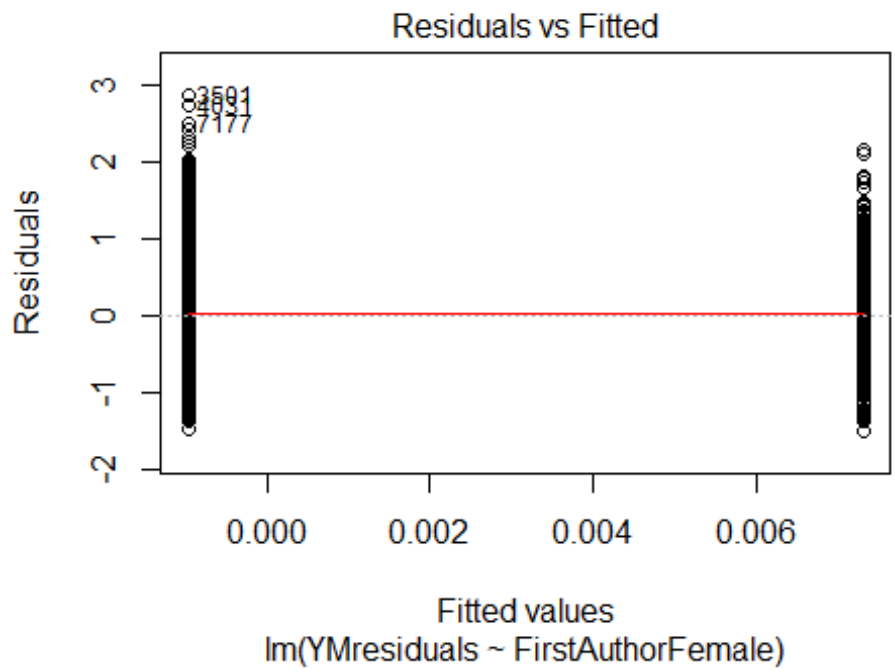

```
## [1] "List of  3 outliers with residuals above  2.5"
##         ScopusId  NLCS Year OneField Fields residuals
## 4876  0034193710 4.050 2000     1500      3     2.744
## 9432 18144385855 3.797 2005     1500      3     2.548
## 9434 18144418629 3.800 2005     1500      3     2.551
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.4007 -0.5625  0.0704  0.5246  2.7381
##
## Coefficients:
##                   Estimate Std. Error t value Pr(>|t|)
## (Intercept)        1.20783    0.03953   30.56   <2e-16 ***
## LastAuthorFemale1  0.00139    0.03492    0.04   0.9681
## Year1997          -0.03318    0.06125   -0.54   0.5880
## Year1998          -0.11932    0.06437   -1.85   0.0638 .
## Year1999           0.19291    0.15495    1.25   0.2132
## Year2000           0.10411    0.06786    1.53   0.1250
## Year2001           0.04156    0.07765    0.54   0.5925
## Year2002           0.14212    0.06763    2.10   0.0356 *
## Year2003           0.18931    0.06251    3.03   0.0025 **
## Year2004          -0.04224    0.06085   -0.69   0.4877
```

```
## Year2005           0.04665    0.06272    0.74   0.4570
## Year2006           0.03367    0.05482    0.61   0.5391
## Year2007           0.04864    0.05349    0.91   0.3633
## Year2008          -0.04319    0.05277   -0.82   0.4131
## Year2009          -0.04732    0.04957   -0.95   0.3398
## Year2010          -0.00717    0.04806   -0.15   0.8814
## Year2011          -0.06364    0.04887   -1.30   0.1929
## Year2012          -0.03618    0.04810   -0.75   0.4519
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.769
## Multiple R-squared:  0.00849,    Adjusted R-squared:  0.00597
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
##  535 weights are ~= 1. The remaining 6197 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.178   0.861   0.946   0.911   0.984   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.49e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                  psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Sample size for the above analysis:  6732"
## [1] ""
## [1] ""
## [1] "##################################"
## [1] "Analysis of AJSC 2103"
## [1] "##################################"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  962  766  749  725  709  679  531  550  581  627  664  748  744  891  908
## 2011 2012
##  901  863
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
```

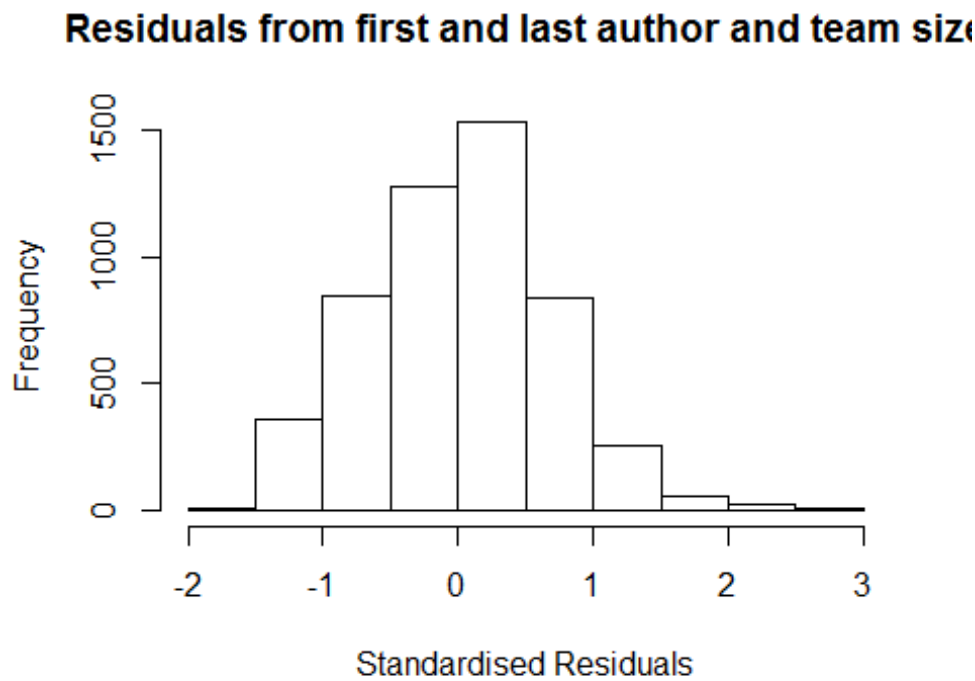
```
##  403  338  277  247  333  186  249  264  315  339  368  432  396  527  529
## 2011 2012
##  511  504
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  344  284  246  208  278  158  212  213  268  282  305  351  325  446  433
## 2011 2012
##  420  418
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
##  Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 190, df = 16, p-value <2e-16
```




```
##
##  Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 0.25, df = 1, p-value = 0.6
```

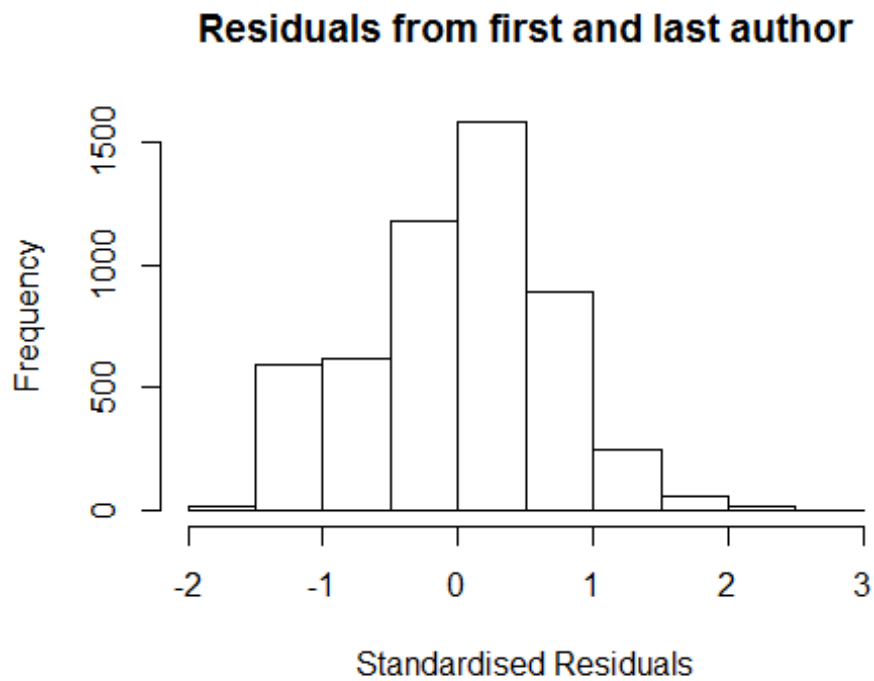



```
## [1] "Female first author team size 2018 geometric mean: 3.25621107011926"
## [1] "Male first author team size 2018 geometric mean: 3.24305795721261"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 20000, p-value = 0.8
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 3.08391865131447"
## [1] "Male last author team size 2018 geometric mean: 3.27712458939744"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 16000, p-value = 0.6
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.028  1           1.014
## LastAuthorFemale  1.017  1           1.008
## UniqueAuthors     1.106  4           1.013
## Year              1.129 16           1.004
```



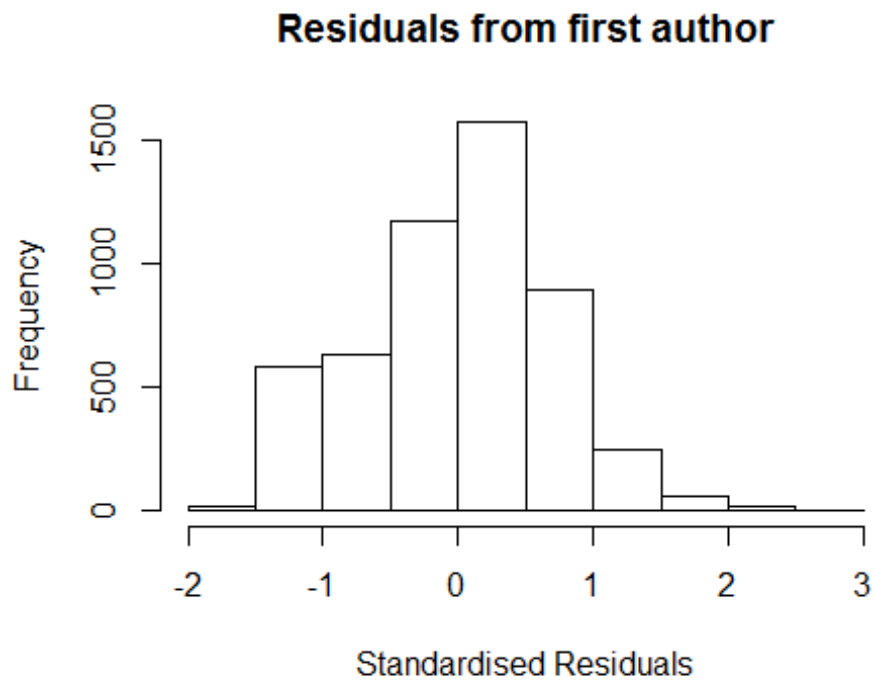

```
## [1] "List of  4 outliers with residuals above  2.5"
##         ScopusId  NLCS Year OneField Fields residuals
## 1503  0030720685 3.368 1997     1500      3     2.574
## 4031  0034193710 4.050 2000     1500      3     2.771
## 7175 18144385855 3.797 2005     1500      3     2.881
## 7177 18144418629 3.800 2005     1500      3     2.884
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
UniqueAuthors +
##     Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##     k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.7025 -0.4509  0.0379  0.4519  2.8836
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)          0.89294    0.04624   19.31  < 2e-16 ***
## FirstAuthorFemale1 -0.00195    0.02969   -0.07   0.9476
## LastAuthorFemale1   0.03678    0.03242    1.13   0.2567
## UniqueAuthors2      0.38040    0.03506   10.85  < 2e-16 ***
## UniqueAuthors3      0.43675    0.03570   12.23  < 2e-16 ***
## UniqueAuthors4      0.51937    0.03833   13.55  < 2e-16 ***
## UniqueAuthors5      0.53729    0.03871   13.88  < 2e-16 ***
```

```
## Year1997           -0.09848    0.06100   -1.61   0.1065
## Year1998           -0.01591    0.06521   -0.24   0.8073
## Year1999            0.04777    0.06566    0.73   0.4670
## Year2000            0.00517    0.07024    0.07   0.9413
## Year2001           -0.03452    0.08759   -0.39   0.6935
## Year2002            0.18580    0.06603    2.81   0.0049 **
## Year2003            0.27223    0.06447    4.22  2.5e-05 ***
## Year2004           -0.00249    0.06224   -0.04   0.9680
## Year2005            0.02349    0.06035    0.39   0.6971
## Year2006            0.01408    0.05501    0.26   0.7981
## Year2007            0.06628    0.05173    1.28   0.2001
## Year2008            0.04019    0.05070    0.79   0.4280
## Year2009           -0.01687    0.04849   -0.35   0.7280
## Year2010           -0.02963    0.04761   -0.62   0.5337
## Year2011           -0.09260    0.04936   -1.88   0.0607 .
## Year2012           -0.09129    0.04742   -1.93   0.0542 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.656
## Multiple R-squared:  0.0806, Adjusted R-squared:  0.0767
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
##  398 weights are ~= 1. The remaining 4793 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0142  0.8560  0.9480  0.9020  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.93e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.018  1           1.009
## LastAuthorFemale  1.006  1           1.003
## Year              1.024 16           1.001
```



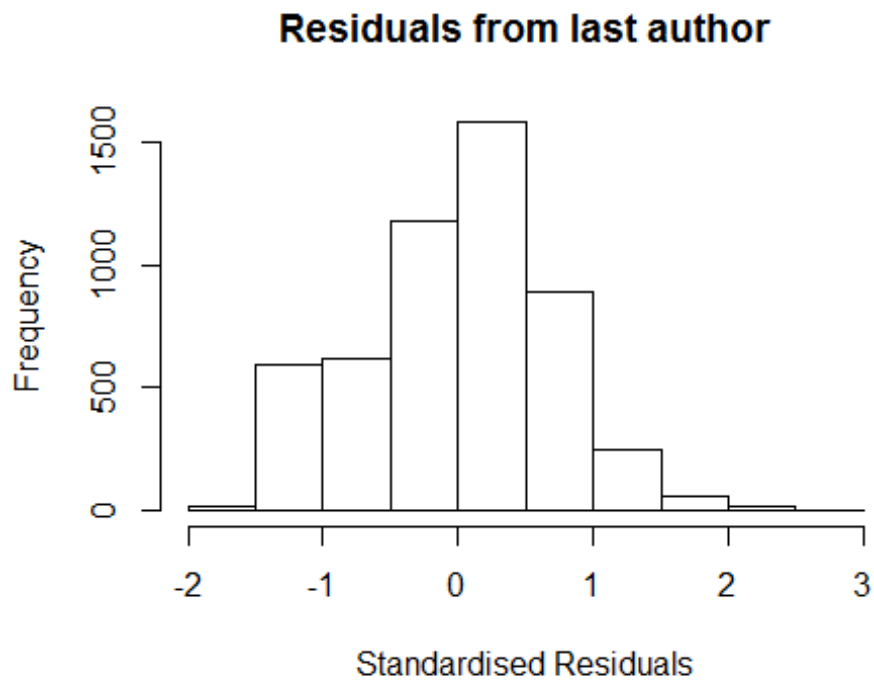

```
## [1] "List of  3 outliers with residuals above  2.5"
##         ScopusId  NLCS Year OneField Fields residuals
## 4031  0034193710 4.050 2000     1500      3     2.844
## 7175 18144385855 3.797 2005     1500      3     2.534
## 7177 18144418629 3.800 2005     1500      3     2.537
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.5276 -0.4588  0.0531  0.4677  2.8441
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)          1.2198     0.0402   30.36  < 2e-16 ***
## FirstAuthorFemale1   0.0242     0.0316    0.77   0.4433
## LastAuthorFemale1    0.0341     0.0351    0.97   0.3310
## Year1997            -0.1241     0.0643   -1.93   0.0538 .
## Year1998            -0.0223     0.0672   -0.33   0.7398
## Year1999             0.0843     0.0692    1.22   0.2233
## Year2000            -0.0140     0.0707   -0.20   0.8433
## Year2001            -0.0185     0.0889   -0.21   0.8349
## Year2002             0.2171     0.0670    3.24   0.0012 **
## Year2003             0.2835     0.0665    4.26  2.1e-05 ***
```

```
## Year2004             0.0375     0.0642    0.58   0.5590
## Year2005             0.0436     0.0664    0.66   0.5118
## Year2006             0.0521     0.0583    0.89   0.3710
## Year2007             0.1139     0.0547    2.08   0.0373 *
## Year2008             0.1006     0.0532    1.89   0.0587 .
## Year2009             0.0530     0.0503    1.05   0.2919
## Year2010             0.0444     0.0492    0.90   0.3665
## Year2011            -0.0171     0.0518   -0.33   0.7409
## Year2012            -0.0218     0.0489   -0.45   0.6550
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.676
## Multiple R-squared:  0.0143, Adjusted R-squared:  0.0109
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
##  453 weights are ~= 1. The remaining 4738 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.037   0.855   0.947   0.900   0.984   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.93e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.018  1           1.009
## Year              1.018 16           1.001
```




```
## [1] "List of  3 outliers with residuals above  2.5"
##         ScopusId  NLCS Year OneField Fields residuals
## 4031  0034193710 4.050 2000     1500      3     2.844
## 7175 18144385855 3.797 2005     1500      3     2.534
## 7177 18144418629 3.800 2005     1500      3     2.537
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.5352 -0.4607  0.0517  0.4658  2.8416
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)          1.2216     0.0401   30.43   <2e-16 ***
## FirstAuthorFemale1   0.0292     0.0318    0.92   0.3597
## Year1997            -0.1229     0.0643   -1.91   0.0560 .
## Year1998            -0.0228     0.0673   -0.34   0.7346
## Year1999             0.0865     0.0693    1.25   0.2120
## Year2000            -0.0132     0.0708   -0.19   0.8526
## Year2001            -0.0176     0.0889   -0.20   0.8434
## Year2002             0.2178     0.0671    3.25   0.0012 **
## Year2003             0.2844     0.0666    4.27    2e-05 ***
## Year2004             0.0396     0.0642    0.62   0.5374
```

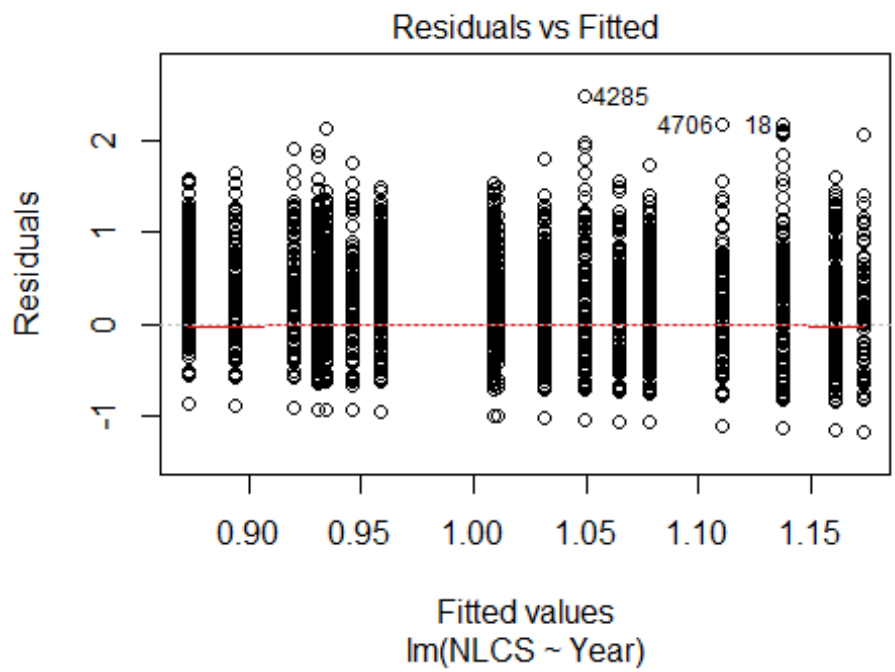
```
## Year2005             0.0452     0.0663    0.68   0.4957
## Year2006             0.0535     0.0583    0.92   0.3586
## Year2007             0.1147     0.0547    2.10   0.0362 *
## Year2008             0.1008     0.0532    1.89   0.0583 .
## Year2009             0.0537     0.0503    1.07   0.2857
## Year2010             0.0452     0.0492    0.92   0.3578
## Year2011            -0.0172     0.0518   -0.33   0.7395
## Year2012            -0.0213     0.0489   -0.44   0.6630
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.675
## Multiple R-squared:  0.0141, Adjusted R-squared:  0.0109
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
##  465 weights are ~= 1. The remaining 4726 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0374  0.8540  0.9470  0.9000  0.9840  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.93e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##                   GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.006  1           1.003
## Year             1.006 16           1.000
```



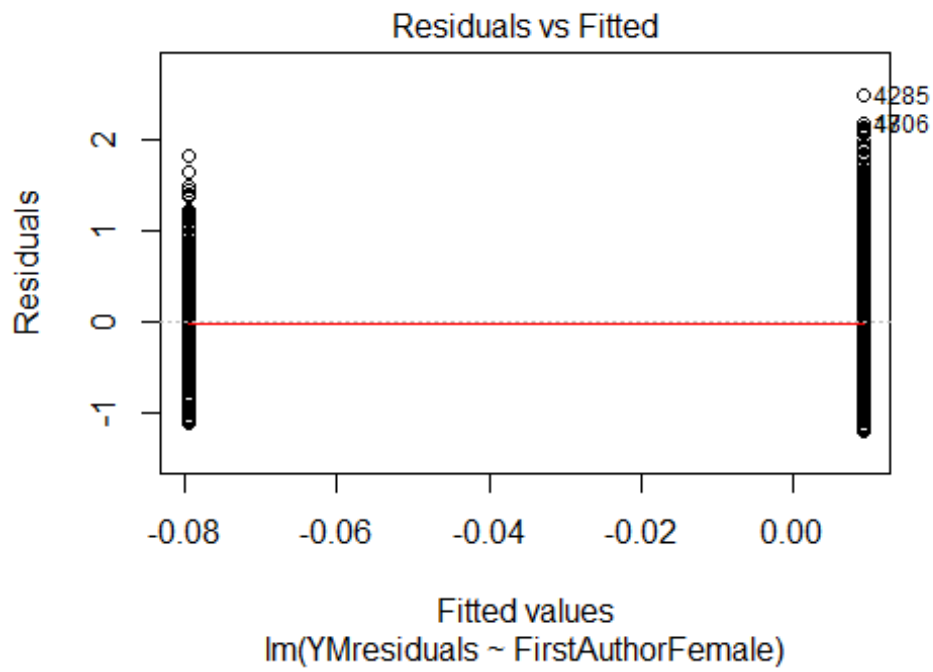

```
## [1] "List of  3 outliers with residuals above  2.5"
##         ScopusId  NLCS Year OneField Fields residuals
## 4031  0034193710 4.050 2000     1500      3     2.844
## 7175 18144385855 3.797 2005     1500      3     2.534
## 7177 18144418629 3.800 2005     1500      3     2.537
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.5059 -0.4582  0.0518  0.4680  2.8419
##
## Coefficients:
##                   Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.2215     0.0402   30.42   <2e-16 ***
## LastAuthorFemale1   0.0387     0.0355    1.09   0.2759
## Year1997           -0.1232     0.0644   -1.91   0.0557 .
## Year1998           -0.0217     0.0673   -0.32   0.7473
## Year1999            0.0851     0.0692    1.23   0.2190
## Year2000           -0.0135     0.0708   -0.19   0.8491
## Year2001           -0.0188     0.0888   -0.21   0.8321
## Year2002            0.2166     0.0671    3.23   0.0012 **
## Year2003            0.2844     0.0665    4.27    2e-05 ***
## Year2004            0.0392     0.0641    0.61   0.5408
```

```
## Year2005            0.0442     0.0665    0.67   0.5061
## Year2006            0.0530     0.0583    0.91   0.3628
## Year2007            0.1151     0.0547    2.10   0.0355 *
## Year2008            0.1020     0.0532    1.92   0.0551 .
## Year2009            0.0550     0.0503    1.09   0.2740
## Year2010            0.0449     0.0492    0.91   0.3614
## Year2011           -0.0162     0.0518   -0.31   0.7540
## Year2012           -0.0213     0.0489   -0.44   0.6633
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.676
## Multiple R-squared:  0.0142, Adjusted R-squared:  0.0109
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
##  461 weights are ~= 1. The remaining 4730 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0378  0.8540  0.9480  0.9000  0.9840  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.93e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                  psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Sample size for the above analysis:  5191"
## [1] ""
## [1] ""
## [1] "##################################"
## [1] "Analysis of AJSC 2104"
## [1] "##################################"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  735  698  652  558  498  384  457  391  366  593  589  517  639  933  655
## 2011 2012
##  634  729
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
```

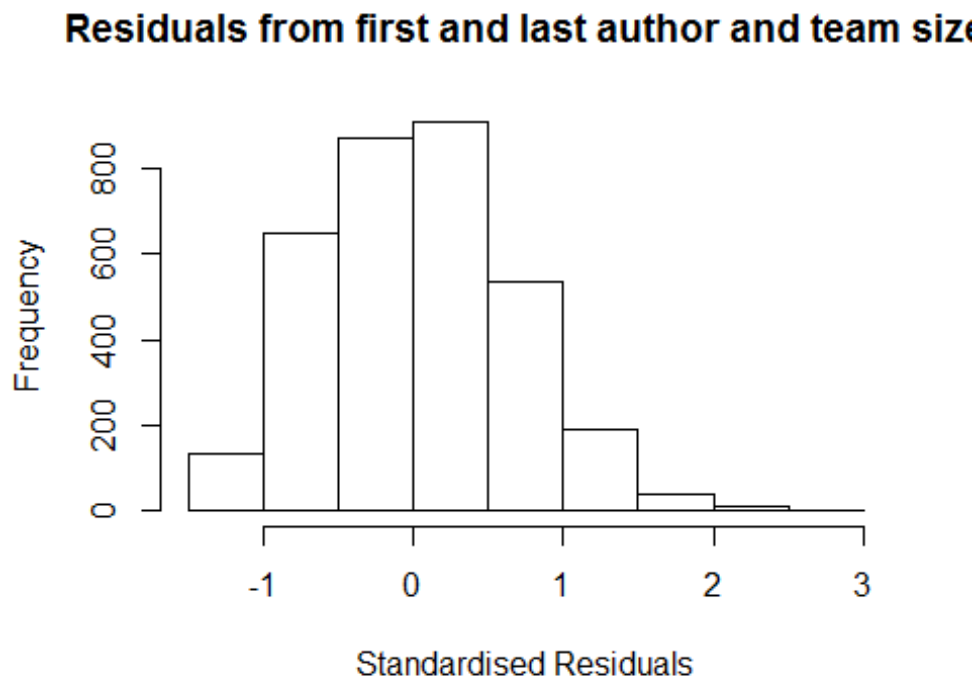
```
##  214  194  154  117  148   88  185  135  171  251  281  230  276  437  355
## 2011 2012
##  334  399
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  183  159  137   92  122   80  150  113  141  210  237  202  239  362  305
## 2011 2012
##  273  329
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
##  Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 28, df = 16, p-value = 0.03
```




```
##
##  Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 0.86, df = 1, p-value = 0.4
```

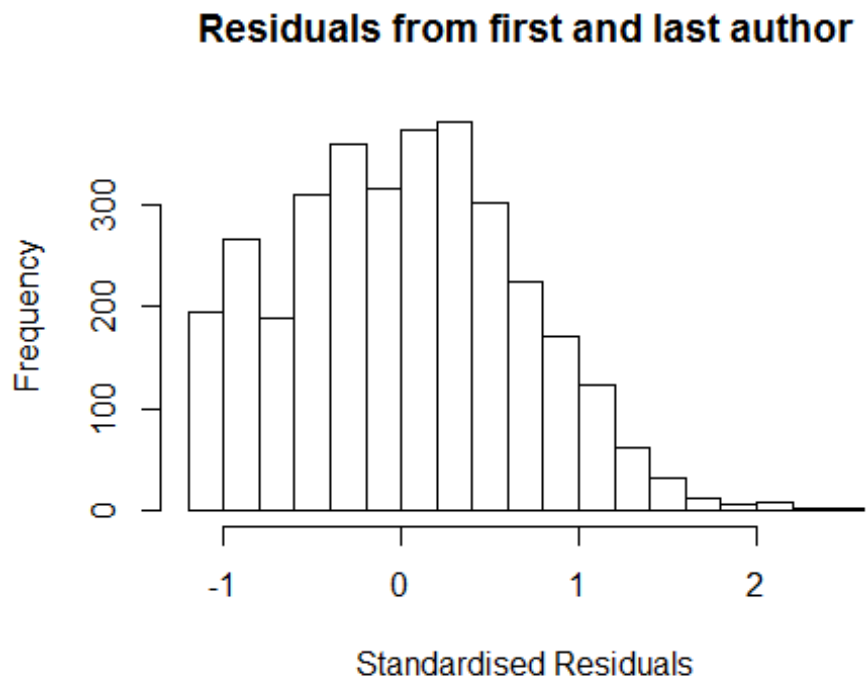



```
## [1] "Female first author team size 2018 geometric mean: 3.07906165106233"
## [1] "Male first author team size 2018 geometric mean: 3.56878510302242"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 9900, p-value = 0.2
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 3.22020807207117"
## [1] "Male last author team size 2018 geometric mean: 3.52717671156147"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 9500, p-value = 0.6
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.091  1           1.044
## LastAuthorFemale  1.072  1           1.035
## UniqueAuthors     1.118  4           1.014
## Year              1.144 16           1.004
```



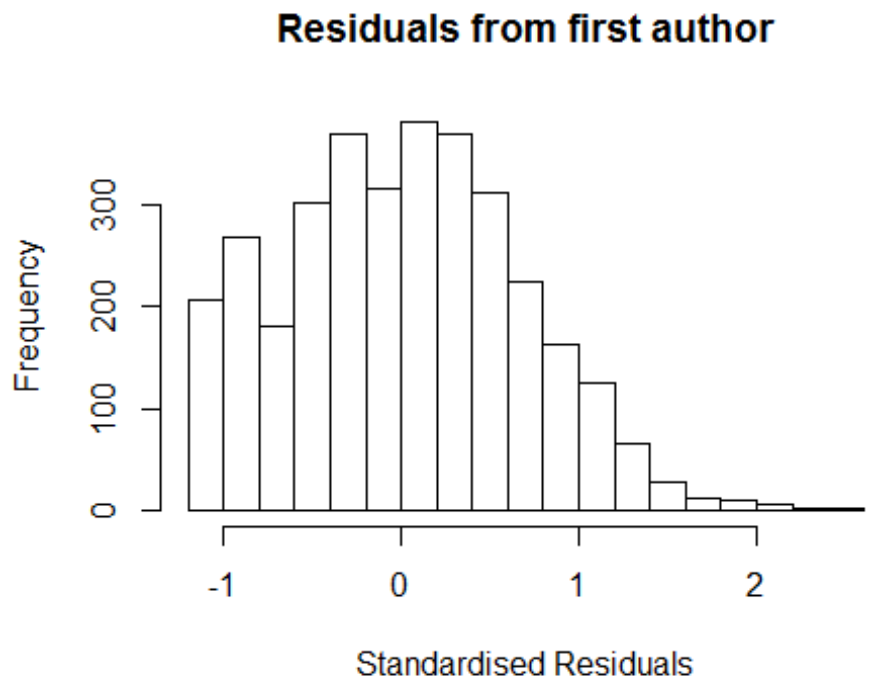

```
## [1] "List of  1 outliers with residuals above  2.5"
##        ScopusId  NLCS Year OneField Fields residuals
## 4285 0036591982 3.537 2002     1500      5     2.685
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
UniqueAuthors +
##     Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##     k.max = 1000))
##  \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.36644 -0.46786  0.00652  0.45976  2.68524
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         0.98928    0.05767   17.15  < 2e-16 ***
## FirstAuthorFemale1 -0.11220    0.03772   -2.97  0.00296 **
## LastAuthorFemale1  -0.17920    0.04291   -4.18  3.0e-05 ***
## UniqueAuthors2      0.14549    0.03501    4.16  3.3e-05 ***
## UniqueAuthors3      0.21084    0.03796    5.55  3.0e-08 ***
## UniqueAuthors4      0.30617    0.04110    7.45  1.2e-13 ***
## UniqueAuthors5      0.38100    0.03771   10.10  < 2e-16 ***
## Year1997           -0.10493    0.07641   -1.37  0.16978
## Year1998           -0.19421    0.08596   -2.26  0.02393 *
## Year1999           -0.10901    0.09157   -1.19  0.23398
```

```
## Year2000           -0.24270    0.08740   -2.78  0.00552 **
## Year2001            0.00507    0.09776    0.05  0.95867
## Year2002           -0.13752    0.08052   -1.71  0.08772 .
## Year2003           -0.00384    0.09124   -0.04  0.96641
## Year2004           -0.06553    0.08121   -0.81  0.41977
## Year2005           -0.29829    0.06809   -4.38  1.2e-05 ***
## Year2006           -0.31869    0.06914   -4.61  4.2e-06 ***
## Year2007           -0.25523    0.07209   -3.54  0.00040 ***
## Year2008           -0.10985    0.06650   -1.65  0.09865 .
## Year2009           -0.27236    0.06210   -4.39  1.2e-05 ***
## Year2010           -0.21609    0.06481   -3.33  0.00087 ***
## Year2011           -0.16678    0.06665   -2.50  0.01238 *
## Year2012           -0.11933    0.06551   -1.82  0.06862 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.676
## Multiple R-squared:  0.0612, Adjusted R-squared:  0.055
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  276 weights are ~= 1. The remaining 3058 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0789  0.8710  0.9490  0.9150  0.9840  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          3.00e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.070  1           1.035
## LastAuthorFemale  1.056  1           1.028
## Year              1.035 16           1.001
```



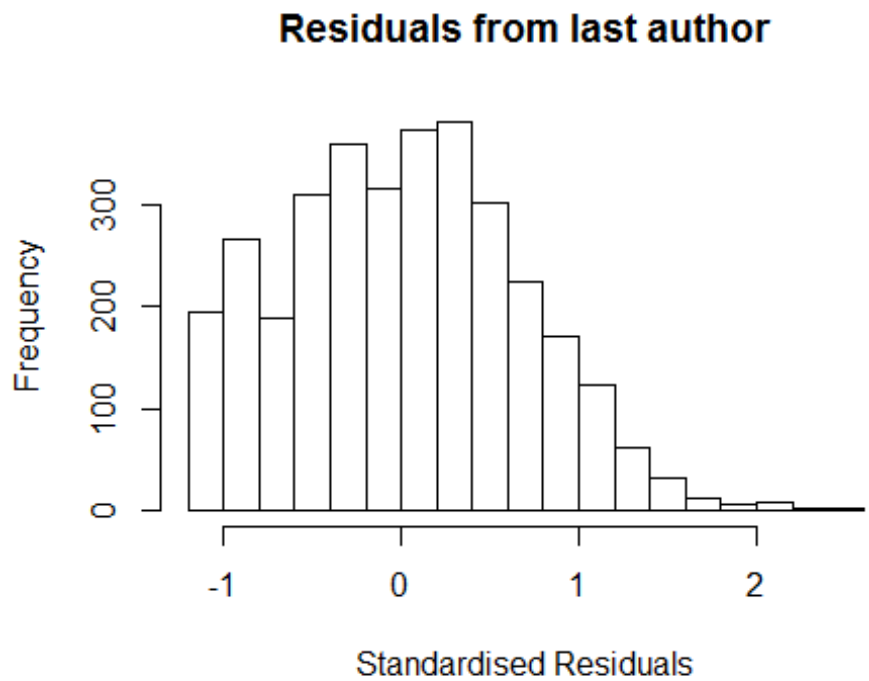

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.1506 -0.4918  0.0173  0.4622  2.4812
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.12677    0.05512   20.44  < 2e-16 ***
## FirstAuthorFemale1 -0.08464    0.03879   -2.18  0.02918 *
## LastAuthorFemale1  -0.17672    0.04425   -3.99  6.7e-05 ***
## Year1997           -0.08359    0.07806   -1.07  0.28430
## Year1998           -0.20033    0.08832   -2.27  0.02337 *
## Year1999           -0.09771    0.09089   -1.08  0.28244
## Year2000           -0.21471    0.08679   -2.47  0.01341 *
## Year2001            0.00813    0.09869    0.08  0.93438
## Year2002           -0.07093    0.08110   -0.87  0.38187
## Year2003            0.02381    0.09119    0.26  0.79403
## Year2004            0.00566    0.08383    0.07  0.94617
## Year2005           -0.25857    0.07065   -3.66  0.00026 ***
```

```
## Year2006           -0.26619    0.07153   -3.72  0.00020 ***
## Year2007           -0.22316    0.07463   -2.99  0.00281 **
## Year2008           -0.08502    0.06837   -1.24  0.21378
## Year2009           -0.20426    0.06360   -3.21  0.00133 **
## Year2010           -0.14774    0.06702   -2.20  0.02756 *
## Year2011           -0.08999    0.06909   -1.30  0.19280
## Year2012           -0.02545    0.06646   -0.38  0.70174
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.701
## Multiple R-squared:  0.0259, Adjusted R-squared:  0.0206
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  264 weights are ~= 1. The remaining 3070 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.185   0.868   0.950   0.918   0.985   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          3.00e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.024  1           1.012
## Year              1.024 16           1.001
```




```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.1331 -0.5034  0.0174  0.4637  2.4922
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.11586    0.05484   20.35  < 2e-16 ***
## FirstAuthorFemale1 -0.12114    0.03900   -3.11  0.00191 **
## Year1997           -0.08367    0.07843   -1.07  0.28617
## Year1998           -0.19102    0.08834   -2.16  0.03067 *
## Year1999           -0.09728    0.09169   -1.06  0.28876
## Year2000           -0.20444    0.08648   -2.36  0.01813 *
## Year2001            0.00885    0.09786    0.09  0.92798
## Year2002           -0.07102    0.08124   -0.87  0.38209
## Year2003            0.01722    0.09154    0.19  0.85079
## Year2004            0.00489    0.08388    0.06  0.95354
## Year2005           -0.25849    0.07046   -3.67  0.00025 ***
## Year2006           -0.26257    0.07090   -3.70  0.00022 ***
```

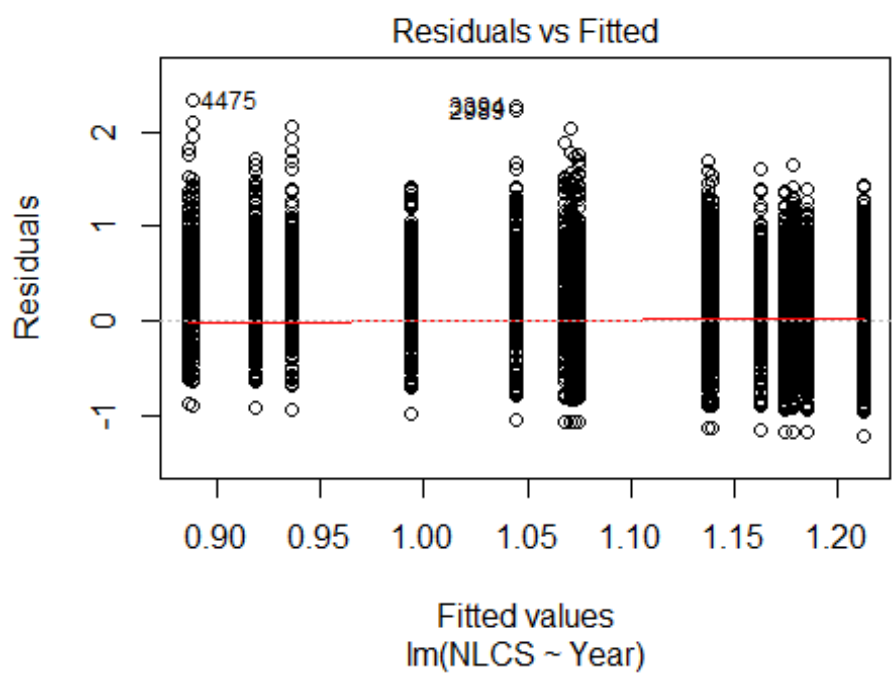
```
## Year2007           -0.21988    0.07475   -2.94  0.00329 **
## Year2008           -0.08486    0.06837   -1.24  0.21464
## Year2009           -0.20487    0.06345   -3.23  0.00126 **
## Year2010           -0.14968    0.06685   -2.24  0.02521 *
## Year2011           -0.09260    0.06896   -1.34  0.17943
## Year2012           -0.02426    0.06643   -0.37  0.71496
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.704
## Multiple R-squared:  0.0211, Adjusted R-squared:  0.0161
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  270 weights are ~= 1. The remaining 3064 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.184   0.870   0.950   0.918   0.985   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          3.00e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##                   GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.009  1           1.004
## Year             1.009 16           1.000
```



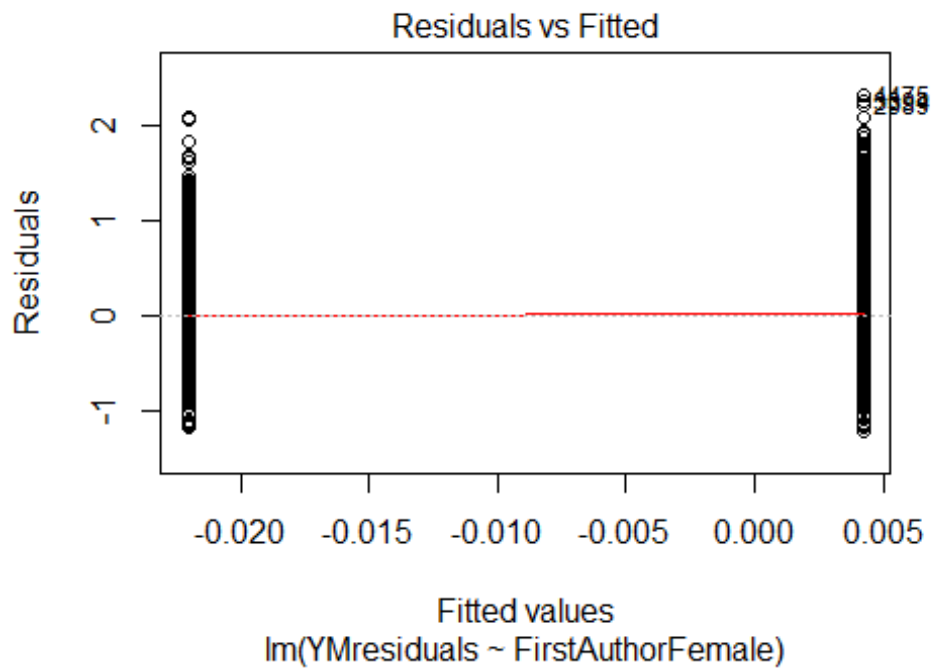

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.1410 -0.5106  0.0175  0.4628  2.4873
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.12306    0.05537   20.28  < 2e-16 ***
## LastAuthorFemale1 -0.19859    0.04386   -4.53  6.2e-06 ***
## Year1997          -0.08252    0.07830   -1.05  0.29200
## Year1998          -0.20268    0.08851   -2.29  0.02209 *
## Year1999          -0.10123    0.09109   -1.11  0.26650
## Year2000          -0.21944    0.08704   -2.52  0.01174 *
## Year2001           0.00511    0.09895    0.05  0.95886
## Year2002          -0.07340    0.08110   -0.91  0.36550
## Year2003           0.01792    0.09158    0.20  0.84490
## Year2004           0.00715    0.08400    0.09  0.93215
## Year2005          -0.26265    0.07087   -3.71  0.00021 ***
## Year2006          -0.27344    0.07174   -3.81  0.00014 ***
```

```
## Year2007          -0.22582    0.07487   -3.02  0.00258 **
## Year2008          -0.08474    0.06868   -1.23  0.21734
## Year2009          -0.20867    0.06397   -3.26  0.00112 **
## Year2010          -0.15277    0.06723   -2.27  0.02314 *
## Year2011          -0.09391    0.06937   -1.35  0.17588
## Year2012          -0.02890    0.06657   -0.43  0.66423
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.701
## Multiple R-squared:  0.0246, Adjusted R-squared:  0.0196
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  259 weights are ~= 1. The remaining 3075 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.182   0.871   0.949   0.918   0.985   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          3.00e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Sample size for the above analysis:  3334"
## [1] ""
## [1] ""
## [1] "####################################"
## [1] "Analysis of AJSC 2105"
## [1] "####################################"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  672  764  650  587  564  580  511  567  578  546  708  780  908 1080 1366
## 2011 2012
## 1661 1622
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  315  357  293  222  275  272  280  290  297  270  339  392  501  618  750
## 2011 2012
```

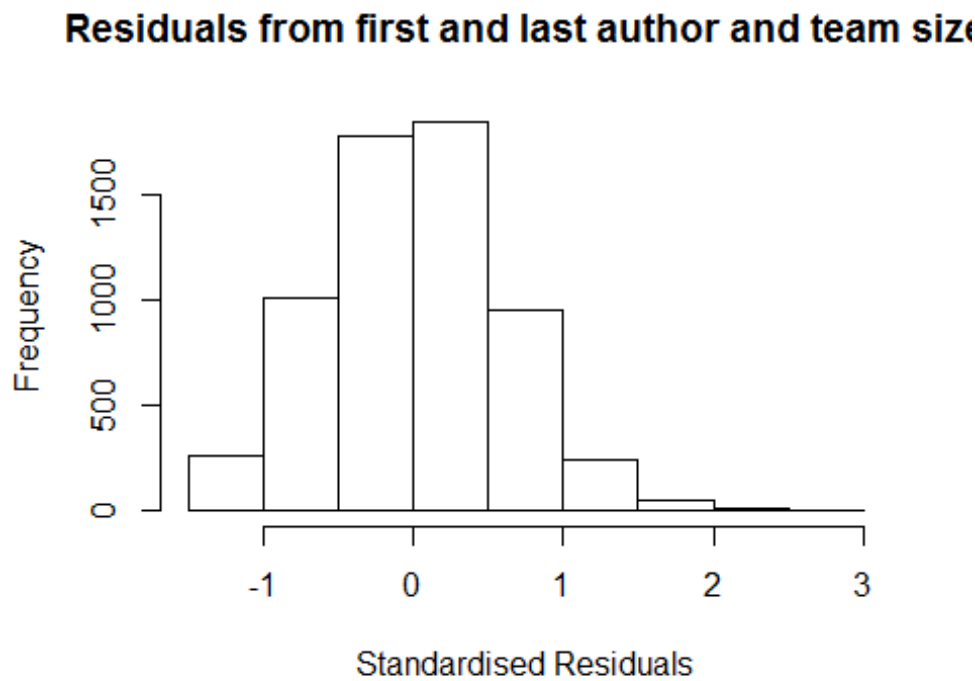
```
##  921  906
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  265  314  265  202  242  242  244  243  246  230  287  324  429  525  607
## 2011 2012
##  740  725
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
##  Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 140, df = 16, p-value <2e-16
```




```
##
##  Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 5.6, df = 1, p-value = 0.02
```

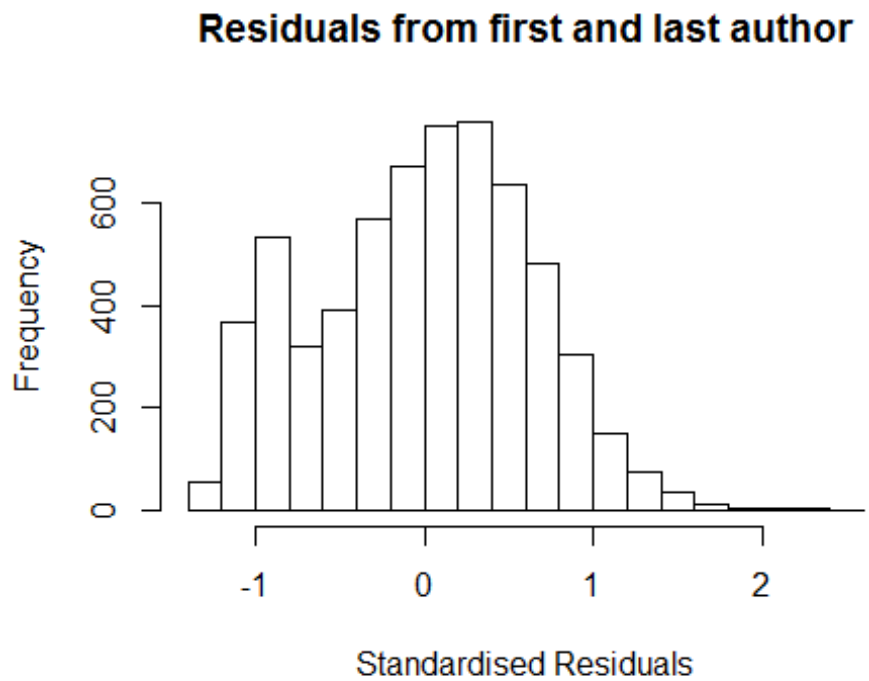



```
## [1] "Female first author team size 2018 geometric mean: 3.26600185097304"
## [1] "Male first author team size 2018 geometric mean: 3.29839149220884"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 110000, p-value = 1
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 2.94908090483912"
## [1] "Male last author team size 2018 geometric mean: 3.38929374670609"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 80000, p-value = 0.007
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##                     GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.130  1           1.063
## LastAuthorFemale  1.143  1           1.069
## UniqueAuthors     1.163  4           1.019
## Year              1.168 16           1.005
```



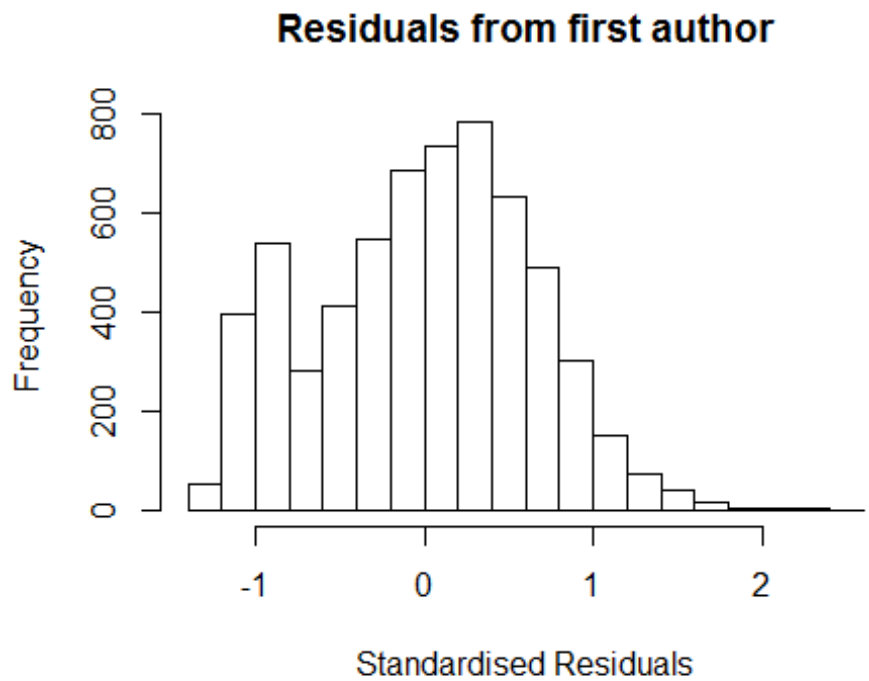

```
## [1] "List of  3 outliers with residuals above  2.5"
##        ScopusId  NLCS Year OneField Fields residuals
## 3394 0034324931 3.324 2000     1606      4     2.712
## 4366 0037054468 2.986 2002     2102      4     2.512
## 4475 0036489148 3.220 2002     2102      4     2.746
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
UniqueAuthors +
##     Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##     k.max = 1000))
##  \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.40037 -0.42658  0.00678  0.40374  2.74597
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)          0.5846     0.0393   14.89  < 2e-16 ***
## FirstAuthorFemale1  -0.0150     0.0207   -0.73  0.46835
## LastAuthorFemale1   -0.0860     0.0241   -3.57  0.00036 ***
## UniqueAuthors2       0.5315     0.0253   21.00  < 2e-16 ***
## UniqueAuthors3       0.5774     0.0254   22.76  < 2e-16 ***
## UniqueAuthors4       0.6761     0.0272   24.83  < 2e-16 ***
## UniqueAuthors5       0.6970     0.0263   26.50  < 2e-16 ***
## Year1997            -0.0529     0.0547   -0.97  0.33422
```

```
## Year1998            -0.0854     0.0534   -1.60  0.11001
## Year1999             0.0605     0.0604    1.00  0.31676
## Year2000             0.0270     0.0624    0.43  0.66568
## Year2001            -0.0500     0.0589   -0.85  0.39621
## Year2002            -0.1106     0.0583   -1.90  0.05771 .
## Year2003             0.0776     0.0564    1.37  0.16933
## Year2004             0.0480     0.0556    0.86  0.38772
## Year2005             0.1310     0.0534    2.45  0.01422 *
## Year2006             0.0732     0.0527    1.39  0.16479
## Year2007             0.1397     0.0483    2.89  0.00383 **
## Year2008             0.0130     0.0463    0.28  0.77824
## Year2009             0.1076     0.0454    2.37  0.01768 *
## Year2010             0.1125     0.0421    2.67  0.00756 **
## Year2011             0.1115     0.0419    2.66  0.00781 **
## Year2012             0.0658     0.0418    1.57  0.11576
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.558
## Multiple R-squared:  0.209,  Adjusted R-squared:  0.206
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
##  2 observations c(1250,1668) are outliers with |weight| = 0 ( < 1.6e-05);
##  462 weights are ~= 1. The remaining 5666 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0059  0.8660  0.9400  0.9000  0.9830  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.63e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.069  1           1.034
## LastAuthorFemale  1.073  1           1.036
## Year              1.017 16           1.001
```



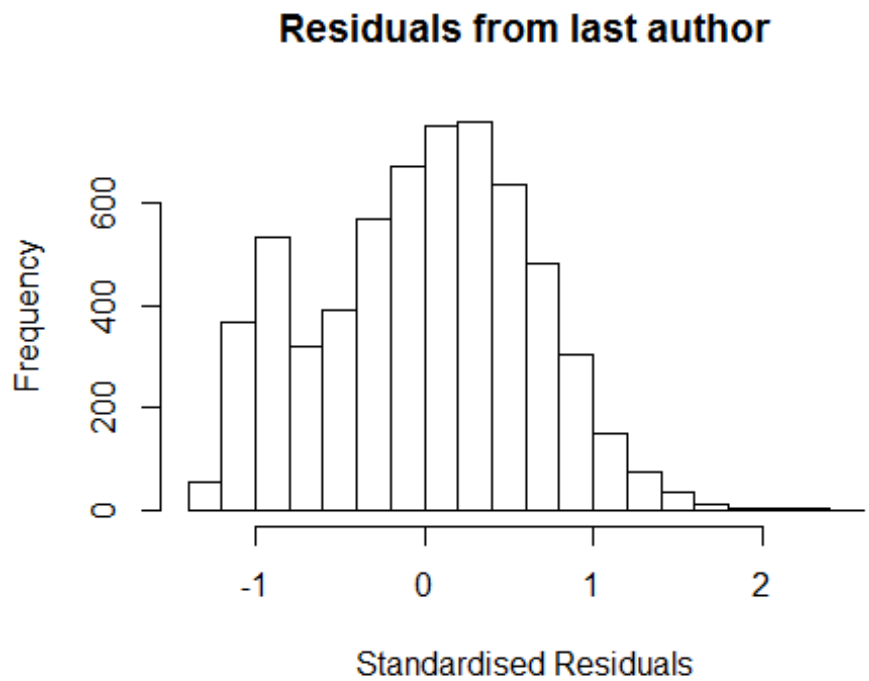

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##    Min     1Q Median     3Q    Max
## -1.231 -0.466  0.050  0.447  2.411
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)           0.9590     0.0429   22.37  < 2e-16 ***
## FirstAuthorFemale1    0.0121     0.0219    0.55   0.5794
## LastAuthorFemale1    -0.1195     0.0260   -4.60  4.3e-06 ***
## Year1997             -0.0737     0.0619   -1.19   0.2339
## Year1998             -0.0967     0.0625   -1.55   0.1220
## Year1999              0.0852     0.0691    1.23   0.2180
## Year2000              0.0400     0.0700    0.57   0.5675
## Year2001             -0.0496     0.0648   -0.77   0.4442
## Year2002             -0.1495     0.0720   -2.08   0.0378 *
## Year2003              0.0971     0.0649    1.50   0.1349
## Year2004              0.0874     0.0638    1.37   0.1705
## Year2005              0.2054     0.0618    3.32   0.0009 ***
```

```
## Year2006             0.1913     0.0593    3.23   0.0013 **
## Year2007             0.2308     0.0553    4.18  3.0e-05 ***
## Year2008             0.1041     0.0524    1.99   0.0470 *
## Year2009             0.1711     0.0526    3.25   0.0011 **
## Year2010             0.2251     0.0484    4.65  3.4e-06 ***
## Year2011             0.2600     0.0479    5.43  5.8e-08 ***
## Year2012             0.2143     0.0480    4.47  8.0e-06 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.652
## Multiple R-squared:  0.0367, Adjusted R-squared:  0.0338
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
##  466 weights are ~= 1. The remaining 5664 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.142   0.860   0.947   0.912   0.985   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.63e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.009  1           1.004
## Year              1.009 16           1.000
```




```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##    Min     1Q Median     3Q    Max
## -1.210 -0.469  0.047  0.452  2.420
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)           0.9515     0.0426   22.36  < 2e-16 ***
## FirstAuthorFemale1   -0.0231     0.0219   -1.06   0.2911
## Year1997             -0.0734     0.0619   -1.19   0.2354
## Year1998             -0.0968     0.0623   -1.55   0.1204
## Year1999              0.0847     0.0688    1.23   0.2188
## Year2000              0.0367     0.0696    0.53   0.5983
## Year2001             -0.0571     0.0647   -0.88   0.3778
## Year2002             -0.1511     0.0720   -2.10   0.0357 *
## Year2003              0.0949     0.0647    1.47   0.1425
## Year2004              0.0802     0.0632    1.27   0.2045
## Year2005              0.2026     0.0617    3.28   0.0010 **
## Year2006              0.1888     0.0592    3.19   0.0014 **
```

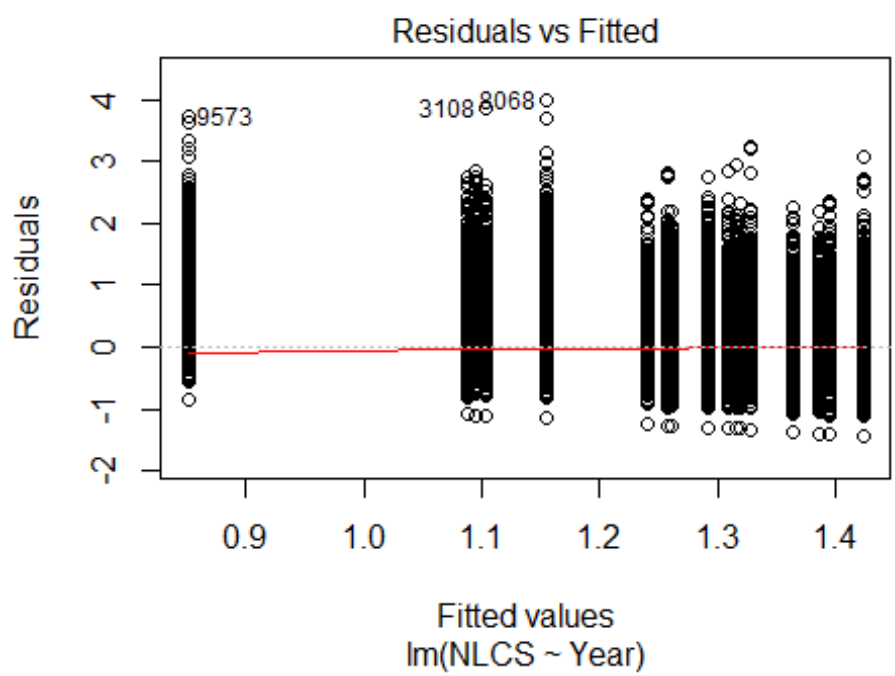
```
## Year2007             0.2294     0.0551    4.16  3.2e-05 ***
## Year2008             0.0981     0.0522    1.88   0.0601 .
## Year2009             0.1683     0.0524    3.21   0.0013 **
## Year2010             0.2233     0.0482    4.63  3.8e-06 ***
## Year2011             0.2589     0.0476    5.44  5.5e-08 ***
## Year2012             0.2133     0.0477    4.47  8.0e-06 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.654
## Multiple R-squared:  0.0334, Adjusted R-squared:  0.0307
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
##  461 weights are ~= 1. The remaining 5669 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.142   0.860   0.948   0.912   0.985   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.63e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##                  GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.01  1           1.005
## Year             1.01 16           1.000
```



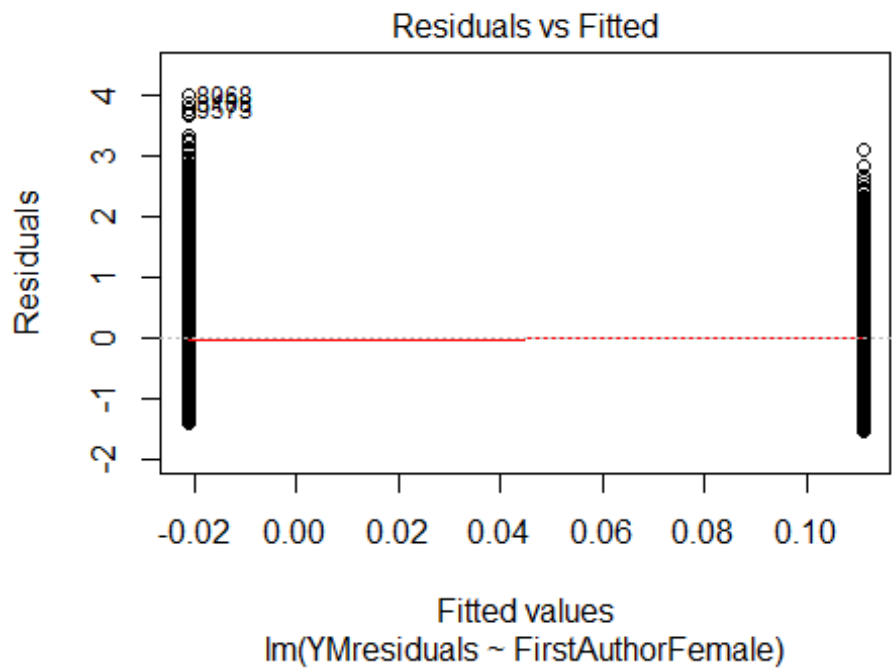

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.2207 -0.4663  0.0494  0.4473  2.4096
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)          0.9602     0.0428   22.44  < 2e-16 ***
## LastAuthorFemale1   -0.1150     0.0252   -4.56  5.2e-06 ***
## Year1997            -0.0741     0.0620   -1.20   0.2316
## Year1998            -0.0967     0.0625   -1.55   0.1218
## Year1999             0.0855     0.0691    1.24   0.2162
## Year2000             0.0397     0.0700    0.57   0.5705
## Year2001            -0.0492     0.0648   -0.76   0.4478
## Year2002            -0.1498     0.0720   -2.08   0.0374 *
## Year2003             0.0967     0.0649    1.49   0.1363
## Year2004             0.0873     0.0638    1.37   0.1713
## Year2005             0.2054     0.0619    3.32   0.0009 ***
## Year2006             0.1911     0.0593    3.22   0.0013 **
```

```
## Year2007            0.2312     0.0553    4.18  2.9e-05 ***
## Year2008            0.1045     0.0524    2.00   0.0461 *
## Year2009            0.1714     0.0526    3.26   0.0011 **
## Year2010            0.2257     0.0484    4.66  3.2e-06 ***
## Year2011            0.2605     0.0479    5.44  5.5e-08 ***
## Year2012            0.2151     0.0480    4.48  7.5e-06 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.652
## Multiple R-squared:  0.0366, Adjusted R-squared:  0.034
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
##  464 weights are ~= 1. The remaining 5666 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.142   0.860   0.947   0.912   0.985   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.63e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Sample size for the above analysis:  6130"
## [1] ""
## [1] ""
## [1] "####################################"
## [1] "Analysis of AJSC 2200"
## [1] "####################################"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 4054 2269 1925 2426 1898 1373 1136  992 1056 1274 1462 1302 1531 1712 1703
## 2011 2012
## 1676 1746
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 2347 1241 1086 1545 1096  666  679  575  615  748  893  826  971 1030 1011
## 2011 2012
```

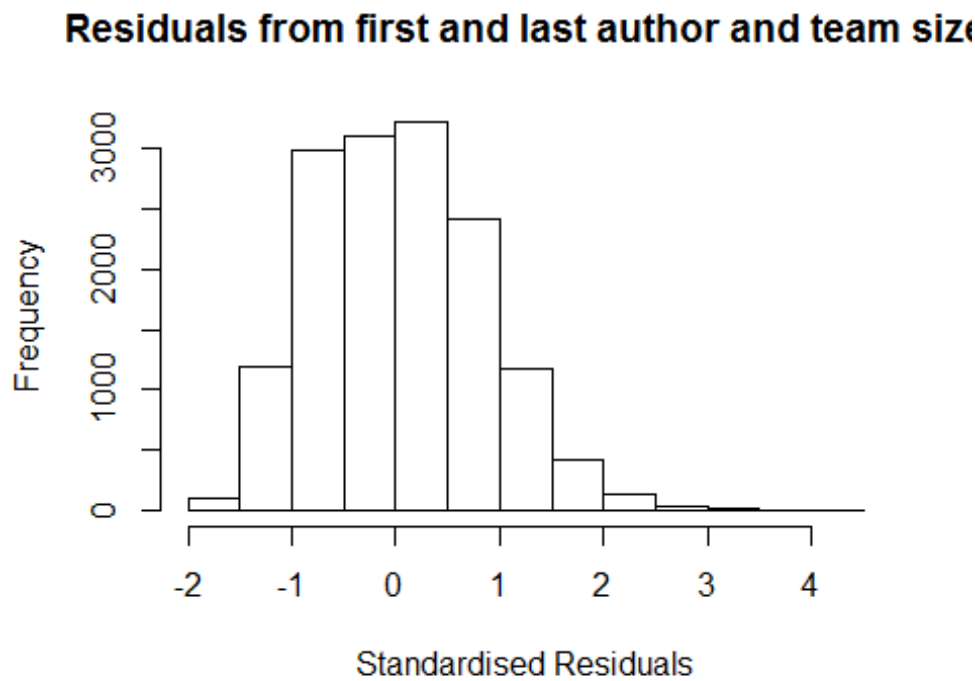
```
##  978 1003
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 2045 1099  958 1379  954  577  560  497  498  618  755  712  821  867  822
## 2011 2012
##  811  816
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
##  Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 160, df = 16, p-value <2e-16
```




```
##
##  Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 0.4, df = 1, p-value = 0.5
```



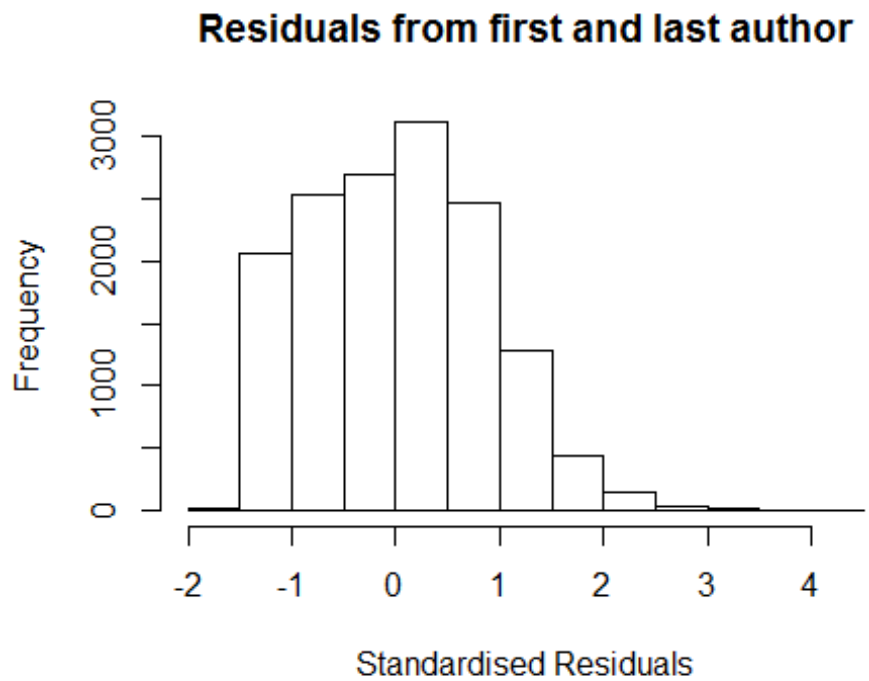

```
## [1] "Female first author team size 2018 geometric mean: 3.1230956612823"
## [1] "Male first author team size 2018 geometric mean: 2.53771454564902"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 19000, p-value = 0.01
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 2.69424608163119"
## [1] "Male last author team size 2018 geometric mean: 2.65368312345395"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 15000, p-value = 0.8
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##                     GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.194  1           1.093
## LastAuthorFemale  1.187  1           1.089
## UniqueAuthors     1.077  4           1.009
## Year              1.083 16           1.002
```




```
## [1] "List of  51 outliers with residuals above  2.5"
##          ScopusId  NLCS Year OneField Fields residuals
## 669   27244447775 3.352 1996     2200      1     2.611
## 1251   0030270991 3.271 1996     2200      1     2.530
## 2847   0002362842 3.459 1996     2200      1     2.598
## 3043   0030105583 3.708 1996     2200      1     2.611
## 3108   0030162959 4.956 1996     2200      1     3.859
## 3140   0030193251 3.423 1996     2200      1     2.562
## 4837   0031271567 3.528 1997     1405      3     2.662
## 5331   0031139222 3.785 1997     2200      1     3.039
## 5844   0030678920 3.755 1997     2200      1     2.501
## 5859   0030687194 3.926 1997     2200      1     2.672
## 6032   0031234125 3.838 1997     2200      1     3.092
## 6315   0031190973 3.602 1997     1700      2     2.856
## 6848   0032207559 3.561 1998     2200      1     2.628
## 6850   0032208441 3.888 1998     2200      1     3.075
## 7546   0032117028 4.290 1998     2200      1     2.969
## 8068   0031728565 5.139 1998     2200      1     3.570
## 8095   0032047559 4.844 1998     2200      1     3.651
## 8184   0041344867 4.120 1998     2200      1     3.307
## 8240   0000763845 3.978 1998     2200      2     2.810
## 8745   0141465932 3.529 1998     2200      3     2.716
## 8891   0033101948 3.622 1999     2200      1     3.061
## 9481   0033316052 2.983 1999     2200      1     2.542
## 9541   0033354342 3.934 1999     2200      1     3.493
## 9546   0033356696 3.551 1999     2200      1     2.990
## 9573   0033366376 4.589 1999     2200      1     4.148
```

```
## 9939   0033349534 4.175 1999     2200      3     3.226
## 10111  0032657184 4.491 1999     2200      1     3.293
## 10359  0000330869 3.184 1999     2200      1     2.743
## 10406  0032598062 4.069 1999     2200      1     3.628
## 10642  0032641366 3.270 1999     2200      1     2.829
## 10689  0032663759 3.118 1999     2200      1     2.557
## 10767  0032759697 2.954 1999     2200      1     2.513
## 10776  0032762062 3.329 1999     2200      1     2.888
## 10824  0033189140 3.567 1999     2200      1     3.006
## 10837  0033334086 3.184 1999     2200      1     2.743
## 10839  0033338157 3.430 1999     2200      1     2.869
## 10890  2942597334 3.405 1999     2200      1     2.844
## 11346  0346727379 3.513 1999     2200      3     2.716
## 12378  0033666418 3.349 2000     2200      1     2.616
## 12856  0038063991 3.722 2000     2200      1     2.989
## 13942  0034765279 4.537 2001     2200      1     2.704
## 14626 33845594450 4.567 2001     2200      1     3.059
## 15517  0036359986 4.076 2002     2200      1     3.050
## 15524  0036373007 4.093 2002     2200      1     3.067
## 15819  0036526755 3.549 2002     2200      1     2.523
## 15916 84893808653 3.950 2002     2200      1     2.568
## 17109  0346780414 3.738 2003     2200      1     2.752
## 22596 34247481878 4.038 2007     1700      2     2.832
## 22597 34247500374 4.020 2007     1700      2     3.169
## 26213 65549134300 3.585 2009     2200      3     2.776
## 27740 78650092372 4.275 2010     2200      3     2.571
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
UniqueAuthors +
##     Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##     k.max = 1000))
##  \--> method = "MM"
## Residuals:
##       Min        1Q    Median        3Q       Max
## -1.833091 -0.647181  0.000738  0.575320  4.147751
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         0.74113    0.02076   35.71  < 2e-16 ***
## FirstAuthorFemale1  0.09555    0.02079    4.60  4.3e-06 ***
## LastAuthorFemale1   0.02447    0.02236    1.09  0.27394
## UniqueAuthors2      0.35562    0.01785   19.93  < 2e-16 ***
## UniqueAuthors3      0.50819    0.02122   23.95  < 2e-16 ***
## UniqueAuthors4      0.55198    0.02597   21.26  < 2e-16 ***
## UniqueAuthors5      0.75690    0.02419   31.28  < 2e-16 ***
## Year1997            0.00493    0.03390    0.15  0.88430
## Year1998            0.07145    0.03674    1.94  0.05185 .
## Year1999           -0.29988    0.03102   -9.67  < 2e-16 ***
```

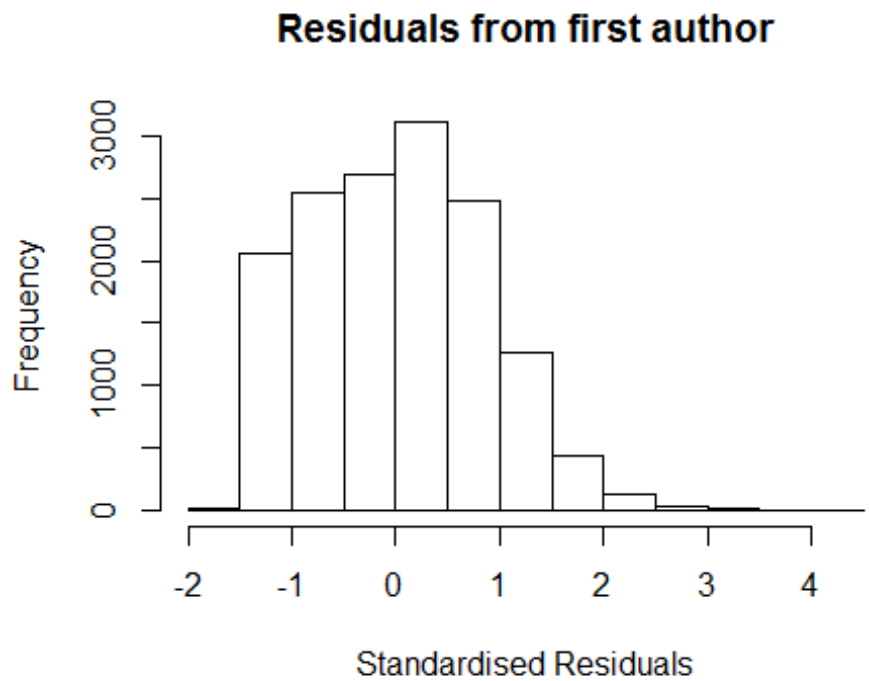
```
## Year2000           -0.00852    0.03293   -0.26  0.79594
## Year2001            0.21507    0.04179    5.15  2.7e-07 ***
## Year2002            0.28514    0.03636    7.84  4.7e-15 ***
## Year2003            0.24445    0.04063    6.02  1.8e-09 ***
## Year2004            0.26079    0.03858    6.76  1.4e-11 ***
## Year2005            0.18379    0.03808    4.83  1.4e-06 ***
## Year2006            0.13737    0.03553    3.87  0.00011 ***
## Year2007            0.10973    0.03532    3.11  0.00189 **
## Year2008            0.12951    0.03222    4.02  5.8e-05 ***
## Year2009            0.06743    0.03278    2.06  0.03967 *
## Year2010            0.11080    0.03334    3.32  0.00089 ***
## Year2011            0.14116    0.03307    4.27  2.0e-05 ***
## Year2012            0.21506    0.03261    6.59  4.4e-11 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.827
## Multiple R-squared:  0.13,   Adjusted R-squared:  0.129
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  observation 4513 is an outlier with |weight| = 0 ( < 6.8e-06);
##  1089 weights are ~= 1. The remaining 13699 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0001  0.8820  0.9420  0.9130  0.9830  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          6.76e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.161  1           1.077
## LastAuthorFemale  1.159  1           1.077
## Year              1.019 16           1.001
```




```
## [1] "List of  47 outliers with residuals above  2.5"
##          ScopusId  NLCS Year OneField Fields residuals
## 3043   0030105583 3.708 1996     2200      1     2.674
## 3108   0030162959 4.956 1996     2200      1     3.922
## 4447   0031374947 3.556 1997     2200      1     2.568
## 4888   0031250434 3.653 1997     2200      1     2.665
## 5135   0030947661 3.738 1997     2200      1     2.602
## 5331   0031139222 3.785 1997     2200      1     2.797
## 5844   0030678920 3.755 1997     2200      1     2.767
## 5849   0030682930 3.504 1997     2200      1     2.516
## 5859   0030687194 3.926 1997     2200      1     2.938
## 6032   0031234125 3.838 1997     2200      1     2.850
## 6127  33748731480 3.560 1997     2200      1     2.572
## 6315   0031190973 3.602 1997     1700      2     2.614
## 6840   0032123731 3.836 1998     2200      1     2.749
## 6850   0032208441 3.888 1998     2200      1     2.822
## 7546   0032117028 4.290 1998     2200      1     3.224
## 8068   0031728565 5.139 1998     2200      1     4.073
## 8095   0032047559 4.844 1998     2200      1     3.757
## 8184   0041344867 4.120 1998     2200      1     3.054
## 8240   0000763845 3.978 1998     2200      2     2.912
## 8891   0033101948 3.622 1999     2200      1     2.705
## 9541   0033354342 3.934 1999     2200      1     3.185
## 9546   0033356696 3.551 1999     2200      1     2.634
## 9563   0033363601 3.357 1999     2200      1     2.608
## 9573   0033366376 4.589 1999     2200      1     3.840
## 9939   0033349534 4.175 1999     2200      3     3.426
```

```
## 10111  0032657184 4.491 1999     2200      1     3.742
## 10406  0032598062 4.069 1999     2200      1     3.320
## 10642  0032641366 3.270 1999     2200      1     2.521
## 10753  0032704301 3.282 1999     2200      1     2.533
## 10776  0032762062 3.329 1999     2200      1     2.580
## 10793  0032970565 3.379 1999     2200      1     2.609
## 10824  0033189140 3.567 1999     2200      1     2.650
## 10839  0033338157 3.430 1999     2200      1     2.513
## 11346  0346727379 3.513 1999     2200      3     2.764
## 11360  0347467670 3.304 1999     2200      3     2.555
## 12099  0034210957 3.685 2000     2200      1     2.678
## 12856  0038063991 3.722 2000     2200      1     2.715
## 13942  0034765279 4.537 2001     2200      1     3.119
## 14626 33845594450 4.567 2001     2200      1     3.317
## 15517  0036359986 4.076 2002     2200      1     2.717
## 15524  0036373007 4.093 2002     2200      1     2.734
## 15916 84893808653 3.950 2002     2200      1     2.591
## 19671 15844407150 4.159 2005     1708      2     2.893
## 22451 34250883216 4.080 2007     2200      1     2.878
## 22596 34247481878 4.038 2007     1700      2     2.857
## 22597 34247500374 4.020 2007     1700      2     2.839
## 27740 78650092372 4.275 2010     2200      3     2.882
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.5381 -0.6939  0.0116  0.6169  4.0729
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)          1.0339     0.0202   51.29  < 2e-16 ***
## FirstAuthorFemale1   0.1475     0.0216    6.82  9.4e-12 ***
## LastAuthorFemale1    0.0205     0.0233    0.88   0.3777
## Year1997            -0.0459     0.0373   -1.23   0.2186
## Year1998             0.0322     0.0394    0.82   0.4136
## Year1999            -0.2845     0.0316   -9.02  < 2e-16 ***
## Year2000            -0.0268     0.0348   -0.77   0.4418
## Year2001             0.2160     0.0438    4.93  8.2e-07 ***
## Year2002             0.3253     0.0379    8.59  < 2e-16 ***
## Year2003             0.2822     0.0427    6.61  4.1e-11 ***
## Year2004             0.2749     0.0417    6.59  4.6e-11 ***
## Year2005             0.2324     0.0398    5.85  5.1e-09 ***
## Year2006             0.1792     0.0379    4.73  2.3e-06 ***
## Year2007             0.1475     0.0380    3.88   0.0001 ***
## Year2008             0.1710     0.0344    4.96  7.0e-07 ***
## Year2009             0.1372     0.0350    3.92  8.9e-05 ***
```

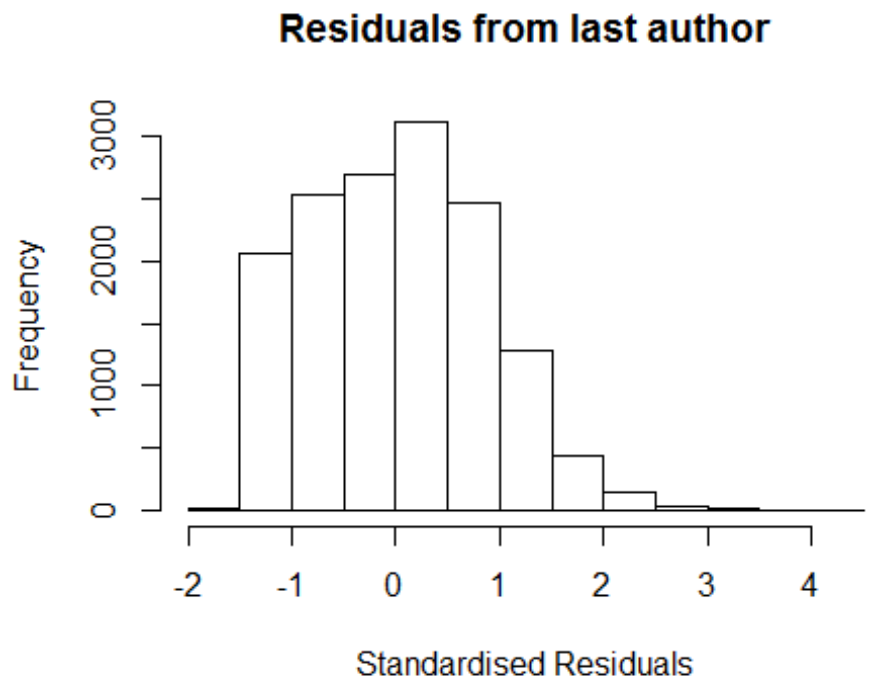
```
## Year2010             0.2120     0.0357    5.93  3.1e-09 ***
## Year2011             0.2434     0.0352    6.92  4.7e-12 ***
## Year2012             0.3361     0.0352    9.56  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.902
## Multiple R-squared:  0.0432, Adjusted R-squared:  0.042
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  1162 weights are ~= 1. The remaining 13627 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.005   0.884   0.943   0.919   0.984   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          6.76e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.011  1           1.006
## Year              1.011 16           1.000
```




```
## [1] "List of  47 outliers with residuals above  2.5"
##          ScopusId  NLCS Year OneField Fields residuals
## 3043   0030105583 3.708 1996     2200      1     2.674
## 3108   0030162959 4.956 1996     2200      1     3.922
## 4447   0031374947 3.556 1997     2200      1     2.568
## 4888   0031250434 3.653 1997     2200      1     2.665
## 5135   0030947661 3.738 1997     2200      1     2.602
## 5331   0031139222 3.785 1997     2200      1     2.797
## 5844   0030678920 3.755 1997     2200      1     2.767
## 5849   0030682930 3.504 1997     2200      1     2.516
## 5859   0030687194 3.926 1997     2200      1     2.938
## 6032   0031234125 3.838 1997     2200      1     2.850
## 6127  33748731480 3.560 1997     2200      1     2.572
## 6315   0031190973 3.602 1997     1700      2     2.614
## 6840   0032123731 3.836 1998     2200      1     2.749
## 6850   0032208441 3.888 1998     2200      1     2.822
## 7546   0032117028 4.290 1998     2200      1     3.224
## 8068   0031728565 5.139 1998     2200      1     4.073
## 8095   0032047559 4.844 1998     2200      1     3.757
## 8184   0041344867 4.120 1998     2200      1     3.054
## 8240   0000763845 3.978 1998     2200      2     2.912
## 8891   0033101948 3.622 1999     2200      1     2.705
## 9541   0033354342 3.934 1999     2200      1     3.185
## 9546   0033356696 3.551 1999     2200      1     2.634
## 9563   0033363601 3.357 1999     2200      1     2.608
## 9573   0033366376 4.589 1999     2200      1     3.840
## 9939   0033349534 4.175 1999     2200      3     3.426
```

```
## 10111  0032657184 4.491 1999     2200      1     3.742
## 10406  0032598062 4.069 1999     2200      1     3.320
## 10642  0032641366 3.270 1999     2200      1     2.521
## 10753  0032704301 3.282 1999     2200      1     2.533
## 10776  0032762062 3.329 1999     2200      1     2.580
## 10793  0032970565 3.379 1999     2200      1     2.609
## 10824  0033189140 3.567 1999     2200      1     2.650
## 10839  0033338157 3.430 1999     2200      1     2.513
## 11346  0346727379 3.513 1999     2200      3     2.764
## 11360  0347467670 3.304 1999     2200      3     2.555
## 12099  0034210957 3.685 2000     2200      1     2.678
## 12856  0038063991 3.722 2000     2200      1     2.715
## 13942  0034765279 4.537 2001     2200      1     3.119
## 14626 33845594450 4.567 2001     2200      1     3.317
## 15517  0036359986 4.076 2002     2200      1     2.717
## 15524  0036373007 4.093 2002     2200      1     2.734
## 15916 84893808653 3.950 2002     2200      1     2.591
## 19671 15844407150 4.159 2005     1708      2     2.893
## 22451 34250883216 4.080 2007     2200      1     2.878
## 22596 34247481878 4.038 2007     1700      2     2.857
## 22597 34247500374 4.020 2007     1700      2     2.839
## 27740 78650092372 4.275 2010     2200      3     2.882
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.5268 -0.6955  0.0102  0.6173  4.0712
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)          1.0348     0.0201   51.37  < 2e-16 ***
## FirstAuthorFemale1   0.1550     0.0202    7.69  1.6e-14 ***
## Year1997            -0.0449     0.0372   -1.21     0.23
## Year1998             0.0330     0.0394    0.84     0.40
## Year1999            -0.2840     0.0316   -9.00  < 2e-16 ***
## Year2000            -0.0263     0.0348   -0.75     0.45
## Year2001             0.2168     0.0438    4.95  7.5e-07 ***
## Year2002             0.3264     0.0378    8.64  < 2e-16 ***
## Year2003             0.2834     0.0427    6.64  3.3e-11 ***
## Year2004             0.2760     0.0417    6.62  3.7e-11 ***
## Year2005             0.2332     0.0397    5.87  4.5e-09 ***
## Year2006             0.1798     0.0379    4.74  2.1e-06 ***
## Year2007             0.1481     0.0380    3.90  9.8e-05 ***
## Year2008             0.1719     0.0344    4.99  6.0e-07 ***
## Year2009             0.1379     0.0350    3.94  8.2e-05 ***
## Year2010             0.2130     0.0357    5.96  2.5e-09 ***
```

```
## Year2011             0.2450     0.0351    6.97  3.2e-12 ***
## Year2012             0.3370     0.0351    9.59  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.902
## Multiple R-squared:  0.0431, Adjusted R-squared:  0.042
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  1162 weights are ~= 1. The remaining 13627 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0052  0.8840  0.9440  0.9190  0.9840  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          6.76e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##                  GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.01  1           1.005
## Year             1.01 16           1.000
```

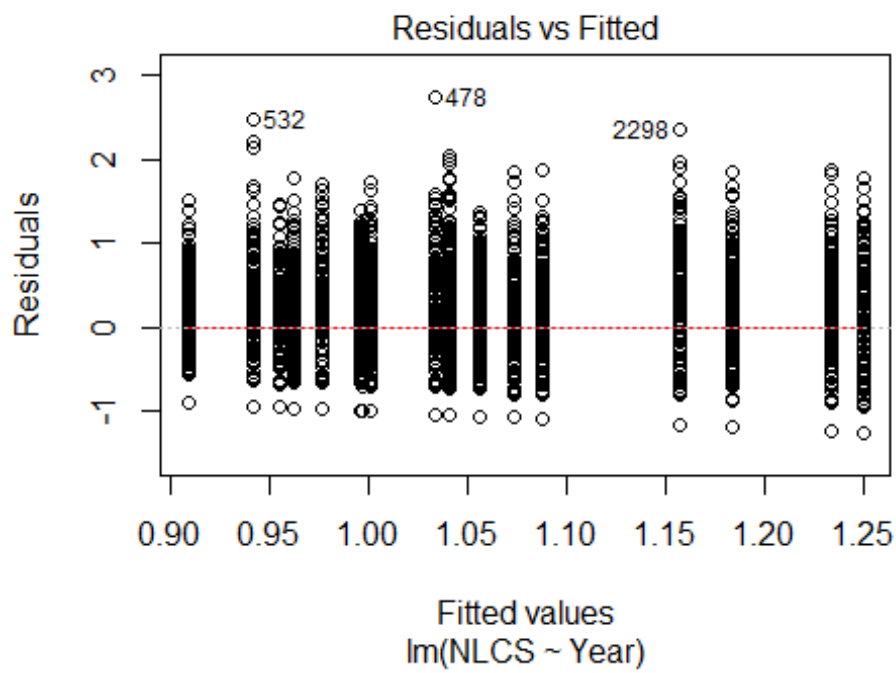



```
## [1] "List of  47 outliers with residuals above  2.5"
##          ScopusId  NLCS Year OneField Fields residuals
## 3043   0030105583 3.708 1996     2200      1     2.674
## 3108   0030162959 4.956 1996     2200      1     3.922
## 4447   0031374947 3.556 1997     2200      1     2.568
## 4888   0031250434 3.653 1997     2200      1     2.665
## 5135   0030947661 3.738 1997     2200      1     2.602
## 5331   0031139222 3.785 1997     2200      1     2.797
## 5844   0030678920 3.755 1997     2200      1     2.767
## 5849   0030682930 3.504 1997     2200      1     2.516
## 5859   0030687194 3.926 1997     2200      1     2.938
## 6032   0031234125 3.838 1997     2200      1     2.850
## 6127  33748731480 3.560 1997     2200      1     2.572
## 6315   0031190973 3.602 1997     1700      2     2.614
## 6840   0032123731 3.836 1998     2200      1     2.749
## 6850   0032208441 3.888 1998     2200      1     2.822
## 7546   0032117028 4.290 1998     2200      1     3.224
## 8068   0031728565 5.139 1998     2200      1     4.073
## 8095   0032047559 4.844 1998     2200      1     3.757
## 8184   0041344867 4.120 1998     2200      1     3.054
## 8240   0000763845 3.978 1998     2200      2     2.912
## 8891   0033101948 3.622 1999     2200      1     2.705
## 9541   0033354342 3.934 1999     2200      1     3.185
## 9546   0033356696 3.551 1999     2200      1     2.634
## 9563   0033363601 3.357 1999     2200      1     2.608
## 9573   0033366376 4.589 1999     2200      1     3.840
## 9939   0033349534 4.175 1999     2200      3     3.426
```

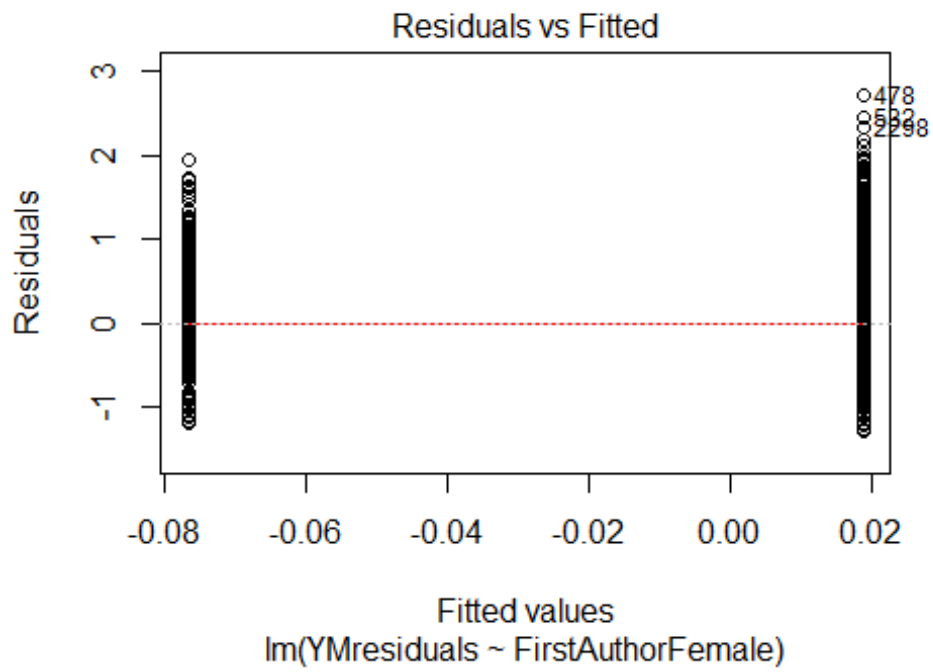
```
## 10111  0032657184 4.491 1999     2200      1     3.742
## 10406  0032598062 4.069 1999     2200      1     3.320
## 10642  0032641366 3.270 1999     2200      1     2.521
## 10753  0032704301 3.282 1999     2200      1     2.533
## 10776  0032762062 3.329 1999     2200      1     2.580
## 10793  0032970565 3.379 1999     2200      1     2.609
## 10824  0033189140 3.567 1999     2200      1     2.650
## 10839  0033338157 3.430 1999     2200      1     2.513
## 11346  0346727379 3.513 1999     2200      3     2.764
## 11360  0347467670 3.304 1999     2200      3     2.555
## 12099  0034210957 3.685 2000     2200      1     2.678
## 12856  0038063991 3.722 2000     2200      1     2.715
## 13942  0034765279 4.537 2001     2200      1     3.119
## 14626 33845594450 4.567 2001     2200      1     3.317
## 15517  0036359986 4.076 2002     2200      1     2.717
## 15524  0036373007 4.093 2002     2200      1     2.734
## 15916 84893808653 3.950 2002     2200      1     2.591
## 19671 15844407150 4.159 2005     1708      2     2.893
## 22451 34250883216 4.080 2007     2200      1     2.878
## 22596 34247481878 4.038 2007     1700      2     2.857
## 22597 34247500374 4.020 2007     1700      2     2.839
## 27740 78650092372 4.275 2010     2200      3     2.882
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.4733 -0.6921  0.0073  0.6225  4.0624
##
## Coefficients:
##                   Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.0470     0.0201   52.09  < 2e-16 ***
## LastAuthorFemale1   0.0820     0.0214    3.83  0.00013 ***
## Year1997           -0.0485     0.0374   -1.30  0.19398
## Year1998            0.0296     0.0394    0.75  0.45247
## Year1999           -0.2834     0.0316   -8.98  < 2e-16 ***
## Year2000           -0.0283     0.0349   -0.81  0.41786
## Year2001            0.2145     0.0438    4.90  9.8e-07 ***
## Year2002            0.3250     0.0380    8.55  < 2e-16 ***
## Year2003            0.2863     0.0427    6.70  2.2e-11 ***
## Year2004            0.2807     0.0418    6.71  2.0e-11 ***
## Year2005            0.2337     0.0397    5.89  4.0e-09 ***
## Year2006            0.1835     0.0380    4.82  1.4e-06 ***
## Year2007            0.1526     0.0382    4.00  6.4e-05 ***
## Year2008            0.1747     0.0346    5.04  4.7e-07 ***
## Year2009            0.1447     0.0351    4.12  3.8e-05 ***
## Year2010            0.2161     0.0359    6.03  1.7e-09 ***
```

```
## Year2011            0.2512     0.0351    7.15  9.2e-13 ***
## Year2012            0.3443     0.0352    9.77  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.904
## Multiple R-squared:  0.0399, Adjusted R-squared:  0.0388
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  1144 weights are ~= 1. The remaining 13645 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0064  0.8820  0.9440  0.9190  0.9840  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          6.76e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Sample size for the above analysis:  14789"
## [1] ""
## [1] ""
## [1] "####################################"
## [1] "Analysis of AJSC 2201"
## [1] "####################################"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  244  231  245  227  222  304  315  247  312  339  379  349  339  295  311
## 2011 2012
##  339  357
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  182  159  187  164  142  161  225  171  203  199  231  224  222  199  216
## 2011 2012
##  245  243
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  167  142  167  147  135  143  203  158  184  163  199  202  196  177  191
```

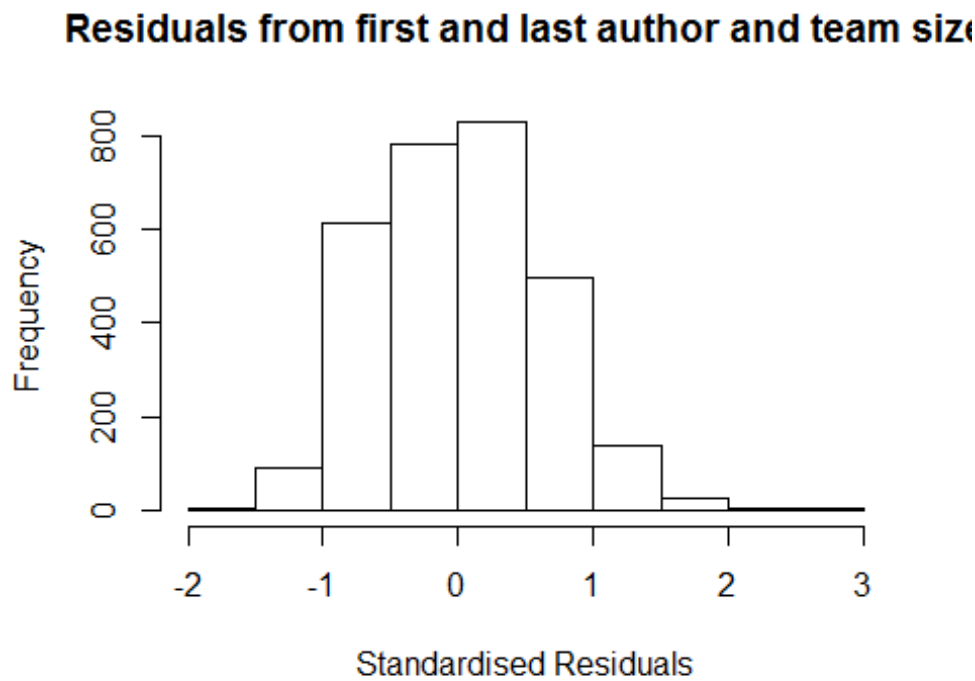
```
## 2011 2012
##  202  211
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
##  Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 72, df = 16, p-value = 4e-09
```




```
##
##  Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 1.3, df = 1, p-value = 0.3
```

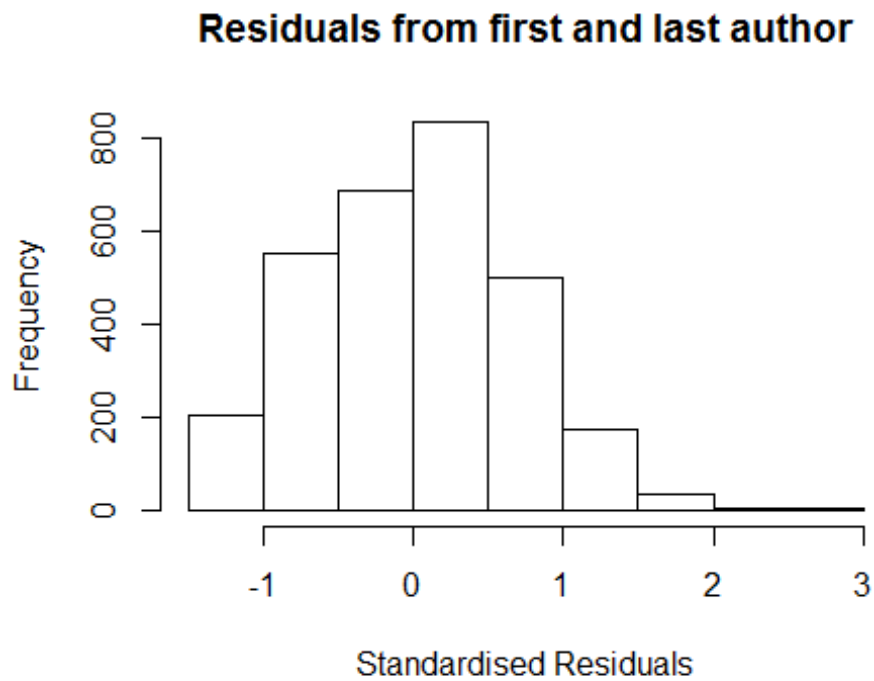



```
## [1] "Female first author team size 2018 geometric mean: 2.61848626763443"
## [1] "Male first author team size 2018 geometric mean: 2.55510454855324"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 7600, p-value = 0.8
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 2.28598149094761"
## [1] "Male last author team size 2018 geometric mean: 2.63655444705472"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 5100, p-value = 0.2
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.626  1           1.275
## LastAuthorFemale  1.623  1           1.274
## UniqueAuthors     1.165  4           1.019
## Year              1.214 16           1.006
```



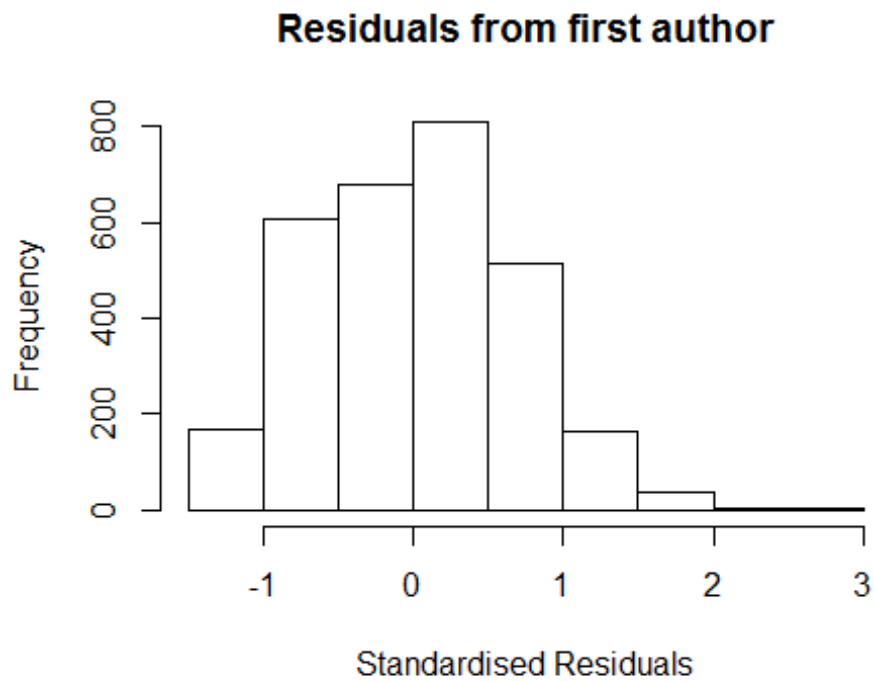

```
## [1] "List of  2 outliers with residuals above  2.5"
##       ScopusId  NLCS Year OneField Fields residuals
## 478 0030643536 3.766 1997     2201      2     2.615
## 753 0032089417 3.145 1998     1000      4     2.546
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
UniqueAuthors +
##     Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##     k.max = 1000))
##  \--> method = "MM"
## Residuals:
##       Min        1Q    Median        3Q       Max
## -1.518866 -0.474588  0.000524  0.444135  2.614827
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)          0.7342     0.0541   13.56   <2e-16 ***
## FirstAuthorFemale1  -0.0236     0.0373   -0.63   0.5273
## LastAuthorFemale1   -0.0834     0.0388   -2.15   0.0315 *
## UniqueAuthors2       0.3477     0.0326   10.66   <2e-16 ***
## UniqueAuthors3       0.5525     0.0371   14.91   <2e-16 ***
## UniqueAuthors4       0.6027     0.0423   14.26   <2e-16 ***
## UniqueAuthors5       0.7154     0.0371   19.30   <2e-16 ***
## Year1997             0.0692     0.0815    0.85   0.3954
## Year1998            -0.1354     0.0734   -1.84   0.0652 .
```

```
## Year1999             0.0911     0.0738    1.23   0.2173
## Year2000            -0.0527     0.0729   -0.72   0.4696
## Year2001             0.2136     0.0774    2.76   0.0058 **
## Year2002             0.0433     0.0783    0.55   0.5805
## Year2003             0.1000     0.0730    1.37   0.1704
## Year2004             0.1034     0.0807    1.28   0.2005
## Year2005             0.1781     0.0810    2.20   0.0280 *
## Year2006            -0.0111     0.0662   -0.17   0.8662
## Year2007            -0.1359     0.0693   -1.96   0.0500 *
## Year2008            -0.0718     0.0682   -1.05   0.2922
## Year2009            -0.0725     0.0683   -1.06   0.2887
## Year2010            -0.0837     0.0718   -1.17   0.2437
## Year2011            -0.0214     0.0682   -0.31   0.7540
## Year2012            -0.2079     0.0691   -3.01   0.0027 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.647
## Multiple R-squared:  0.181,  Adjusted R-squared:  0.175
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  238 weights are ~= 1. The remaining 2749 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.066   0.882   0.946   0.914   0.985   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          3.35e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.622  1           1.274
## LastAuthorFemale  1.613  1           1.270
## Year              1.058 16           1.002
```



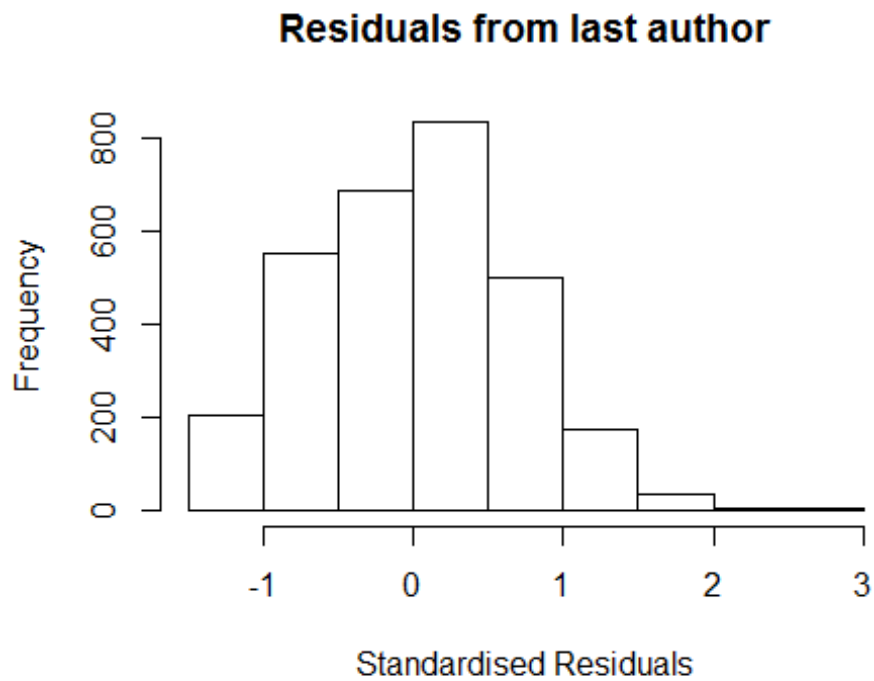

```
## [1] "List of  2 outliers with residuals above  2.5"
##       ScopusId  NLCS Year OneField Fields residuals
## 478 0030643536 3.766 1997     2201      2     2.748
## 532 0032117425 3.423 1998     2201      2     2.554
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.2546 -0.5020  0.0296  0.4823  2.7483
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)          0.95281    0.05387   17.69  < 2e-16 ***
## FirstAuthorFemale1 -0.00821    0.03982   -0.21  0.83662
## LastAuthorFemale1  -0.15217    0.04154   -3.66  0.00025 ***
## Year1997            0.06494    0.08475    0.77  0.44362
## Year1998           -0.08380    0.07707   -1.09  0.27698
## Year1999            0.13168    0.07783    1.69  0.09077 .
## Year2000            0.01785    0.07476    0.24  0.81135
## Year2001            0.28561    0.08398    3.40  0.00068 ***
## Year2002            0.05622    0.08018    0.70  0.48327
## Year2003            0.19423    0.07869    2.47  0.01363 *
## Year2004            0.17324    0.09014    1.92  0.05471 .
```

```
## Year2005            0.30180    0.08171    3.69  0.00022 ***
## Year2006            0.12084    0.06883    1.76  0.07923 .
## Year2007            0.02090    0.07153    0.29  0.77021
## Year2008            0.03386    0.07138    0.47  0.63530
## Year2009            0.05292    0.07024    0.75  0.45131
## Year2010            0.06690    0.07361    0.91  0.36350
## Year2011            0.11910    0.07169    1.66  0.09674 .
## Year2012           -0.03361    0.07113   -0.47  0.63662
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.727
## Multiple R-squared:  0.0281, Adjusted R-squared:  0.0222
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
##  238 weights are ~= 1. The remaining 2749 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.121   0.874   0.950   0.919   0.985   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          3.35e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                  psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.035  1           1.017
## Year              1.035 16           1.001
```




```
## [1] "List of  2 outliers with residuals above  2.5"
##       ScopusId  NLCS Year OneField Fields residuals
## 478 0030643536 3.766 1997     2201      2     2.748
## 532 0032117425 3.423 1998     2201      2     2.554
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.2477 -0.5209  0.0288  0.4847  2.7561
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)           0.9442     0.0541   17.46  < 2e-16 ***
## FirstAuthorFemale1   -0.0924     0.0326   -2.83  0.00466 **
## Year1997              0.0657     0.0848    0.78  0.43832
## Year1998             -0.0843     0.0775   -1.09  0.27641
## Year1999              0.1323     0.0783    1.69  0.09134 .
## Year2000              0.0133     0.0748    0.18  0.85927
## Year2001              0.2860     0.0841    3.40  0.00068 ***
## Year2002              0.0543     0.0802    0.68  0.49830
## Year2003              0.1977     0.0789    2.51  0.01227 *
## Year2004              0.1720     0.0906    1.90  0.05756 .
## Year2005              0.3035     0.0820    3.70  0.00022 ***
```

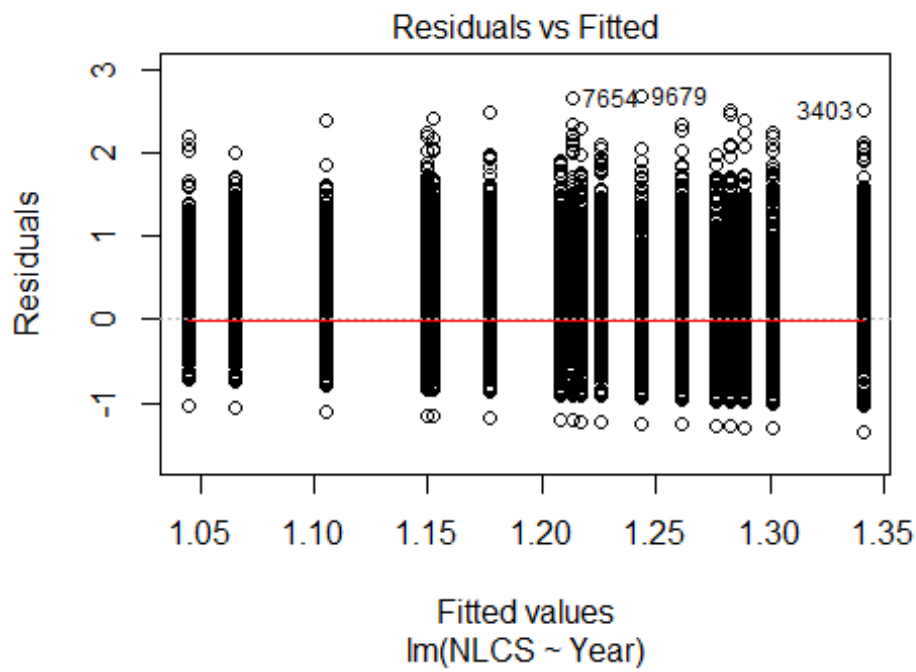
```
## Year2006             0.1207     0.0691    1.75  0.08094 .
## Year2007             0.0185     0.0717    0.26  0.79592
## Year2008             0.0296     0.0713    0.42  0.67799
## Year2009             0.0528     0.0708    0.74  0.45634
## Year2010             0.0570     0.0744    0.77  0.44344
## Year2011             0.1139     0.0719    1.58  0.11311
## Year2012            -0.0472     0.0712   -0.66  0.50685
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.728
## Multiple R-squared:  0.0235, Adjusted R-squared:  0.0179
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
##  259 weights are ~= 1. The remaining 2728 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.120   0.866   0.950   0.919   0.985   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          3.35e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##                   GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.031  1           1.015
## Year             1.031 16           1.001
```



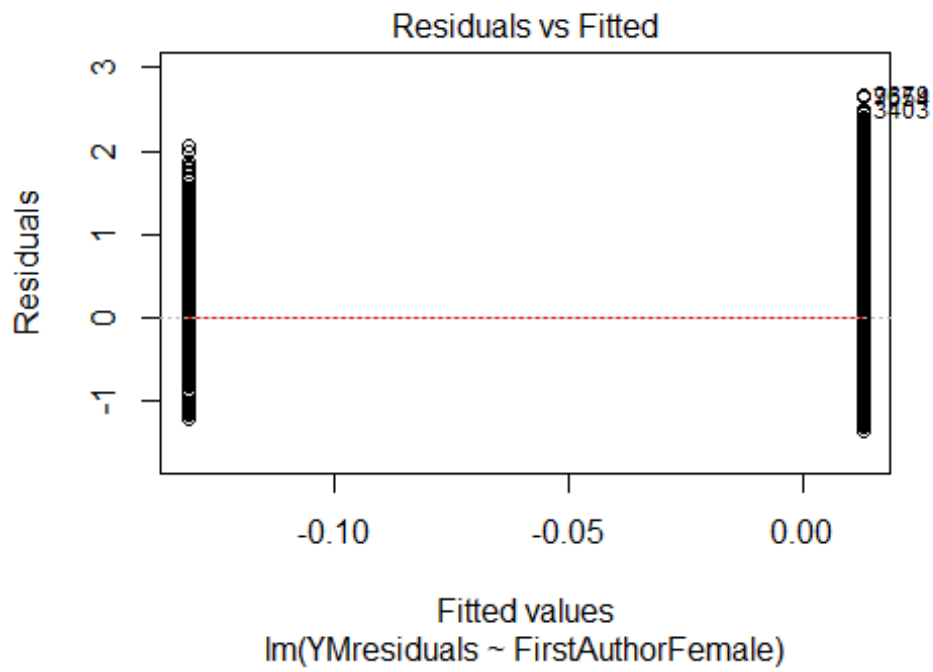

```
## [1] "List of  2 outliers with residuals above  2.5"
##       ScopusId  NLCS Year OneField Fields residuals
## 478 0030643536 3.766 1997     2201      2     2.748
## 532 0032117425 3.423 1998     2201      2     2.554
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##    Min     1Q Median     3Q    Max
## -1.254 -0.504  0.029  0.485  2.749
##
## Coefficients:
##                   Estimate Std. Error t value Pr(>|t|)
## (Intercept)         0.9523     0.0538   17.70  < 2e-16 ***
## LastAuthorFemale1  -0.1570     0.0333   -4.72  2.5e-06 ***
## Year1997            0.0651     0.0848    0.77  0.44273
## Year1998           -0.0837     0.0771   -1.09  0.27738
## Year1999            0.1314     0.0778    1.69  0.09111 .
## Year2000            0.0182     0.0747    0.24  0.80749
## Year2001            0.2857     0.0840    3.40  0.00068 ***
## Year2002            0.0561     0.0802    0.70  0.48386
## Year2003            0.1942     0.0787    2.47  0.01363 *
## Year2004            0.1730     0.0901    1.92  0.05504 .
## Year2005            0.3017     0.0817    3.69  0.00023 ***
```

```
## Year2006            0.1207     0.0688    1.75  0.07942 .
## Year2007            0.0202     0.0714    0.28  0.77741
## Year2008            0.0335     0.0713    0.47  0.63830
## Year2009            0.0521     0.0701    0.74  0.45734
## Year2010            0.0661     0.0735    0.90  0.36858
## Year2011            0.1185     0.0716    1.65  0.09809 .
## Year2012           -0.0338     0.0711   -0.48  0.63474
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.726
## Multiple R-squared:  0.0281, Adjusted R-squared:  0.0225
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
##  238 weights are ~= 1. The remaining 2749 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.121   0.874   0.950   0.919   0.985   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          3.35e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Sample size for the above analysis:  2987"
## [1] ""
## [1] ""
## [1] "####################################"
## [1] "Analysis of AJSC 2202"
## [1] "####################################"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 1548 1354 1218 1135 1225 1277 1143 1110 1060 1024 1106 1219 1060 1305 1209
## 2011 2012
## 1208 1105
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  735  661  604  649  651  509  633  557  553  595  631  752  590  780  726
```

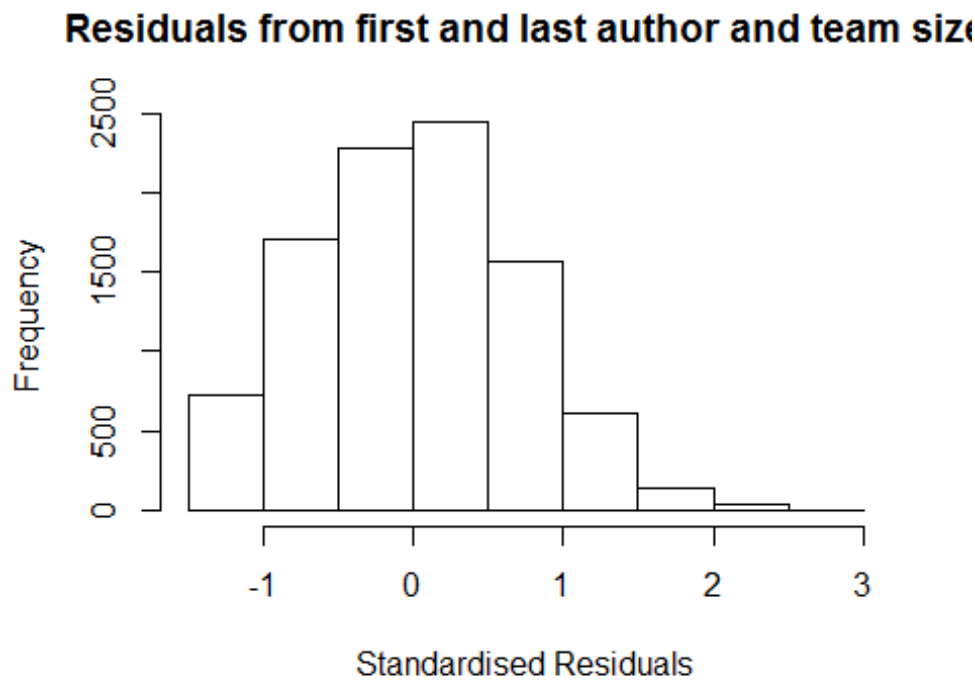
```
## 2011 2012
##  776  698
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  632  571  517  571  557  434  532  458  463  505  554  638  517  659  630
## 2011 2012
##  672  602
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
##  Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 78, df = 16, p-value = 5e-10
```




```
##
##  Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 1.1, df = 1, p-value = 0.3
```

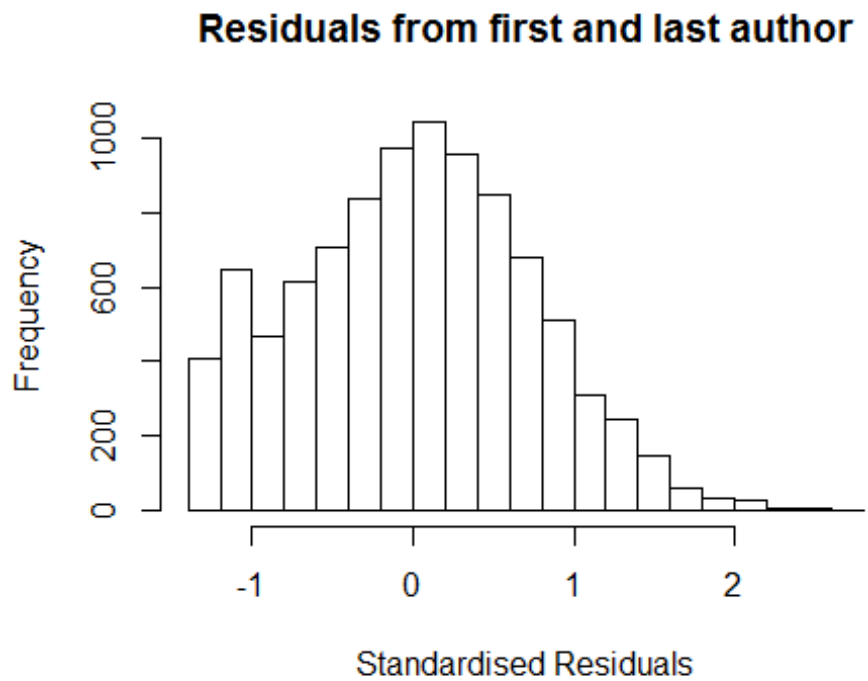



```
## [1] "Female first author team size 2018 geometric mean: 3.03577299387658"
## [1] "Male first author team size 2018 geometric mean: 2.49047895848809"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 17000, p-value = 0.007
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 2.77169349110016"
## [1] "Male last author team size 2018 geometric mean: 2.52641097517993"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 14000, p-value = 0.2
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.148  1           1.071
## LastAuthorFemale  1.146  1           1.070
## UniqueAuthors     1.047  4           1.006
## Year              1.060 16           1.002
```



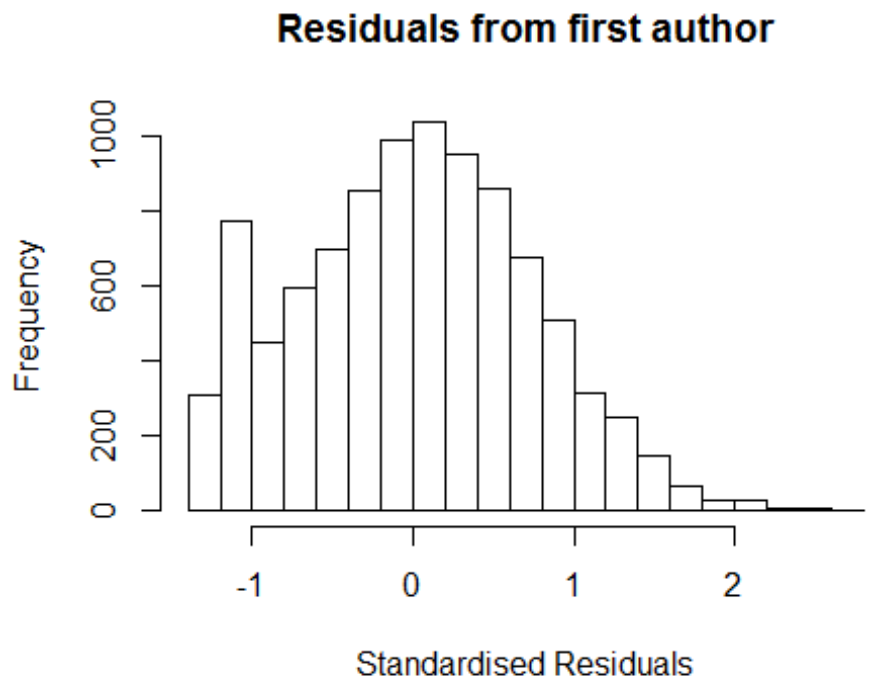

```
## [1] "List of  4 outliers with residuals above  2.5"
##          ScopusId  NLCS Year OneField Fields residuals
## 7928  33845588040 3.436 2001     2202      2     2.565
## 8657   0036591982 3.537 2002     1500      5     2.513
## 9679   1642401353 3.922 2003     2202      2     2.926
## 14577 34347379461 3.514 2007     2202      2     2.584
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
UniqueAuthors +
##     Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##     k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.4948 -0.5116  0.0101  0.4937  2.9260
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)          0.90937    0.03350   27.14  < 2e-16 ***
## FirstAuthorFemale1 -0.11684    0.02731   -4.28  1.9e-05 ***
## LastAuthorFemale1  -0.14221    0.02840   -5.01  5.6e-07 ***
## UniqueAuthors2       0.33420    0.01960   17.05  < 2e-16 ***
## UniqueAuthors3       0.39804    0.02221   17.92  < 2e-16 ***
## UniqueAuthors4       0.37068    0.02904   12.76  < 2e-16 ***
## UniqueAuthors5       0.33545    0.03082   10.89  < 2e-16 ***
```

```
## Year1997            0.07258    0.04450    1.63  0.10297
## Year1998            0.18741    0.04827    3.88  0.00010 ***
## Year1999            0.10268    0.04456    2.30  0.02121 *
## Year2000            0.02339    0.04386    0.53  0.59382
## Year2001           -0.03872    0.05399   -0.72  0.47329
## Year2002            0.11436    0.04793    2.39  0.01705 *
## Year2003            0.08665    0.04654    1.86  0.06265 .
## Year2004            0.09994    0.04508    2.22  0.02664 *
## Year2005            0.07233    0.04463    1.62  0.10511
## Year2006            0.09310    0.04533    2.05  0.04003 *
## Year2007            0.02095    0.04172    0.50  0.61563
## Year2008           -0.00970    0.04448   -0.22  0.82746
## Year2009            0.00787    0.04252    0.19  0.85322
## Year2010           -0.07394    0.04091   -1.81  0.07075 .
## Year2011           -0.11078    0.04047   -2.74  0.00620 **
## Year2012           -0.13870    0.04169   -3.33  0.00088 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.739
## Multiple R-squared:  0.0659, Adjusted R-squared:  0.0638
## Convergence in 10 IRWLS iterations
##
## Robustness weights:
##  798 weights are ~= 1. The remaining 8714 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0811  0.8730  0.9490  0.9140  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.05e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.109  1           1.053
## LastAuthorFemale  1.107  1           1.052
## Year              1.016 16           1.000
```



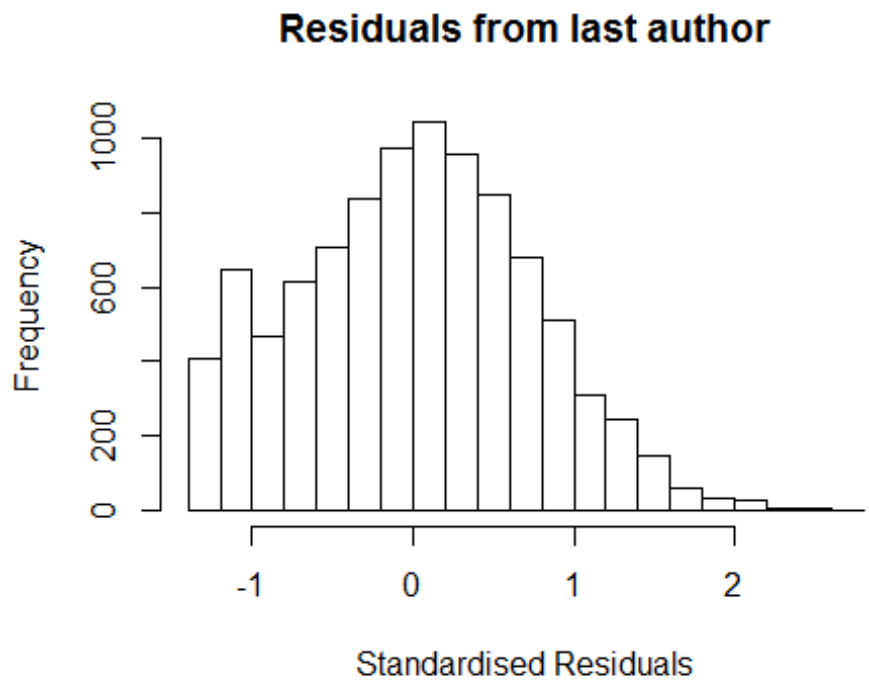

```
## [1] "List of  4 outliers with residuals above  2.5"
##          ScopusId  NLCS Year OneField Fields residuals
## 3403   0000737072 3.862 1998     2202      1     2.516
## 5155   0032593386 3.604 1999     1705      4     2.517
## 9679   1642401353 3.922 2003     2202      2     2.688
## 14101 33646437267 3.788 2006     2202      1     2.520
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.3456 -0.5246  0.0205  0.5058  2.6875
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.14858    0.03223   35.64  < 2e-16 ***
## FirstAuthorFemale1 -0.11365    0.02798   -4.06  4.9e-05 ***
## LastAuthorFemale1  -0.15723    0.02942   -5.34  9.3e-08 ***
## Year1997            0.05195    0.04526    1.15   0.2511
## Year1998            0.19699    0.04984    3.95  7.8e-05 ***
## Year1999            0.09591    0.04522    2.12   0.0340 *
## Year2000            0.04039    0.04443    0.91   0.3633
## Year2001            0.01806    0.05423    0.33   0.7391
## Year2002            0.13789    0.04819    2.86   0.0042 **
```

```
## Year2003            0.08592    0.04724    1.82   0.0690 .
## Year2004            0.12296    0.04648    2.65   0.0082 **
## Year2005            0.10805    0.04618    2.34   0.0193 *
## Year2006            0.11965    0.04649    2.57   0.0101 *
## Year2007            0.05434    0.04306    1.26   0.2070
## Year2008           -0.00565    0.04670   -0.12   0.9037
## Year2009            0.02125    0.04396    0.48   0.6288
## Year2010           -0.04489    0.04283   -1.05   0.2945
## Year2011           -0.07658    0.04204   -1.82   0.0686 .
## Year2012           -0.10300    0.04337   -2.37   0.0176 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.76
## Multiple R-squared:  0.0183, Adjusted R-squared:  0.0165
## Convergence in 10 IRWLS iterations
##
## Robustness weights:
##  809 weights are ~= 1. The remaining 8703 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.185   0.870   0.949   0.914   0.985   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.05e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                    psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##                     GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.008  1           1.004
## Year              1.008 16           1.000
```




```
## [1] "List of  4 outliers with residuals above  2.5"
##          ScopusId  NLCS Year OneField Fields residuals
## 3403   0000737072 3.862 1998     2202      1     2.516
## 5155   0032593386 3.604 1999     1705      4     2.517
## 9679   1642401353 3.922 2003     2202      2     2.688
## 14101 33646437267 3.788 2006     2202      1     2.520
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.3382 -0.5219  0.0215  0.5032  2.6963
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)          1.1413     0.0321   35.54  < 2e-16 ***
## FirstAuthorFemale1  -0.1634     0.0276   -5.93  3.1e-09 ***
## Year1997             0.0526     0.0451    1.17   0.2434
## Year1998             0.1969     0.0499    3.95  7.9e-05 ***
## Year1999             0.0967     0.0453    2.14   0.0326 *
## Year2000             0.0392     0.0444    0.88   0.3771
## Year2001             0.0162     0.0541    0.30   0.7642
## Year2002             0.1365     0.0482    2.83   0.0046 **
## Year2003             0.0844     0.0472    1.79   0.0735 .
```

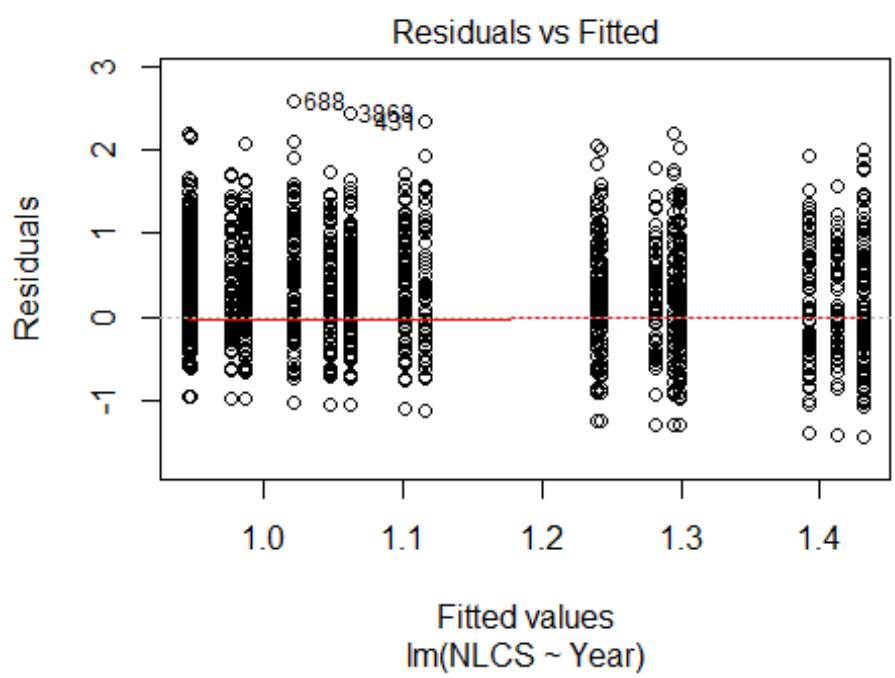
```
## Year2004             0.1257     0.0465    2.70   0.0068 **
## Year2005             0.1073     0.0461    2.33   0.0201 *
## Year2006             0.1198     0.0466    2.57   0.0101 *
## Year2007             0.0543     0.0431    1.26   0.2083
## Year2008            -0.0114     0.0467   -0.24   0.8074
## Year2009             0.0227     0.0440    0.52   0.6060
## Year2010            -0.0481     0.0428   -1.13   0.2605
## Year2011            -0.0788     0.0420   -1.88   0.0608 .
## Year2012            -0.1047     0.0433   -2.42   0.0157 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.76
## Multiple R-squared:  0.0155, Adjusted R-squared:  0.0138
## Convergence in 10 IRWLS iterations
##
## Robustness weights:
##  823 weights are ~= 1. The remaining 8689 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.182   0.868   0.949   0.914   0.985   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.05e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##                   GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.007  1           1.004
## Year             1.007 16           1.000
```



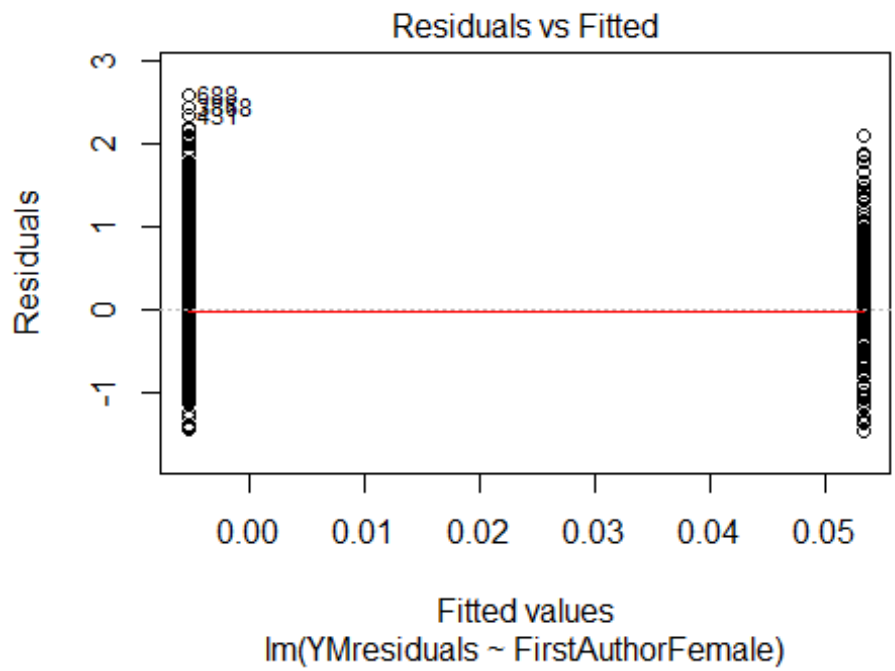

```
## [1] "List of  4 outliers with residuals above  2.5"
##          ScopusId  NLCS Year OneField Fields residuals
## 3403   0000737072 3.862 1998     2202      1     2.516
## 5155   0032593386 3.604 1999     1705      4     2.517
## 9679   1642401353 3.922 2003     2202      2     2.688
## 14101 33646437267 3.788 2006     2202      1     2.520
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.3406 -0.5196  0.0211  0.5051  2.6947
##
## Coefficients:
##                   Estimate Std. Error t value Pr(>|t|)
## (Intercept)        1.14364    0.03223   35.48  < 2e-16 ***
## LastAuthorFemale1 -0.19708    0.02877   -6.85  7.8e-12 ***
## Year1997           0.05227    0.04525    1.16   0.2480
## Year1998           0.19691    0.04987    3.95  7.9e-05 ***
## Year1999           0.09682    0.04532    2.14   0.0327 *
## Year2000           0.03883    0.04448    0.87   0.3826
## Year2001           0.01726    0.05427    0.32   0.7505
## Year2002           0.13663    0.04821    2.83   0.0046 **
## Year2003           0.08366    0.04732    1.77   0.0771 .
```

```
## Year2004           0.12258    0.04657    2.63   0.0085 **
## Year2005           0.10736    0.04626    2.32   0.0203 *
## Year2006           0.11657    0.04654    2.50   0.0123 *
## Year2007           0.05151    0.04307    1.20   0.2317
## Year2008          -0.00937    0.04680   -0.20   0.8413
## Year2009           0.01577    0.04392    0.36   0.7196
## Year2010          -0.04772    0.04294   -1.11   0.2665
## Year2011          -0.07975    0.04213   -1.89   0.0584 .
## Year2012          -0.10851    0.04340   -2.50   0.0124 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.76
## Multiple R-squared:  0.0167, Adjusted R-squared:  0.0149
## Convergence in 10 IRWLS iterations
##
## Robustness weights:
##  796 weights are ~= 1. The remaining 8716 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.183   0.869   0.950   0.915   0.986   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.05e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Sample size for the above analysis:  9512"
## [1] ""
## [1] ""
## [1] "####################################"
## [1] "Analysis of AJSC 2203"
## [1] "####################################"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  186  186  152  242  229  241  162  158  212  185  203  254  255  542  373
## 2011 2012
##  445  394
##
```

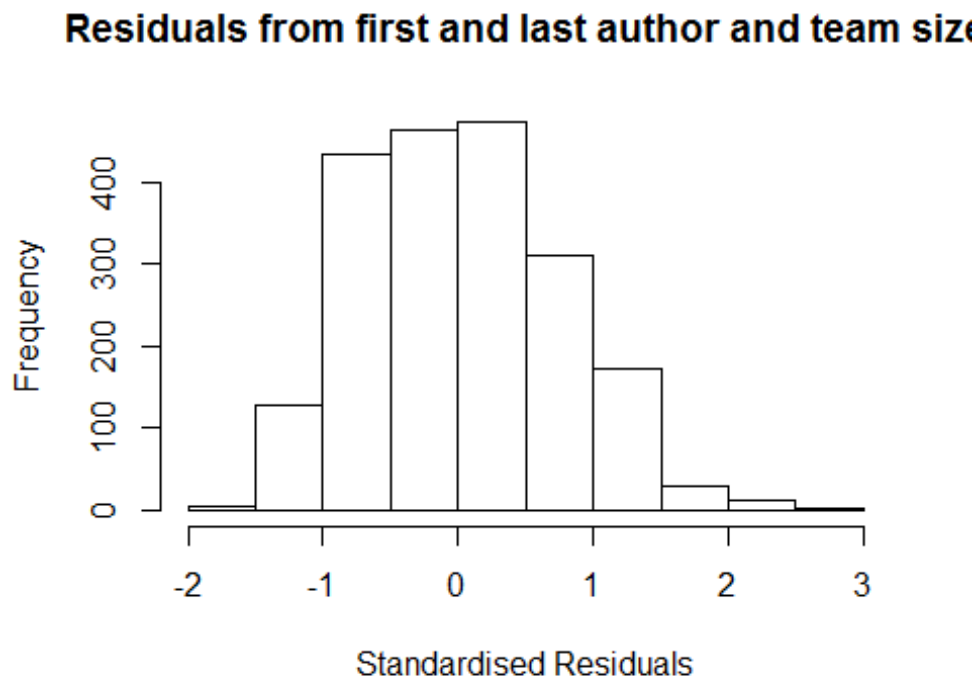
```
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  103  115   81  169  135  104   90   82  101   78   78  133  114  325  219
## 2011 2012
##  256  238
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##   91  103   69  153  120   89   76   68   77   63   58   97   98  273  175
## 2011 2012
##  217  199
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
##  Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 40, df = 16, p-value = 8e-04
```




```
##
##  Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 0.67, df = 1, p-value = 0.4
```

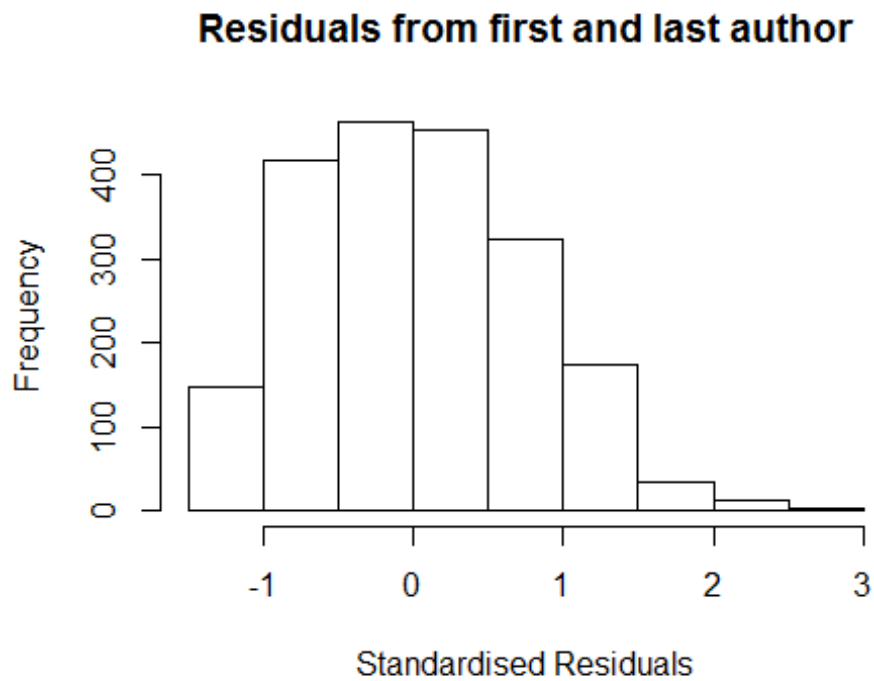



```
## [1] "Female first author team size 2018 geometric mean: 2.96193199953013"
## [1] "Male first author team size 2018 geometric mean: 2.59429069121508"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 2200, p-value = 0.3
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 2.5584763855261"
## [1] "Male last author team size 2018 geometric mean: 2.66793010393555"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 1600, p-value = 0.8
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.128  1           1.062
## LastAuthorFemale  1.126  1           1.061
## UniqueAuthors     1.226  4           1.026
## Year              1.265 16           1.007
```



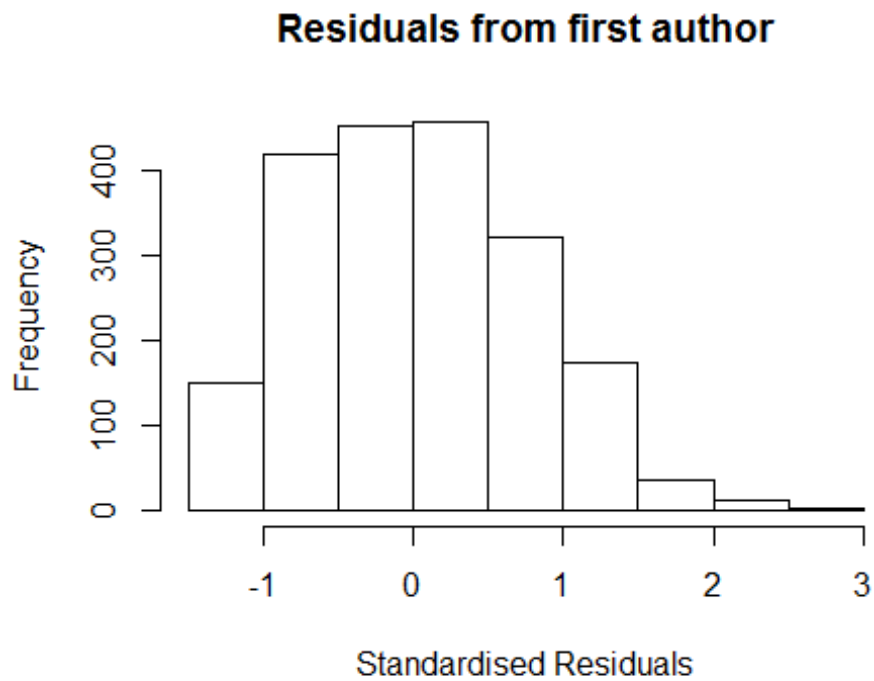

```
## [1] "List of  1 outliers with residuals above  2.5"
##       ScopusId  NLCS Year OneField Fields residuals
## 688 0032593386 3.604 1999     1705      4     2.537
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
UniqueAuthors +
##     Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##     k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.5623 -0.5620 -0.0156  0.5257  2.5375
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)          0.7675     0.0935    8.21  3.9e-16 ***
## FirstAuthorFemale1   0.0229     0.0620    0.37   0.7118
## LastAuthorFemale1   -0.0658     0.0699   -0.94   0.3471
## UniqueAuthors2       0.4005     0.0459    8.72  < 2e-16 ***
## UniqueAuthors3       0.4314     0.0506    8.52  < 2e-16 ***
## UniqueAuthors4       0.5191     0.0652    7.96  2.8e-15 ***
## UniqueAuthors5       0.3751     0.0630    5.96  3.1e-09 ***
## Year1997             0.2572     0.1261    2.04   0.0415 *
## Year1998             0.0340     0.1428    0.24   0.8120
## Year1999            -0.0357     0.1175   -0.30   0.7616
```

```
## Year2000             0.0973     0.1162    0.84   0.4022
## Year2001             0.3587     0.1324    2.71   0.0068 **
## Year2002             0.1971     0.1318    1.50   0.1350
## Year2003             0.1689     0.1273    1.33   0.1846
## Year2004             0.1318     0.1195    1.10   0.2701
## Year2005             0.1967     0.1343    1.46   0.1432
## Year2006             0.2756     0.1303    2.11   0.0346 *
## Year2007             0.1355     0.1195    1.13   0.2569
## Year2008            -0.1824     0.1147   -1.59   0.1118
## Year2009            -0.2097     0.0998   -2.10   0.0359 *
## Year2010            -0.1006     0.1012   -0.99   0.3204
## Year2011            -0.1601     0.0999   -1.60   0.1092
## Year2012            -0.2346     0.1032   -2.27   0.0231 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.778
## Multiple R-squared:  0.099,  Adjusted R-squared:  0.0891
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  174 weights are ~= 1. The remaining 1852 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.266   0.873   0.948   0.917   0.985   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          4.94e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.086  1           1.042
## LastAuthorFemale  1.093  1           1.045
## Year              1.050 16           1.002
```



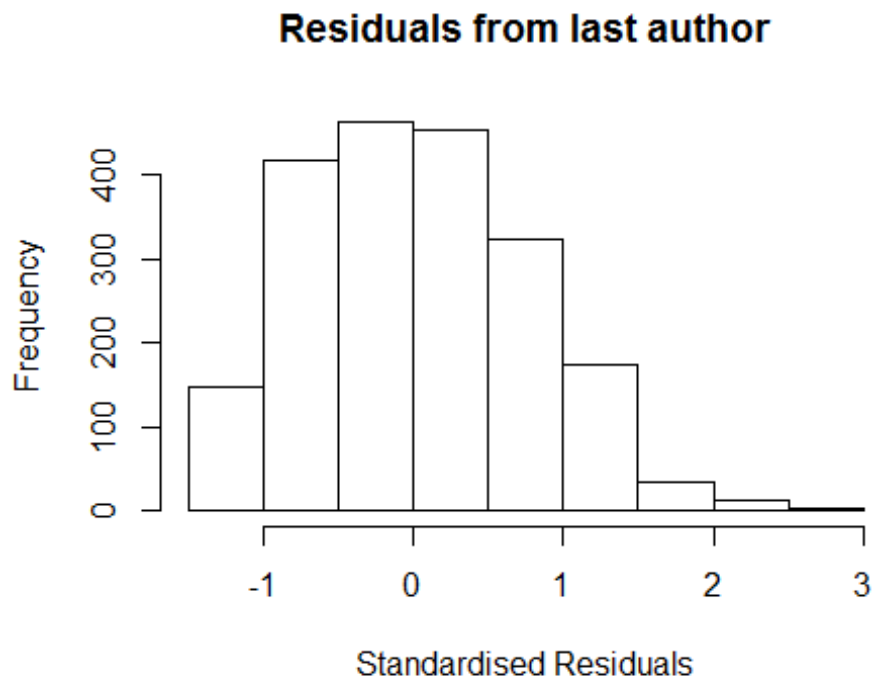

```
## [1] "List of  1 outliers with residuals above  2.5"
##       ScopusId  NLCS Year OneField Fields residuals
## 688 0032593386 3.604 1999     1705      4     2.736
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.4184 -0.5532 -0.0123  0.5369  2.7357
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)           1.0507     0.0875   12.01   <2e-16 ***
## FirstAuthorFemale1    0.0520     0.0643    0.81   0.4191
## LastAuthorFemale1    -0.0647     0.0738   -0.88   0.3802
## Year1997              0.1954     0.1229    1.59   0.1121
## Year1998             -0.0413     0.1454   -0.28   0.7764
## Year1999             -0.1177     0.1187   -0.99   0.3216
## Year2000              0.0130     0.1212    0.11   0.9144
## Year2001              0.3677     0.1353    2.72   0.0066 **
## Year2002              0.2385     0.1281    1.86   0.0627 .
## Year2003              0.1885     0.1312    1.44   0.1510
## Year2004              0.1415     0.1206    1.17   0.2410
## Year2005              0.2510     0.1367    1.84   0.0665 .
```

```
## Year2006             0.3110     0.1304    2.38   0.0172 *
## Year2007             0.1884     0.1210    1.56   0.1197
## Year2008            -0.1630     0.1163   -1.40   0.1613
## Year2009            -0.1465     0.0976   -1.50   0.1336
## Year2010            -0.0556     0.1013   -0.55   0.5830
## Year2011            -0.1029     0.0993   -1.04   0.3003
## Year2012            -0.1641     0.1016   -1.62   0.1062
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.801
## Multiple R-squared:  0.0445, Adjusted R-squared:  0.0359
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
##  170 weights are ~= 1. The remaining 1856 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.220   0.878   0.951   0.917   0.985   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          4.94e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.021  1           1.010
## Year              1.021 16           1.001
```




```
## [1] "List of  1 outliers with residuals above  2.5"
##       ScopusId  NLCS Year OneField Fields residuals
## 688 0032593386 3.604 1999     1705      4     2.736
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.4128 -0.5572 -0.0096  0.5402  2.6727
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)           1.0479     0.0875   11.97   <2e-16 ***
## FirstAuthorFemale1    0.0350     0.0633    0.55    0.580
## Year1997              0.1924     0.1230    1.56    0.118
## Year1998             -0.0393     0.1454   -0.27    0.787
## Year1999             -0.1166     0.1186   -0.98    0.326
## Year2000              0.0121     0.1213    0.10    0.921
## Year2001              0.3649     0.1353    2.70    0.007 **
## Year2002              0.2377     0.1280    1.86    0.063 .
## Year2003              0.1876     0.1315    1.43    0.154
## Year2004              0.1430     0.1207    1.18    0.237
## Year2005              0.2527     0.1369    1.85    0.065 .
## Year2006              0.3118     0.1307    2.39    0.017 *
```

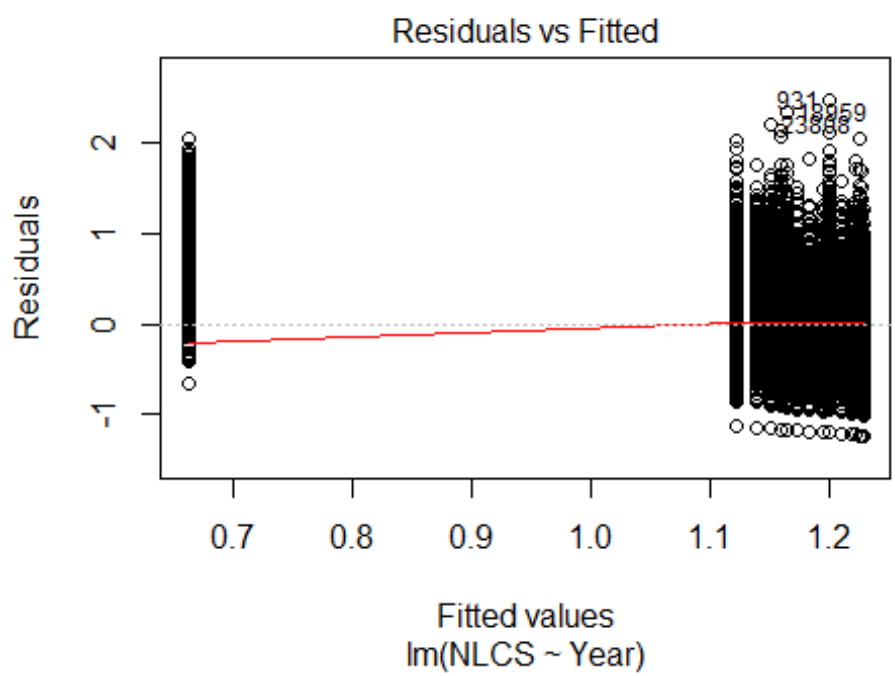
```
## Year2007             0.1868     0.1210    1.54    0.123
## Year2008            -0.1640     0.1163   -1.41    0.159
## Year2009            -0.1479     0.0978   -1.51    0.131
## Year2010            -0.0541     0.1014   -0.53    0.594
## Year2011            -0.1035     0.0995   -1.04    0.298
## Year2012            -0.1668     0.1016   -1.64    0.101
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.802
## Multiple R-squared:  0.0439, Adjusted R-squared:  0.0359
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  169 weights are ~= 1. The remaining 1857 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.244   0.878   0.951   0.917   0.985   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          4.94e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##                   GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.026  1           1.013
## Year             1.026 16           1.001
```



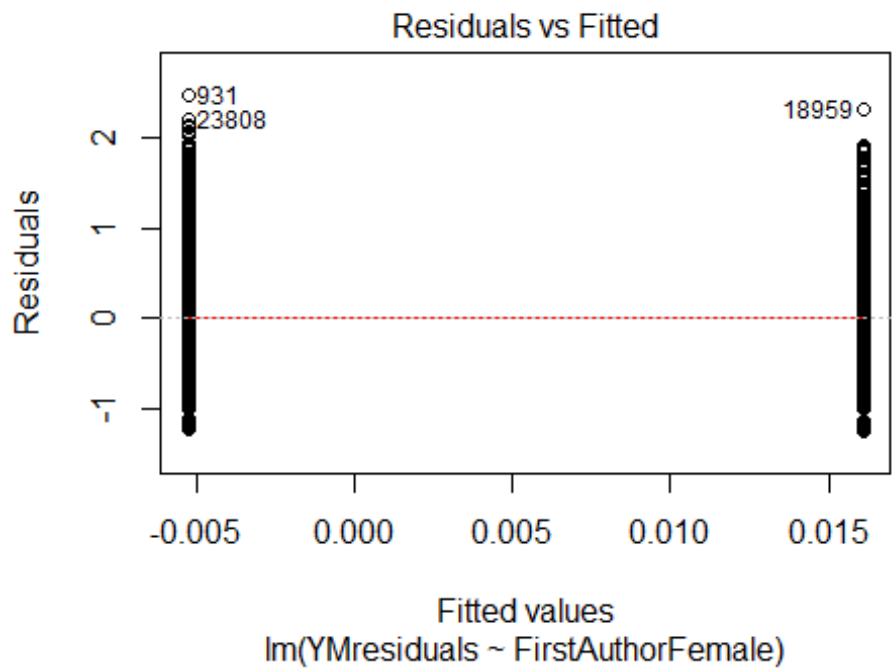

```
## [1] "List of  1 outliers with residuals above  2.5"
##       ScopusId  NLCS Year OneField Fields residuals
## 688 0032593386 3.604 1999     1705      4     2.736
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.4201 -0.5560 -0.0106  0.5374  2.7184
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)          1.0537     0.0873   12.08   <2e-16 ***
## LastAuthorFemale1   -0.0487     0.0707   -0.69   0.4908
## Year1997             0.1980     0.1228    1.61   0.1070
## Year1998            -0.0398     0.1456   -0.27   0.7847
## Year1999            -0.1193     0.1185   -1.01   0.3141
## Year2000             0.0119     0.1211    0.10   0.9216
## Year2001             0.3664     0.1352    2.71   0.0068 **
## Year2002             0.2388     0.1279    1.87   0.0620 .
## Year2003             0.1897     0.1309    1.45   0.1476
## Year2004             0.1430     0.1204    1.19   0.2354
## Year2005             0.2524     0.1364    1.85   0.0644 .
## Year2006             0.3118     0.1309    2.38   0.0173 *
```

```
## Year2007            0.1881     0.1210    1.55   0.1204
## Year2008           -0.1590     0.1162   -1.37   0.1712
## Year2009           -0.1462     0.0975   -1.50   0.1338
## Year2010           -0.0535     0.1012   -0.53   0.5969
## Year2011           -0.1032     0.0992   -1.04   0.2982
## Year2012           -0.1635     0.1014   -1.61   0.1072
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.801
## Multiple R-squared:  0.0442, Adjusted R-squared:  0.0361
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
##  159 weights are ~= 1. The remaining 1867 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.226   0.880   0.951   0.917   0.986   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          4.94e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Sample size for the above analysis:  2026"
## [1] ""
## [1] ""
## [1] "####################################"
## [1] "Analysis of AJSC 2204"
## [1] "####################################"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  915  956 1050  887 1852  948 1013  970 1167 1392 1569 2015 2287 2580 2812
## 2011 2012
## 2938 3033
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  580  534  568  497 1137  403  572  593  710  851  936 1214 1410 1480 1698
## 2011 2012
```

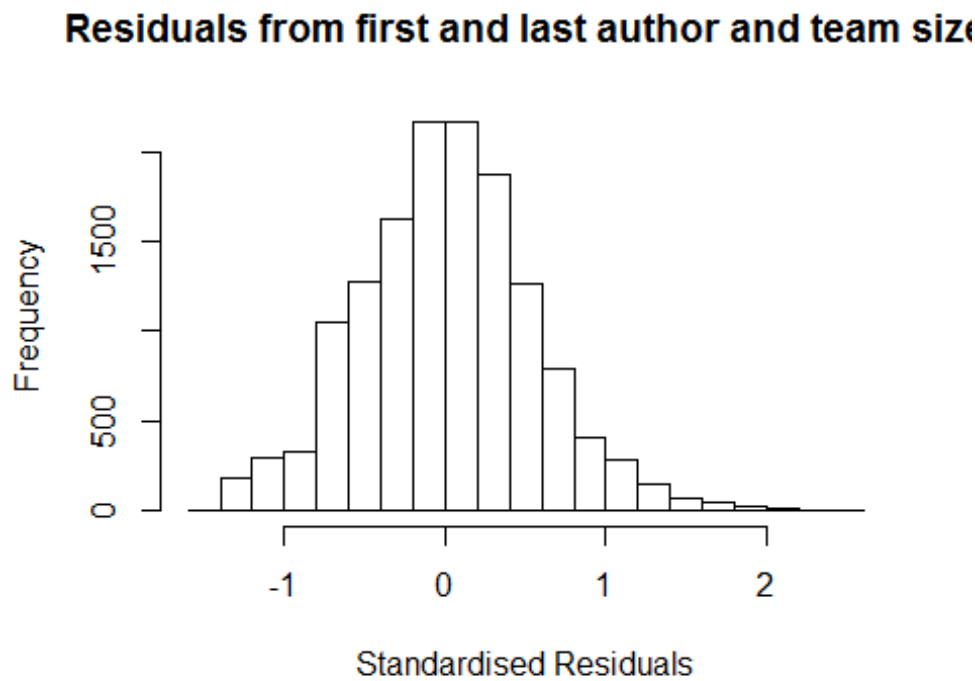
```
## 1775 1821
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  519  462  485  430  973  353  492  497  584  700  778 1013 1148 1200 1384
## 2011 2012
## 1463 1487
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
##  Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 580, df = 16, p-value <2e-16
```




```
##
##  Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 24, df = 1, p-value = 9e-07
```

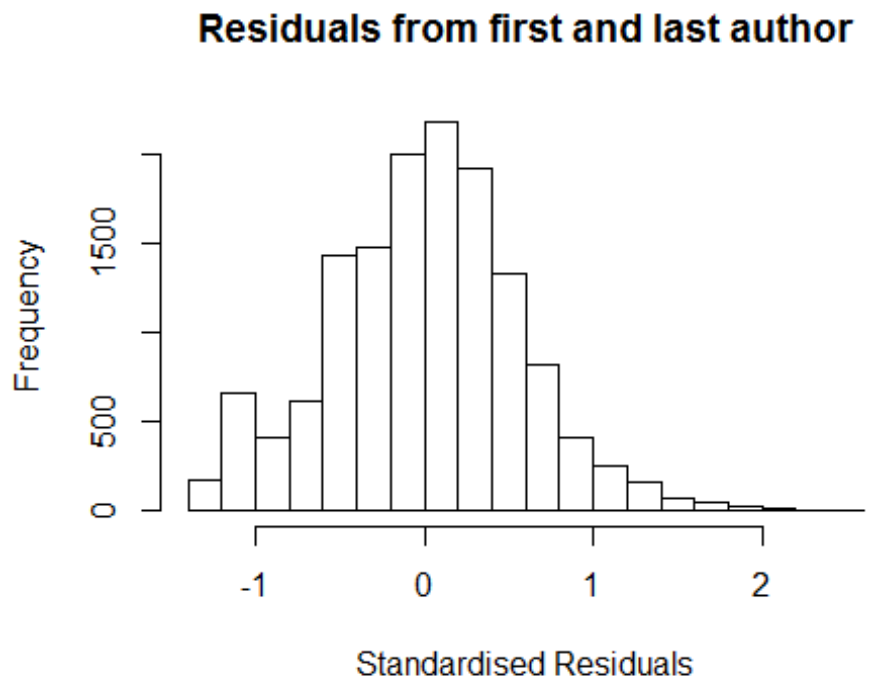



```
## [1] "Female first author team size 2018 geometric mean: 4.66078368238678"
## [1] "Male first author team size 2018 geometric mean: 4.35496402859526"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 280000, p-value = 0.03
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 4.2906857418799"
## [1] "Male last author team size 2018 geometric mean: 4.49550699158764"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 2e+05, p-value = 0.2
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##                     GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.037  1           1.018
## LastAuthorFemale  1.029  1           1.014
## UniqueAuthors     1.091  4           1.011
## Year              1.112 16           1.003
```



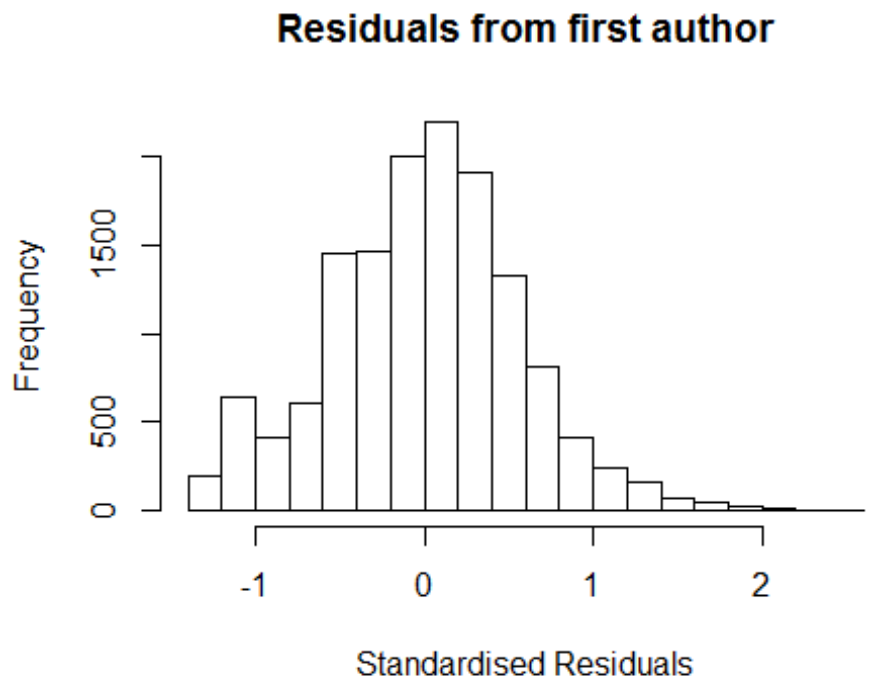

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
UniqueAuthors +
##     Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##     k.max = 1000))
##  \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.42304 -0.34460  0.00818  0.34300  2.46343
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)          0.7751     0.0328   23.63  < 2e-16 ***
## FirstAuthorFemale1   0.0158     0.0103    1.53   0.1253
## LastAuthorFemale1   -0.0169     0.0127   -1.33   0.1822
## UniqueAuthors2       0.4365     0.0217   20.11  < 2e-16 ***
## UniqueAuthors3       0.5000     0.0209   23.91  < 2e-16 ***
## UniqueAuthors4       0.5382     0.0212   25.38  < 2e-16 ***
## UniqueAuthors5       0.6088     0.0195   31.26  < 2e-16 ***
## Year1997             0.0392     0.0413    0.95   0.3433
## Year1998            -0.0764     0.0424   -1.80   0.0720 .
## Year1999            -0.0751     0.0391   -1.92   0.0549 .
```

```
## Year2000            -0.7435     0.0437  -16.99  < 2e-16 ***
## Year2001            -0.0085     0.0416   -0.20   0.8381
## Year2002            -0.0350     0.0382   -0.92   0.3590
## Year2003            -0.0160     0.0371   -0.43   0.6660
## Year2004            -0.0652     0.0354   -1.84   0.0659 .
## Year2005            -0.0698     0.0343   -2.04   0.0418 *
## Year2006            -0.0795     0.0344   -2.31   0.0207 *
## Year2007            -0.0462     0.0341   -1.36   0.1753
## Year2008            -0.1019     0.0335   -3.04   0.0024 **
## Year2009            -0.0832     0.0326   -2.56   0.0106 *
## Year2010            -0.1293     0.0326   -3.97  7.4e-05 ***
## Year2011            -0.1270     0.0323   -3.93  8.6e-05 ***
## Year2012            -0.1674     0.0328   -5.10  3.5e-07 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.488
## Multiple R-squared:  0.187,  Adjusted R-squared:  0.186
## Convergence in 17 IRWLS iterations
##
## Robustness weights:
##  2 observations c(472,1921) are outliers with |weight| = 0 ( < 7.2e-06);
##  1212 weights are ~= 1. The remaining 12754 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0008  0.8500  0.9450  0.8900  0.9840  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          7.16e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.019  1           1.009
## LastAuthorFemale  1.011  1           1.006
## Year              1.023 16           1.001
```



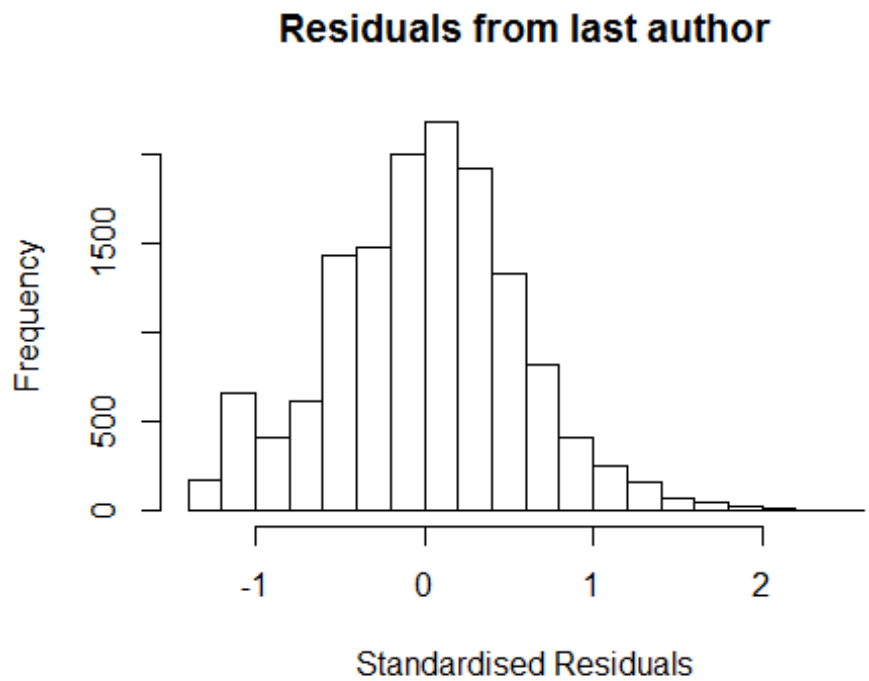

```
## [1] "List of  1 outliers with residuals above  2.5"
##       ScopusId  NLCS Year OneField Fields residuals
## 931 0029670262 3.675 1996     1305      6      2.51
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.2689 -0.3666  0.0209  0.3543  2.5102
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)          1.16480    0.03283   35.48   <2e-16 ***
## FirstAuthorFemale1   0.03963    0.01065    3.72   0.0002 ***
## LastAuthorFemale1   -0.02933    0.01331   -2.20   0.0276 *
## Year1997             0.06361    0.04497    1.41   0.1573
## Year1998            -0.01925    0.04574   -0.42   0.6739
## Year1999            -0.00631    0.04235   -0.15   0.8815
## Year2000            -0.70718    0.04603  -15.37   <2e-16 ***
## Year2001             0.05199    0.04525    1.15   0.2506
## Year2002             0.05370    0.04129    1.30   0.1935
## Year2003             0.08152    0.03962    2.06   0.0397 *
## Year2004             0.02669    0.03859    0.69   0.4891
## Year2005             0.04964    0.03727    1.33   0.1829
```

```
## Year2006            0.01020    0.03727    0.27   0.7843
## Year2007            0.06451    0.03711    1.74   0.0822 .
## Year2008            0.00584    0.03701    0.16   0.8746
## Year2009            0.03370    0.03567    0.94   0.3448
## Year2010           -0.00631    0.03577   -0.18   0.8600
## Year2011           -0.01671    0.03588   -0.47   0.6414
## Year2012           -0.04638    0.03610   -1.28   0.1989
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.496
## Multiple R-squared:  0.0992, Adjusted R-squared:  0.098
## Convergence in 16 IRWLS iterations
##
## Robustness weights:
##  observation 472 is an outlier with |weight| = 0 ( < 7.2e-06);
##  1112 weights are ~= 1. The remaining 12855 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0011  0.8520  0.9430  0.8840  0.9840  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          7.16e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.016  1           1.008
## Year              1.016 16           1.001
```




```
## [1] "List of  1 outliers with residuals above  2.5"
##       ScopusId  NLCS Year OneField Fields residuals
## 931 0029670262 3.675 1996     1305      6      2.51
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.2785 -0.3666  0.0203  0.3545  2.5127
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)          1.16228    0.03283   35.40  < 2e-16 ***
## FirstAuthorFemale1   0.03580    0.01071    3.34  0.00083 ***
## Year1997             0.06369    0.04503    1.41  0.15732
## Year1998            -0.01932    0.04573   -0.42  0.67268
## Year1999            -0.00653    0.04236   -0.15  0.87752
## Year2000            -0.70743    0.04606  -15.36  < 2e-16 ***
## Year2001             0.05087    0.04520    1.13  0.26041
## Year2002             0.05347    0.04130    1.29  0.19546
## Year2003             0.08045    0.03962    2.03  0.04234 *
## Year2004             0.02546    0.03858    0.66  0.50926
## Year2005             0.04816    0.03726    1.29  0.19618
## Year2006             0.00954    0.03726    0.26  0.79799
```

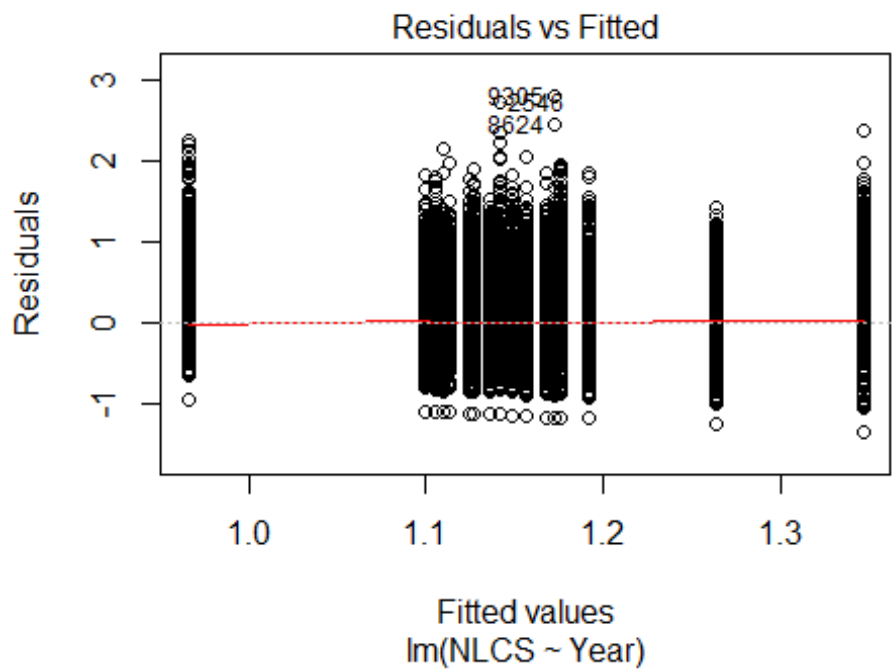
```
## Year2007            0.06335    0.03711    1.71  0.08780 .
## Year2008            0.00478    0.03700    0.13  0.89723
## Year2009            0.03231    0.03566    0.91  0.36498
## Year2010           -0.00824    0.03574   -0.23  0.81765
## Year2011           -0.01846    0.03586   -0.51  0.60684
## Year2012           -0.04786    0.03609   -1.33  0.18475
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.496
## Multiple R-squared:  0.0989, Adjusted R-squared:  0.0978
## Convergence in 16 IRWLS iterations
##
## Robustness weights:
##  observation 472 is an outlier with |weight| = 0 ( < 7.2e-06);
##  1098 weights are ~= 1. The remaining 12869 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0031  0.8520  0.9430  0.8840  0.9840  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          7.16e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##                   GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.009  1           1.005
## Year             1.009 16           1.000
```



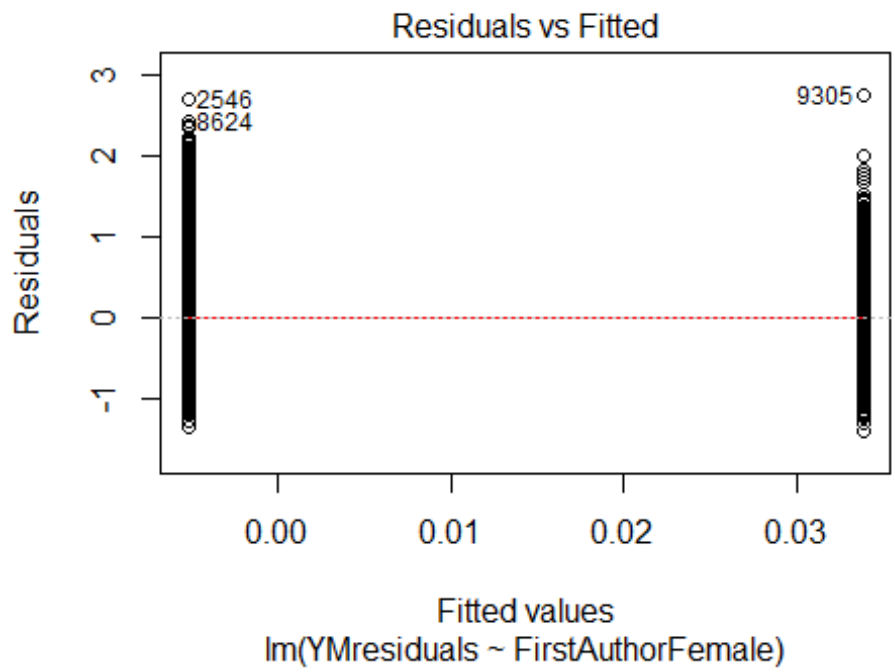

```
## [1] "List of  1 outliers with residuals above  2.5"
##       ScopusId  NLCS Year OneField Fields residuals
## 931 0029670262 3.675 1996     1305      6      2.51
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.2546 -0.3685  0.0226  0.3542  2.5046
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.17044    0.03294   35.53   <2e-16 ***
## LastAuthorFemale1 -0.02185    0.01332   -1.64    0.101
## Year1997            0.06352    0.04513    1.41    0.159
## Year1998           -0.01885    0.04585   -0.41    0.681
## Year1999           -0.00427    0.04250   -0.10    0.920
## Year2000           -0.70487    0.04612  -15.28   <2e-16 ***
## Year2001            0.05340    0.04531    1.18    0.239
## Year2002            0.05537    0.04142    1.34    0.181
## Year2003            0.08417    0.03975    2.12    0.034 *
## Year2004            0.03193    0.03867    0.83    0.409
## Year2005            0.05268    0.03739    1.41    0.159
## Year2006            0.01302    0.03743    0.35    0.728
```

```
## Year2007           0.06710    0.03724    1.80    0.072 .
## Year2008           0.00987    0.03711    0.27    0.790
## Year2009           0.03863    0.03578    1.08    0.280
## Year2010          -0.00220    0.03587   -0.06    0.951
## Year2011          -0.01168    0.03599   -0.32    0.745
## Year2012          -0.04159    0.03621   -1.15    0.251
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.497
## Multiple R-squared:  0.0983, Adjusted R-squared:  0.0972
## Convergence in 16 IRWLS iterations
##
## Robustness weights:
##  observation 472 is an outlier with |weight| = 0 ( < 7.2e-06);
##  1110 weights are ~= 1. The remaining 12857 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0022  0.8520  0.9420  0.8840  0.9840  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          7.16e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Sample size for the above analysis:  13968"
## [1] ""
## [1] ""
## [1] "####################################"
## [1] "Analysis of AJSC 2205"
## [1] "####################################"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 1532 1215 1162 1393 1096 1162 1045 1035 1074 1126 1256 1259 1368 1515 1562
## 2011 2012
## 1468 1660
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  858  679  704  899  579  490  618  590  632  607  693  717  761  815  885
```

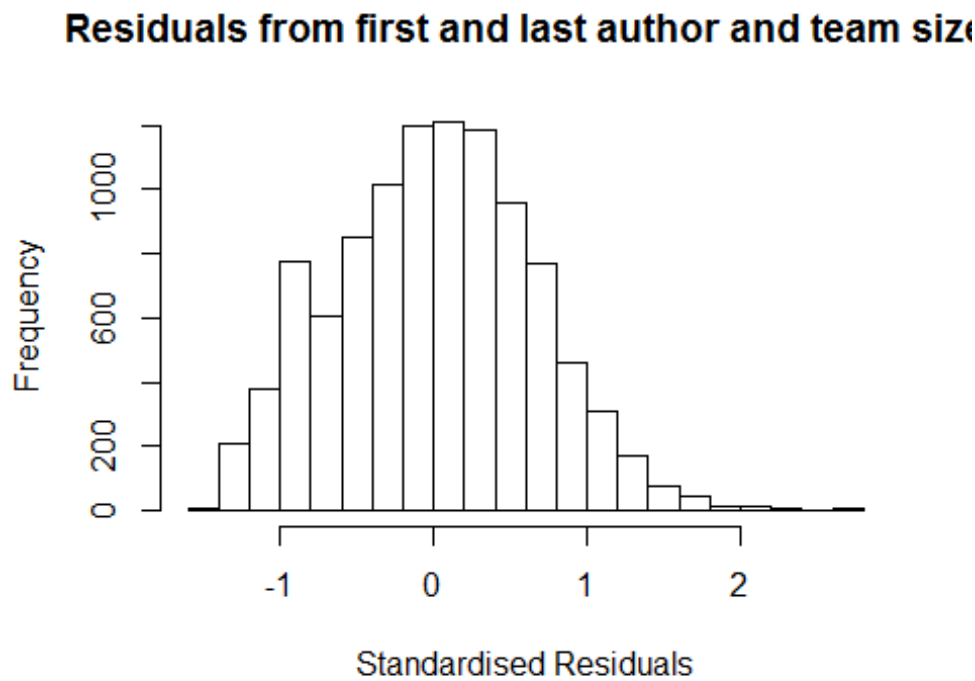
```
## 2011 2012
##  858  934
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  744  585  611  780  482  414  501  488  516  494  579  599  625  647  723
## 2011 2012
##  697  768
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
##  Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 280, df = 16, p-value <2e-16
```




```
##
##  Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 2.3, df = 1, p-value = 0.1
```

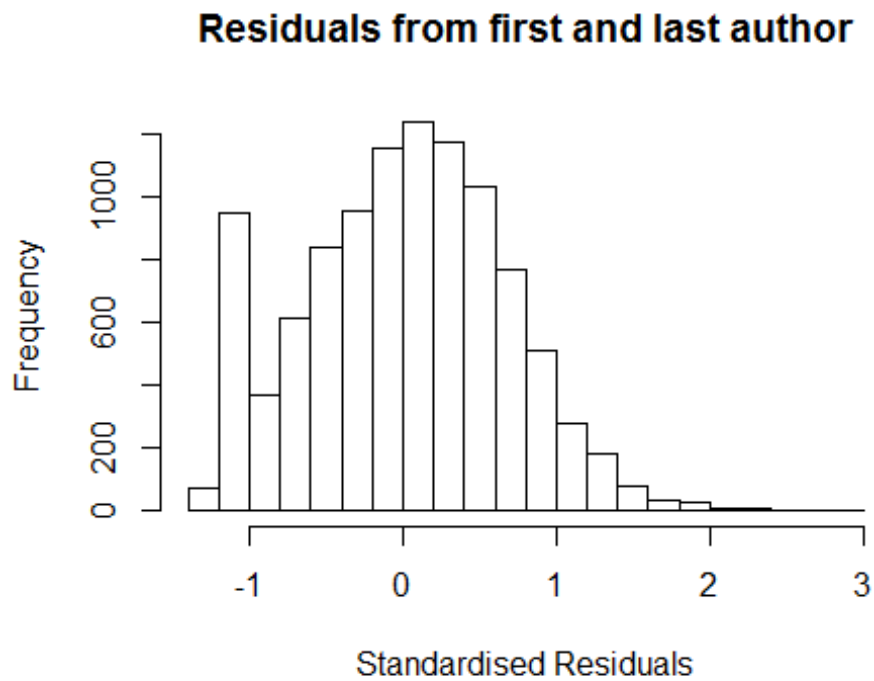



```
## [1] "Female first author team size 2018 geometric mean: 2.84237370055861"
## [1] "Male first author team size 2018 geometric mean: 2.64026008535671"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 95000, p-value = 0.06
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 2.80045199862022"
## [1] "Male last author team size 2018 geometric mean: 2.64807711572491"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 94000, p-value = 0.1
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.054  1           1.026
## LastAuthorFemale  1.045  1           1.022
## UniqueAuthors     1.070  4           1.009
## Year              1.073 16           1.002
```



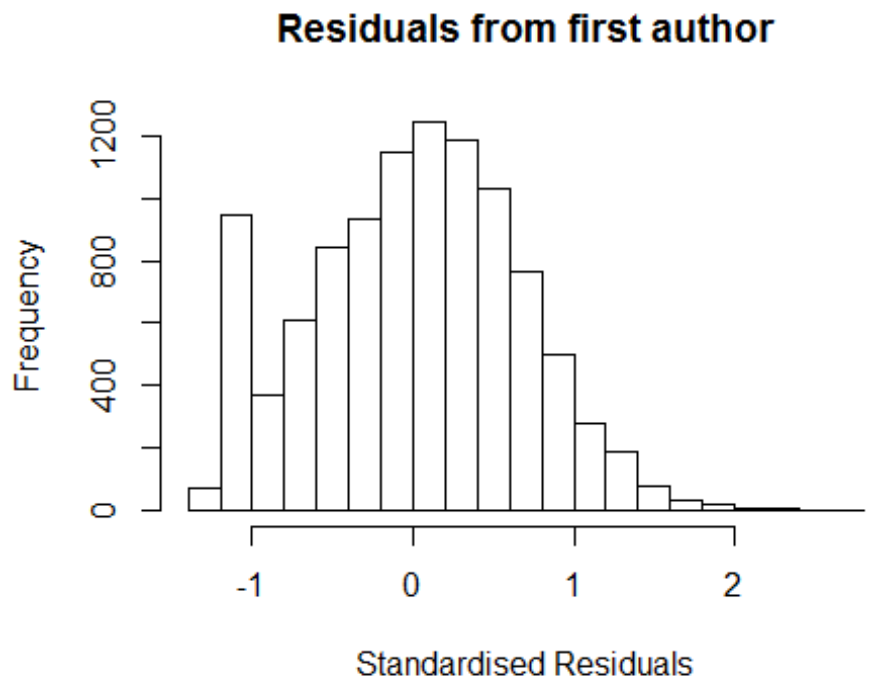

```
## [1] "List of  3 outliers with residuals above  2.5"
##        ScopusId  NLCS Year OneField Fields residuals
## 1680 0031097861 3.492 1997     2205      4     2.612
## 2546 0031233750 3.858 1997     2205      3     2.656
## 9305 0036504271 3.965 2002     2205      4     2.742
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
UniqueAuthors +
##     Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##     k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.4198 -0.4640  0.0127  0.4462  2.7420
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)          0.6886     0.0321   21.46  < 2e-16 ***
## FirstAuthorFemale1   0.0211     0.0201    1.05  0.29359
## LastAuthorFemale1   -0.0318     0.0219   -1.45  0.14684
## UniqueAuthors2       0.3535     0.0190   18.64  < 2e-16 ***
## UniqueAuthors3       0.3747     0.0203   18.43  < 2e-16 ***
## UniqueAuthors4       0.3638     0.0248   14.64  < 2e-16 ***
## UniqueAuthors5       0.3606     0.0279   12.94  < 2e-16 ***
## Year1997             0.1915     0.0454    4.22  2.5e-05 ***
```

```
## Year1998             0.1725     0.0437    3.95  7.8e-05 ***
## Year1999             0.1474     0.0405    3.64  0.00027 ***
## Year2000             0.1717     0.0442    3.88  0.00010 ***
## Year2001             0.3565     0.0480    7.43  1.1e-13 ***
## Year2002             0.1845     0.0436    4.23  2.3e-05 ***
## Year2003             0.1790     0.0446    4.01  6.0e-05 ***
## Year2004             0.1670     0.0394    4.24  2.2e-05 ***
## Year2005             0.2814     0.0410    6.87  6.8e-12 ***
## Year2006             0.1788     0.0385    4.65  3.4e-06 ***
## Year2007             0.1209     0.0390    3.10  0.00192 **
## Year2008             0.1174     0.0391    3.00  0.00270 **
## Year2009             0.1316     0.0392    3.36  0.00079 ***
## Year2010             0.1347     0.0377    3.57  0.00035 ***
## Year2011             0.1083     0.0387    2.80  0.00515 **
## Year2012             0.0959     0.0372    2.58  0.00998 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.658
## Multiple R-squared:  0.0664, Adjusted R-squared:  0.0644
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
##  863 weights are ~= 1. The remaining 9390 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0436  0.8660  0.9480  0.9120  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          9.75e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##                     GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.041  1           1.021
## LastAuthorFemale  1.030  1           1.015
## Year              1.022 16           1.001
```



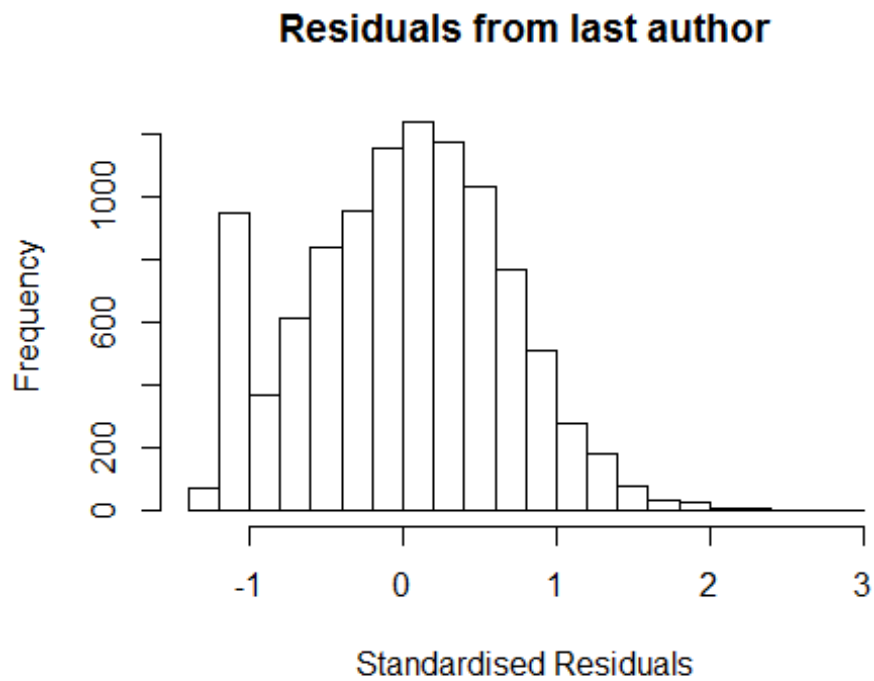

```
## [1] "List of  2 outliers with residuals above  2.5"
##        ScopusId  NLCS Year OneField Fields residuals
## 2546 0031233750 3.858 1997     2205      3     2.774
## 9305 0036504271 3.965 2002     2205      4     2.810
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.3363 -0.4603  0.0307  0.4564  2.8103
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)           0.9049     0.0308   29.40  < 2e-16 ***
## FirstAuthorFemale1    0.0490     0.0202    2.43    0.015 *
## LastAuthorFemale1    -0.0204     0.0219   -0.93    0.351
## Year1997              0.1998     0.0468    4.27  2.0e-05 ***
## Year1998              0.1979     0.0451    4.38  1.2e-05 ***
## Year1999              0.1870     0.0408    4.58  4.7e-06 ***
## Year2000              0.1962     0.0464    4.23  2.4e-05 ***
## Year2001              0.3824     0.0495    7.73  1.2e-14 ***
## Year2002              0.2212     0.0452    4.90  9.8e-07 ***
## Year2003              0.2125     0.0455    4.67  3.1e-06 ***
## Year2004              0.2375     0.0402    5.91  3.6e-09 ***
```

```
## Year2005             0.3580     0.0415    8.62  < 2e-16 ***
## Year2006             0.2359     0.0392    6.02  1.8e-09 ***
## Year2007             0.1826     0.0397    4.60  4.3e-06 ***
## Year2008             0.1745     0.0400    4.36  1.3e-05 ***
## Year2009             0.2056     0.0399    5.15  2.6e-07 ***
## Year2010             0.2142     0.0384    5.57  2.5e-08 ***
## Year2011             0.1774     0.0394    4.51  6.6e-06 ***
## Year2012             0.1929     0.0377    5.11  3.2e-07 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.677
## Multiple R-squared:  0.0131, Adjusted R-squared:  0.0114
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
##  879 weights are ~= 1. The remaining 9374 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0458  0.8690  0.9490  0.9120  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          9.75e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.017  1           1.009
## Year              1.017 16           1.001
```




```
## [1] "List of  2 outliers with residuals above  2.5"
##        ScopusId  NLCS Year OneField Fields residuals
## 2546 0031233750 3.858 1997     2205      3     2.774
## 9305 0036504271 3.965 2002     2205      4     2.810
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.3303 -0.4594  0.0313  0.4562  2.7961
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)           0.9037     0.0310   29.16  < 2e-16 ***
## FirstAuthorFemale1    0.0444     0.0200    2.22    0.027 *
## Year1997              0.2001     0.0469    4.27  2.0e-05 ***
## Year1998              0.1977     0.0452    4.37  1.3e-05 ***
## Year1999              0.1860     0.0410    4.54  5.7e-06 ***
## Year2000              0.1953     0.0466    4.19  2.8e-05 ***
## Year2001              0.3822     0.0497    7.69  1.6e-14 ***
## Year2002              0.2208     0.0453    4.87  1.1e-06 ***
## Year2003              0.2120     0.0457    4.64  3.5e-06 ***
## Year2004              0.2369     0.0404    5.86  4.7e-09 ***
## Year2005              0.3574     0.0418    8.56  < 2e-16 ***
```

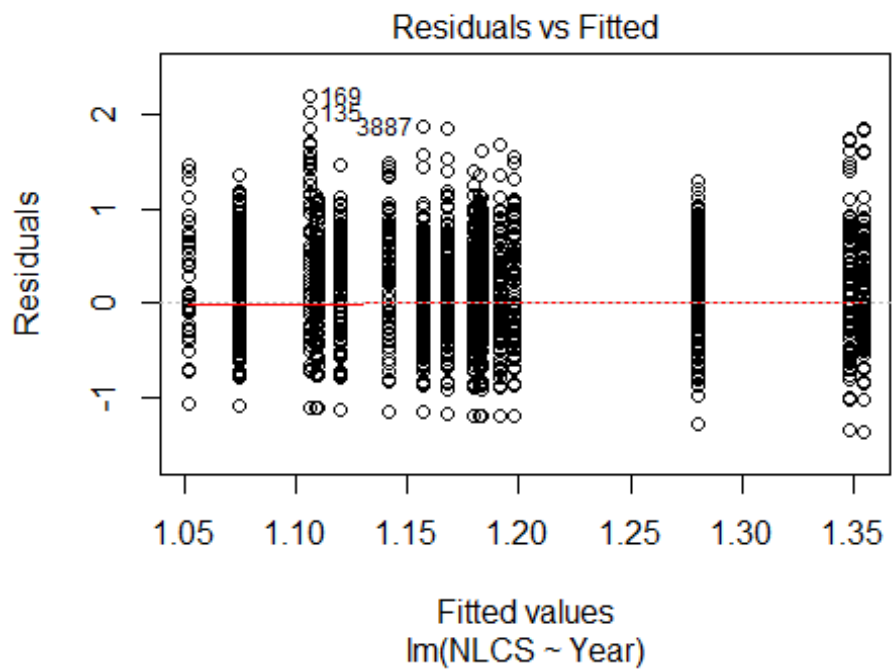
```
## Year2006             0.2350     0.0394    5.97  2.5e-09 ***
## Year2007             0.1818     0.0399    4.55  5.4e-06 ***
## Year2008             0.1741     0.0402    4.33  1.5e-05 ***
## Year2009             0.2053     0.0401    5.12  3.1e-07 ***
## Year2010             0.2136     0.0386    5.53  3.3e-08 ***
## Year2011             0.1772     0.0396    4.48  7.7e-06 ***
## Year2012             0.1919     0.0379    5.06  4.2e-07 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.678
## Multiple R-squared:  0.013,  Adjusted R-squared:  0.0113
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
##  868 weights are ~= 1. The remaining 9385 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0508  0.8700  0.9490  0.9120  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          9.75e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##                   GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.006  1           1.003
## Year             1.006 16           1.000
```



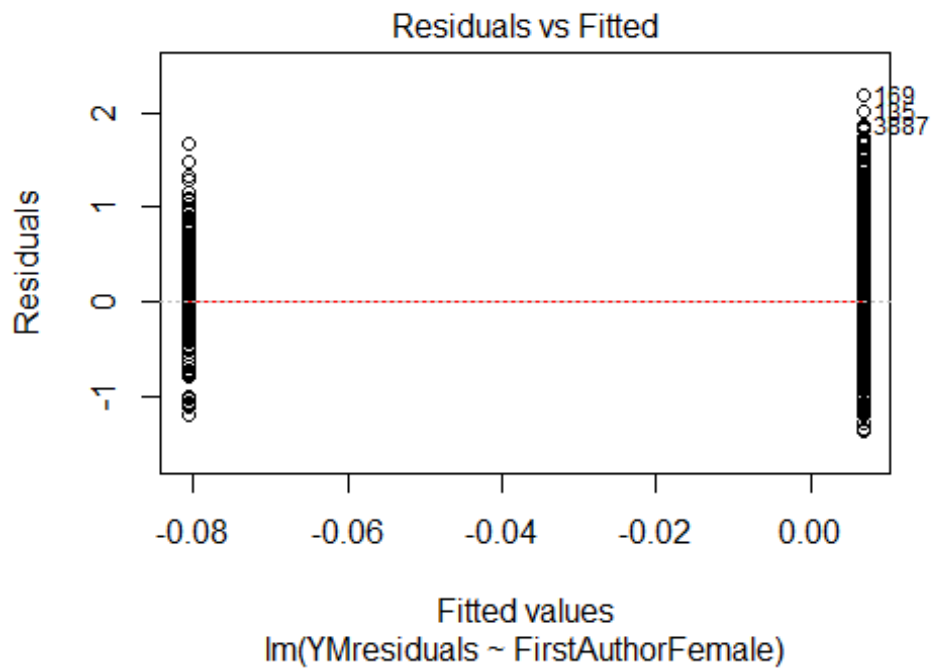

```
## [1] "List of  2 outliers with residuals above  2.5"
##        ScopusId  NLCS Year OneField Fields residuals
## 2546 0031233750 3.858 1997     2205      3     2.774
## 9305 0036504271 3.965 2002     2205      4     2.810
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.2927 -0.4629  0.0302  0.4588  2.8436
##
## Coefficients:
##                   Estimate Std. Error t value Pr(>|t|)
## (Intercept)        0.90832    0.03079   29.50  < 2e-16 ***
## LastAuthorFemale1 -0.00756    0.02170   -0.35     0.73
## Year1997           0.20062    0.04689    4.28  1.9e-05 ***
## Year1998           0.19811    0.04517    4.39  1.2e-05 ***
## Year1999           0.18752    0.04083    4.59  4.4e-06 ***
## Year2000           0.19744    0.04647    4.25  2.2e-05 ***
## Year2001           0.38442    0.04947    7.77  8.6e-15 ***
## Year2002           0.22060    0.04525    4.88  1.1e-06 ***
## Year2003           0.21313    0.04558    4.68  3.0e-06 ***
## Year2004           0.23830    0.04022    5.92  3.2e-09 ***
## Year2005           0.35999    0.04154    8.67  < 2e-16 ***
```

```
## Year2006           0.23947    0.03918    6.11  1.0e-09 ***
## Year2007           0.18443    0.03975    4.64  3.5e-06 ***
## Year2008           0.17647    0.04000    4.41  1.0e-05 ***
## Year2009           0.20640    0.03991    5.17  2.4e-07 ***
## Year2010           0.21736    0.03843    5.66  1.6e-08 ***
## Year2011           0.18094    0.03935    4.60  4.3e-06 ***
## Year2012           0.19632    0.03767    5.21  1.9e-07 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.677
## Multiple R-squared:  0.0126, Adjusted R-squared:  0.0109
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
##  897 weights are ~= 1. The remaining 9356 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0381  0.8700  0.9490  0.9120  0.9840  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          9.75e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Sample size for the above analysis:  10253"
## [1] ""
## [1] ""
## [1] "####################################"
## [1] "Analysis of AJSC 2206"
## [1] "####################################"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  343  260  354  326  334  406  414  421  490  492  528  551  462  296  255
## 2011 2012
##  295  286
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##   85   49  122  128  116  108  158  139  200  206  234  262  222  133  107
```

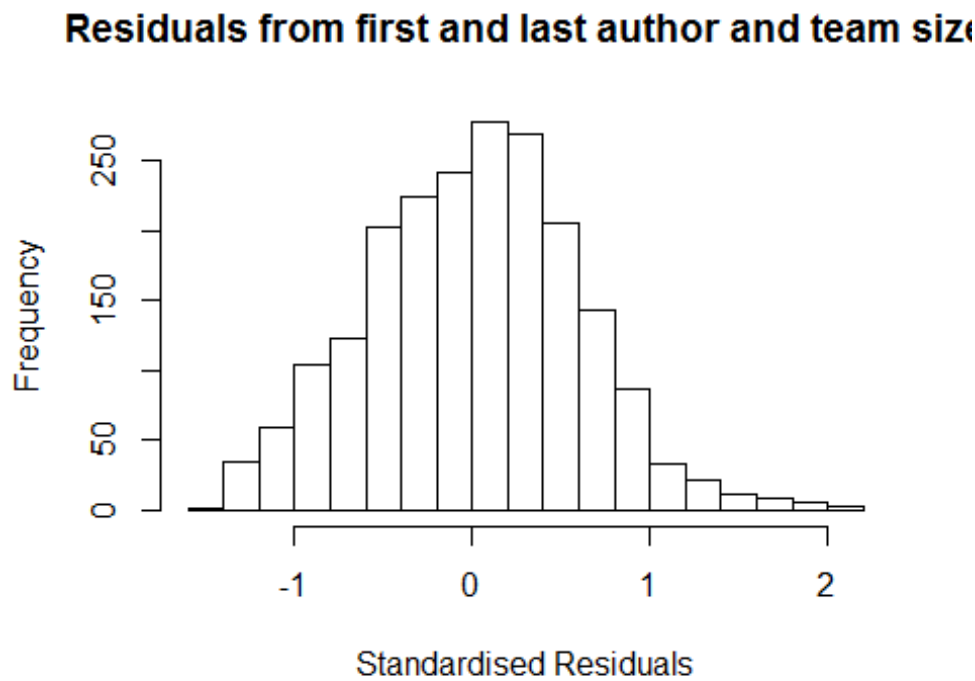
```
## 2011 2012
##  132  133
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##   76   43  106   98  100   90  131  119  163  163  179  216  179  106   84
## 2011 2012
##  100  103
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
##  Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 110, df = 16, p-value <2e-16
```




```
##
##  Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 0.0084, df = 1, p-value = 0.9
```

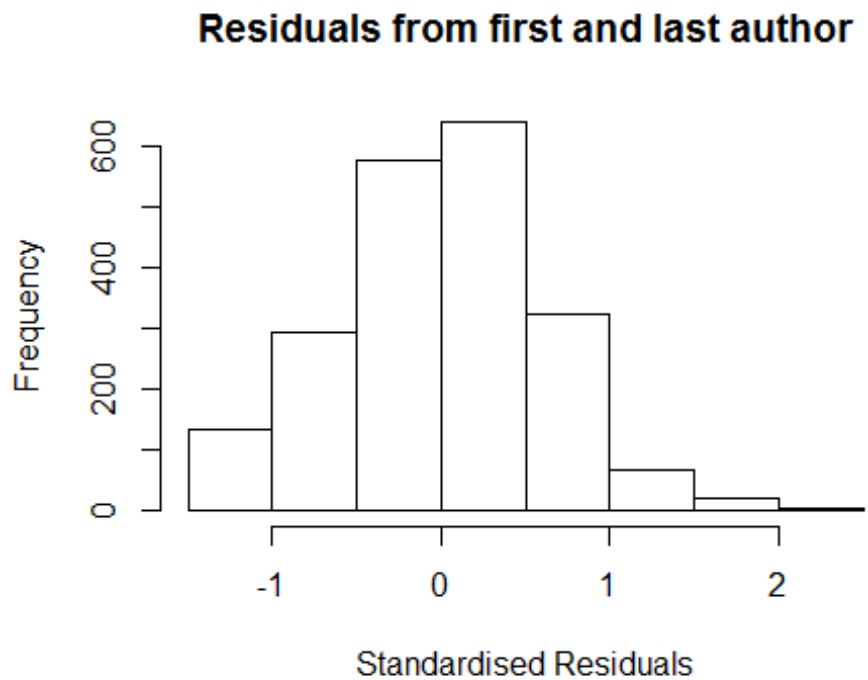



```
## [1] "Female first author team size 2018 geometric mean: 2.99379516552391"
## [1] "Male first author team size 2018 geometric mean: 2.49237191543506"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 880, p-value = 0.2
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 2.85002318705687"
## [1] "Male last author team size 2018 geometric mean: 2.49907693025835"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 980, p-value = 0.3
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##                     GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.122  1           1.059
## LastAuthorFemale  1.153  1           1.074
## UniqueAuthors     1.241  4           1.027
## Year              1.295 16           1.008
```



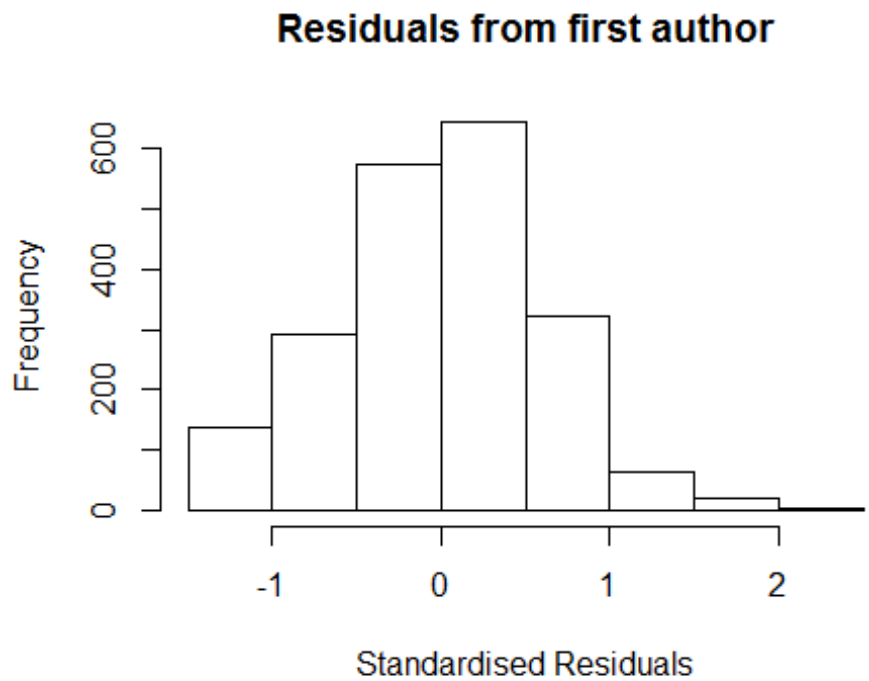

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
UniqueAuthors +
##     Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##     k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.4468 -0.4142  0.0335  0.4028  2.1717
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)          0.7365     0.1314    5.61  2.3e-08 ***
## FirstAuthorFemale1  -0.0473     0.0543   -0.87    0.384
## LastAuthorFemale1   -0.0827     0.0559   -1.48    0.139
## UniqueAuthors2       0.3340     0.0358    9.34  < 2e-16 ***
## UniqueAuthors3       0.3918     0.0407    9.62  < 2e-16 ***
## UniqueAuthors4       0.4011     0.0574    6.99  3.6e-12 ***
## UniqueAuthors5       0.4770     0.0664    7.19  9.3e-13 ***
## Year1997             0.0526     0.1671    0.31    0.753
## Year1998             0.3093     0.1521    2.03    0.042 *
## Year1999             0.1670     0.1511    1.10    0.269
```

```
## Year2000             0.1262     0.1520    0.83    0.406
## Year2001             0.2902     0.1544    1.88    0.060 .
## Year2002             0.1803     0.1454    1.24    0.215
## Year2003             0.2342     0.1433    1.63    0.102
## Year2004             0.2403     0.1389    1.73    0.084 .
## Year2005             0.1186     0.1388    0.85    0.393
## Year2006             0.1629     0.1374    1.18    0.236
## Year2007             0.0870     0.1383    0.63    0.529
## Year2008             0.1759     0.1388    1.27    0.205
## Year2009             0.1512     0.1463    1.03    0.302
## Year2010             0.1078     0.1498    0.72    0.472
## Year2011             0.0917     0.1463    0.63    0.531
## Year2012             0.0650     0.1505    0.43    0.666
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.598
## Multiple R-squared:  0.0846, Adjusted R-squared:  0.0747
## Convergence in 19 IRWLS iterations
##
## Robustness weights:
##  153 weights are ~= 1. The remaining 1903 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.159   0.871   0.952   0.909   0.985   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          4.86e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.098  1           1.048
## LastAuthorFemale  1.130  1           1.063
## Year              1.061 16           1.002
```



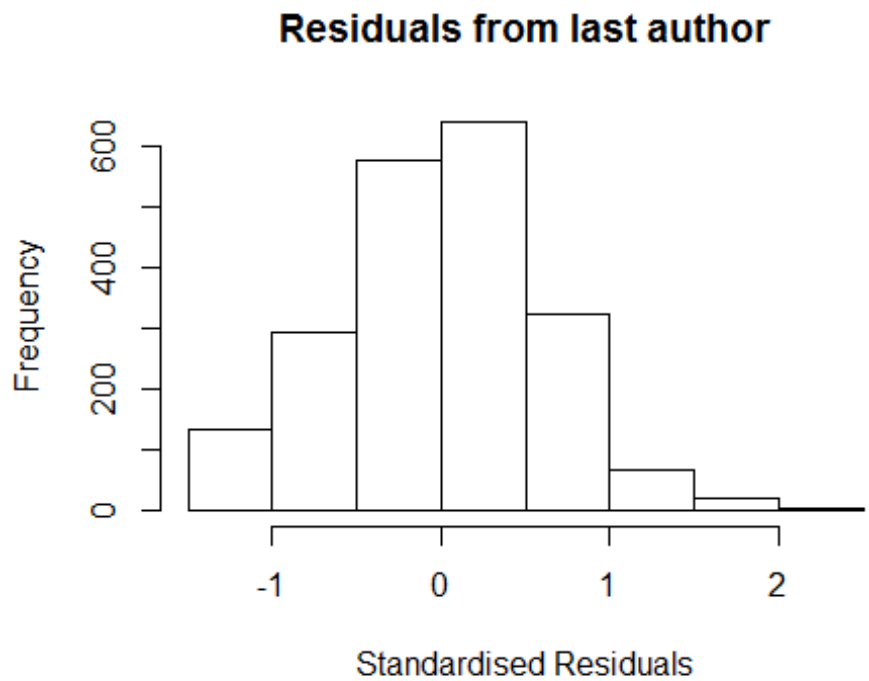

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.3095 -0.4127  0.0183  0.4125  2.4124
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)           0.8876     0.1416    6.27  4.4e-10 ***
## FirstAuthorFemale1   -0.0283     0.0556   -0.51   0.6112
## LastAuthorFemale1    -0.0850     0.0572   -1.49   0.1374
## Year1997              0.1288     0.1776    0.73   0.4683
## Year1998              0.4219     0.1597    2.64   0.0083 **
## Year1999              0.2309     0.1602    1.44   0.1496
## Year2000              0.2866     0.1615    1.77   0.0761 .
## Year2001              0.4097     0.1642    2.50   0.0127 *
## Year2002              0.2617     0.1547    1.69   0.0909 .
## Year2003              0.3025     0.1530    1.98   0.0482 *
## Year2004              0.3562     0.1478    2.41   0.0160 *
## Year2005              0.2359     0.1477    1.60   0.1104
```

```
## Year2006             0.3042     0.1462    2.08   0.0376 *
## Year2007             0.1981     0.1476    1.34   0.1798
## Year2008             0.3083     0.1479    2.08   0.0372 *
## Year2009             0.3273     0.1536    2.13   0.0332 *
## Year2010             0.2704     0.1583    1.71   0.0876 .
## Year2011             0.2415     0.1549    1.56   0.1192
## Year2012             0.2308     0.1583    1.46   0.1451
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.616
## Multiple R-squared:  0.0184, Adjusted R-squared:  0.00972
## Convergence in 19 IRWLS iterations
##
## Robustness weights:
##  164 weights are ~= 1. The remaining 1892 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.091   0.864   0.952   0.908   0.984   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          4.86e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.013  1           1.007
## Year              1.013 16           1.000
```




```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.3058 -0.4121  0.0213  0.4110  2.4129
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)           0.8871     0.1419    6.25  4.9e-10 ***
## FirstAuthorFemale1   -0.0508     0.0546   -0.93   0.3525
## Year1997              0.1293     0.1778    0.73   0.4673
## Year1998              0.4187     0.1599    2.62   0.0089 **
## Year1999              0.2279     0.1606    1.42   0.1561
## Year2000              0.2865     0.1616    1.77   0.0764 .
## Year2001              0.4063     0.1641    2.48   0.0134 *
## Year2002              0.2562     0.1551    1.65   0.0987 .
## Year2003              0.2981     0.1532    1.95   0.0518 .
## Year2004              0.3523     0.1482    2.38   0.0175 *
## Year2005              0.2340     0.1481    1.58   0.1142
## Year2006              0.3011     0.1466    2.05   0.0401 *
```

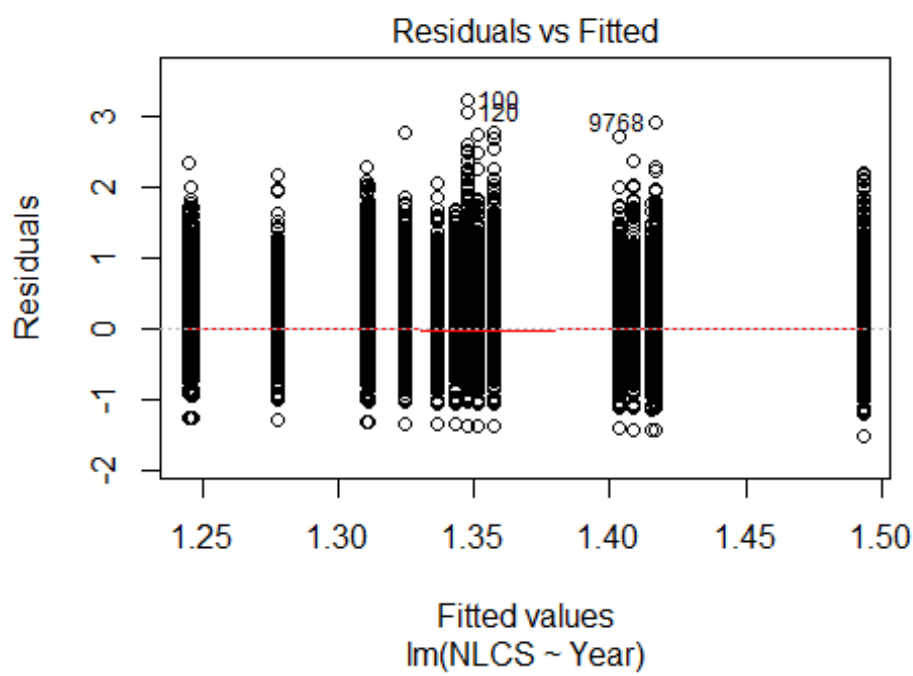
```
## Year2007             0.1950     0.1479    1.32   0.1876
## Year2008             0.3060     0.1483    2.06   0.0392 *
## Year2009             0.3225     0.1539    2.10   0.0362 *
## Year2010             0.2550     0.1578    1.62   0.1063
## Year2011             0.2397     0.1553    1.54   0.1227
## Year2012             0.2243     0.1583    1.42   0.1567
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.616
## Multiple R-squared:  0.0174, Adjusted R-squared:  0.00917
## Convergence in 19 IRWLS iterations
##
## Robustness weights:
##  161 weights are ~= 1. The remaining 1895 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0909  0.8640  0.9510  0.9080  0.9840  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          4.86e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##                   GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.041  1           1.020
## Year             1.041 16           1.001
```



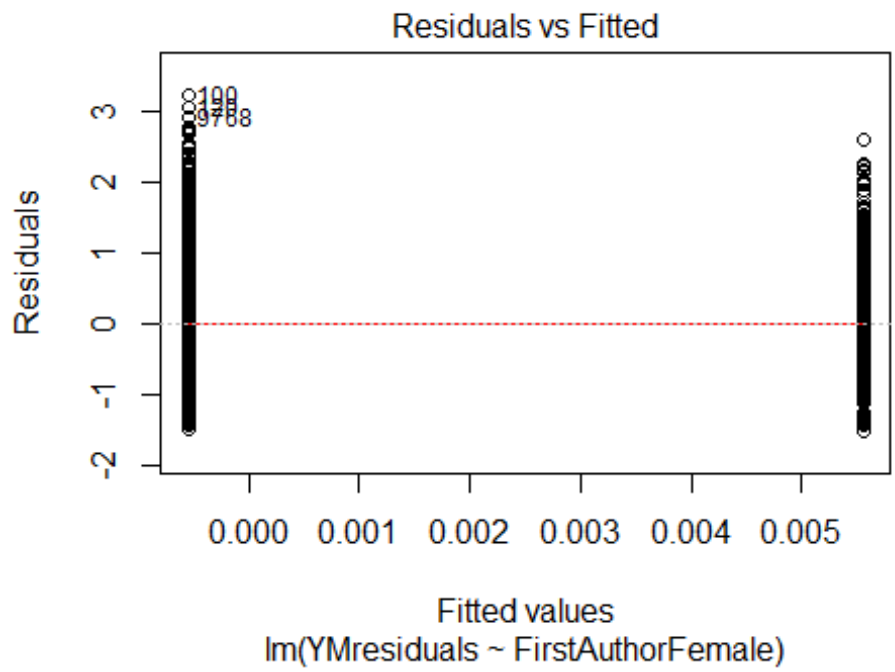

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.3086 -0.4123  0.0197  0.4149  2.4136
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)          0.8864     0.1418    6.25  4.9e-10 ***
## LastAuthorFemale1   -0.0933     0.0553   -1.69   0.0919 .
## Year1997             0.1300     0.1777    0.73   0.4646
## Year1998             0.4222     0.1599    2.64   0.0083 **
## Year1999             0.2305     0.1604    1.44   0.1508
## Year2000             0.2859     0.1615    1.77   0.0769 .
## Year2001             0.4089     0.1643    2.49   0.0129 *
## Year2002             0.2611     0.1548    1.69   0.0919 .
## Year2003             0.3021     0.1531    1.97   0.0486 *
## Year2004             0.3559     0.1479    2.41   0.0162 *
## Year2005             0.2355     0.1479    1.59   0.1114
## Year2006             0.3038     0.1464    2.07   0.0381 *
```

```
## Year2007            0.1975     0.1477    1.34   0.1813
## Year2008            0.3079     0.1480    2.08   0.0376 *
## Year2009            0.3271     0.1537    2.13   0.0334 *
## Year2010            0.2712     0.1585    1.71   0.0872 .
## Year2011            0.2398     0.1550    1.55   0.1219
## Year2012            0.2316     0.1585    1.46   0.1441
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.616
## Multiple R-squared:  0.0182, Adjusted R-squared:  0.01
## Convergence in 19 IRWLS iterations
##
## Robustness weights:
##  161 weights are ~= 1. The remaining 1895 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0901  0.8640  0.9520  0.9080  0.9840  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          4.86e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Sample size for the above analysis:  2056"
## [1] ""
## [1] ""
## [1] "####################################"
## [1] "Analysis of AJSC 2207"
## [1] "####################################"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 1393 1137 1321 1273 1526 1572 1390  958 1030 1008 1244 1316 1341 1444 1252
## 2011 2012
## 1263 1200
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  745  508  734  711  877  762  773  525  549  554  684  687  720  823  690
## 2011 2012
```

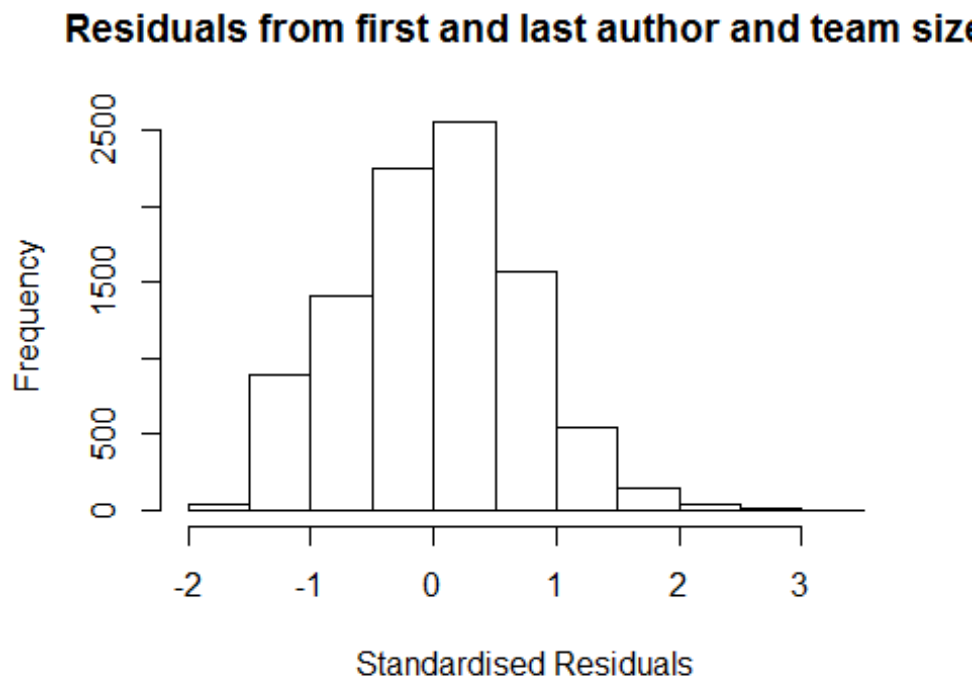
```
##  709  659
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  596  430  611  563  723  615  649  425  450  446  550  559  581  672  546
## 2011 2012
##  545  502
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
##  Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 270, df = 16, p-value <2e-16
```




```
##
##  Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 0.016, df = 1, p-value = 0.9
```



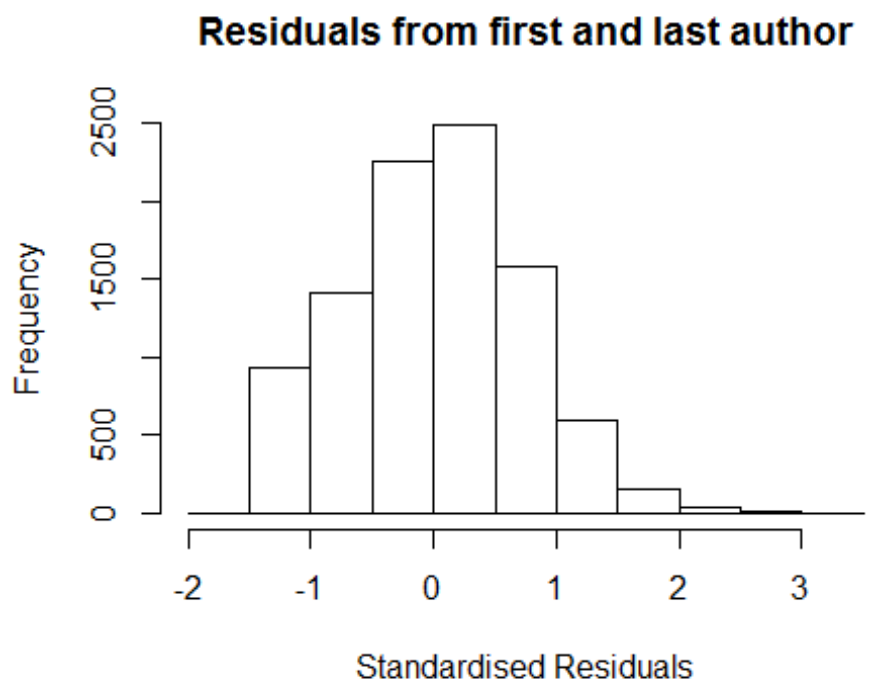

```
## [1] "Female first author team size 2018 geometric mean: 2.62245088585112"
## [1] "Male first author team size 2018 geometric mean: 2.55524948267149"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 3000, p-value = 0.7
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 2.47729778001086"
## [1] "Male last author team size 2018 geometric mean: 2.57306485475464"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 3000, p-value = 0.8
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.064  1           1.031
## LastAuthorFemale  1.061  1           1.030
## UniqueAuthors     1.062  4           1.007
## Year              1.071 16           1.002
```




```
## [1] "List of  14 outliers with residuals above  2.5"
##          ScopusId  NLCS Year OneField Fields residuals
## 100    0030211964 4.574 1996     1702      2     3.461
## 120    0030261036 4.407 1996     2207      2     2.916
## 243    0030142764 3.833 1996     1702      4     2.720
## 1681   0030643777 3.593 1997     2207      3     2.534
## 2827   0032049914 3.632 1998     2207      2     2.582
## 2908   0032308533 4.056 1998     2207      2     2.651
## 2976   0032047505 3.905 1998     1706      3     2.799
## 3085   0032204063 4.140 1998     1702      4     2.656
## 5601   4243148480 4.090 1999     1702      5     2.973
## 9601   0036811662 3.649 2002     1702      4     2.565
## 10705  0036608987 3.709 2002     2207      3     2.568
## 11511  0141607824 4.104 2003     1702      4     2.662
## 12457  4644244041 4.109 2004     1706      3     2.702
## 15612 33644997478 3.797 2006     1706      3     2.650
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
UniqueAuthors +
##     Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##     k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.6050 -0.4951  0.0217  0.4882  3.4609
```

```
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.11306    0.04037   27.57   <2e-16 ***
## FirstAuthorFemale1  0.01499    0.02818    0.53   0.5948
## LastAuthorFemale1  -0.07165    0.03148   -2.28   0.0229 *
## UniqueAuthors2      0.29894    0.02030   14.72   <2e-16 ***
## UniqueAuthors3      0.37772    0.02321   16.27   <2e-16 ***
## UniqueAuthors4      0.38594    0.02920   13.22   <2e-16 ***
## UniqueAuthors5      0.39309    0.03142   12.51   <2e-16 ***
## Year1997           -0.05455    0.06142   -0.89   0.3745
## Year1998           -0.00658    0.05335   -0.12   0.9018
## Year1999            0.00383    0.05193    0.07   0.9413
## Year2000           -0.02983    0.04895   -0.61   0.5423
## Year2001            0.09884    0.05068    1.95   0.0512 .
## Year2002            0.02774    0.05034    0.55   0.5817
## Year2003           -0.04854    0.05184   -0.94   0.3491
## Year2004           -0.00453    0.05104   -0.09   0.9292
## Year2005            0.04934    0.05062    0.97   0.3297
## Year2006            0.03416    0.04901    0.70   0.4858
## Year2007           -0.02204    0.04785   -0.46   0.6451
## Year2008           -0.06270    0.04809   -1.30   0.1924
## Year2009           -0.12985    0.04693   -2.77   0.0057 **
## Year2010           -0.14206    0.04749   -2.99   0.0028 **
## Year2011           -0.12251    0.04828   -2.54   0.0112 *
## Year2012           -0.15839    0.04888   -3.24   0.0012 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.719
## Multiple R-squared:  0.0515, Adjusted R-squared:  0.0493
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
##  observation 50 is an outlier with |weight| = 0 ( < 1.1e-05);
##  846 weights are ~= 1. The remaining 8616 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.049   0.866   0.949   0.907   0.985   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.06e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                    psi           subsampling                   cov
```

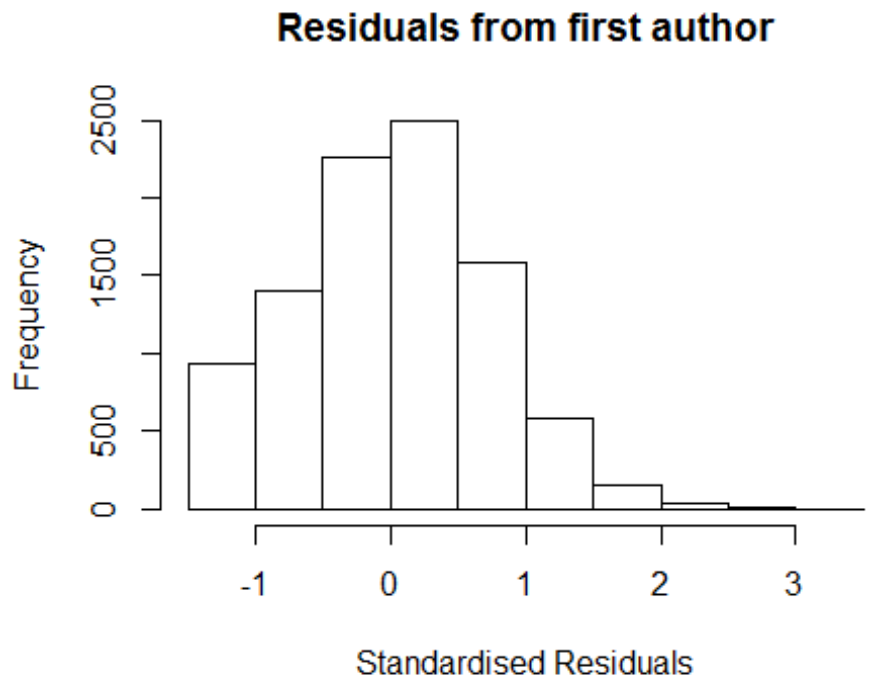
```
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.050  1           1.025
## LastAuthorFemale  1.053  1           1.026
## Year              1.018 16           1.001
```




```
## [1] "List of  10 outliers with residuals above  2.5"
##         ScopusId  NLCS Year OneField Fields residuals
## 13    0029679044 3.953 1996     1702      2     2.609
## 100   0030211964 4.574 1996     1702      2     3.270
## 120   0030261036 4.407 1996     2207      2     3.103
## 243   0030142764 3.833 1996     1702      4     2.529
## 2908  0032308533 4.056 1998     2207      2     2.736
## 2976  0032047505 3.905 1998     1706      3     2.585
## 3085  0032204063 4.140 1998     1702      4     2.820
## 5601  4243148480 4.090 1999     1702      5     2.754
## 11511 0141607824 4.104 2003     1702      4     2.814
## 12457 4644244041 4.109 2004     1706      3     2.731
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
```

```
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##    Min     1Q Median     3Q    Max
## -1.515 -0.497  0.029  0.501  3.270
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.30448    0.03970   32.86   <2e-16 ***
## FirstAuthorFemale1  0.03959    0.02831    1.40   0.1620
## LastAuthorFemale1  -0.06361    0.03163   -2.01   0.0443 *
## Year1997           -0.04843    0.06290   -0.77   0.4413
## Year1998            0.01542    0.05500    0.28   0.7792
## Year1999            0.03195    0.05276    0.61   0.5449
## Year2000            0.01284    0.04905    0.26   0.7936
## Year2001            0.17114    0.05155    3.32   0.0009 ***
## Year2002            0.09350    0.05104    1.83   0.0670 .
## Year2003           -0.01492    0.05348   -0.28   0.7802
## Year2004            0.07315    0.05139    1.42   0.1546
## Year2005            0.10465    0.05194    2.01   0.0440 *
## Year2006            0.10116    0.04966    2.04   0.0417 *
## Year2007            0.04402    0.04898    0.90   0.3689
## Year2008            0.00991    0.04901    0.20   0.8397
## Year2009           -0.04139    0.04720   -0.88   0.3805
## Year2010           -0.06694    0.04846   -1.38   0.1672
## Year2011           -0.04582    0.04879   -0.94   0.3477
## Year2012           -0.08046    0.04962   -1.62   0.1049
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.732
## Multiple R-squared:  0.00883,    Adjusted R-squared:  0.00694
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
##  791 weights are ~= 1. The remaining 8672 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0082  0.8670  0.9500  0.9070  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.06e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                    psi           subsampling                   cov
```

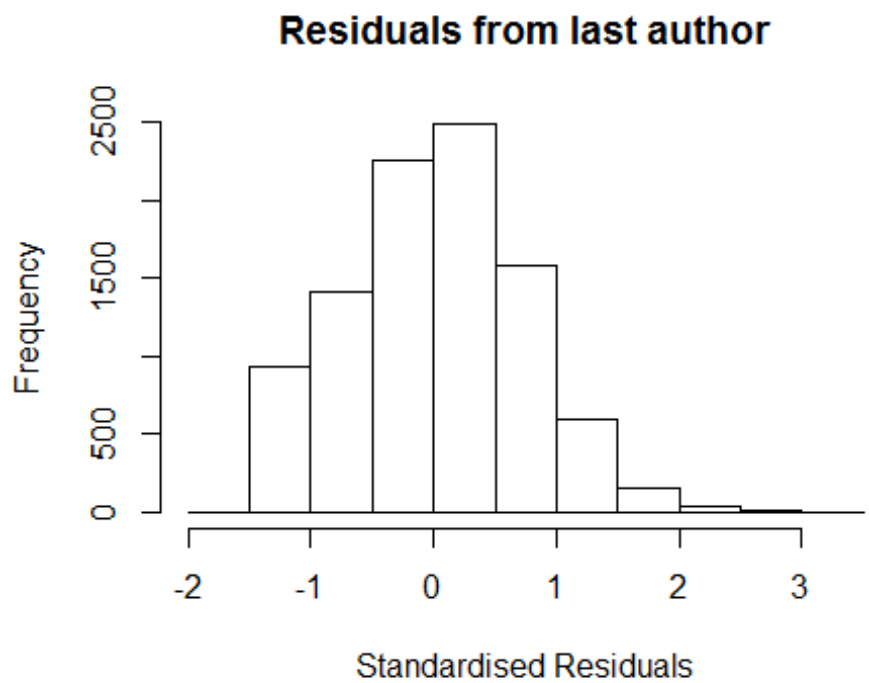
```
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.007  1           1.004
## Year              1.007 16           1.000
```




```
## [1] "List of  10 outliers with residuals above  2.5"
##         ScopusId  NLCS Year OneField Fields residuals
## 13    0029679044 3.953 1996     1702      2     2.609
## 100   0030211964 4.574 1996     1702      2     3.270
## 120   0030261036 4.407 1996     2207      2     3.103
## 243   0030142764 3.833 1996     1702      4     2.529
## 2908  0032308533 4.056 1998     2207      2     2.736
## 2976  0032047505 3.905 1998     1706      3     2.585
## 3085  0032204063 4.140 1998     1702      4     2.820
## 5601  4243148480 4.090 1999     1702      5     2.754
## 11511 0141607824 4.104 2003     1702      4     2.814
## 12457 4644244041 4.109 2004     1706      3     2.731
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
```

```
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.4970 -0.4942  0.0279  0.5006  3.2729
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)          1.3011     0.0397   32.80  < 2e-16 ***
## FirstAuthorFemale1   0.0241     0.0279    0.86  0.38807
## Year1997            -0.0477     0.0630   -0.76  0.44917
## Year1998             0.0159     0.0550    0.29  0.77245
## Year1999             0.0326     0.0528    0.62  0.53731
## Year2000             0.0125     0.0491    0.26  0.79852
## Year2001             0.1718     0.0516    3.33  0.00087 ***
## Year2002             0.0933     0.0510    1.83  0.06749 .
## Year2003            -0.0154     0.0535   -0.29  0.77346
## Year2004             0.0739     0.0514    1.44  0.15039
## Year2005             0.1030     0.0520    1.98  0.04771 *
## Year2006             0.1019     0.0497    2.05  0.04022 *
## Year2007             0.0437     0.0491    0.89  0.37305
## Year2008             0.0100     0.0491    0.20  0.83850
## Year2009            -0.0401     0.0472   -0.85  0.39598
## Year2010            -0.0655     0.0485   -1.35  0.17683
## Year2011            -0.0468     0.0489   -0.96  0.33828
## Year2012            -0.0813     0.0496   -1.64  0.10147
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.732
## Multiple R-squared:  0.00838,    Adjusted R-squared:  0.00659
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
##  781 weights are ~= 1. The remaining 8682 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0079  0.8670  0.9500  0.9080  0.9860  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.06e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
```

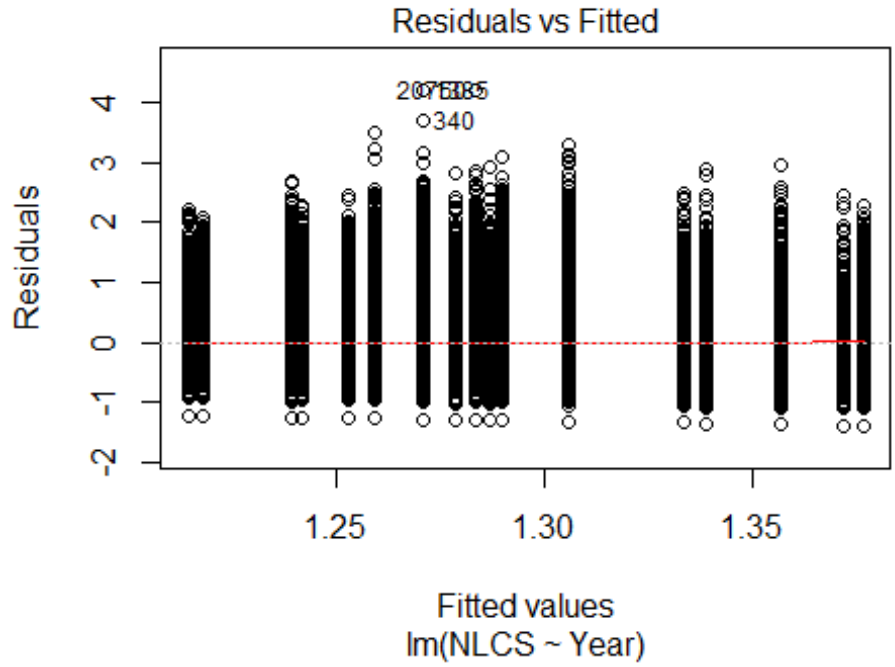
```
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##                  GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.01  1           1.005
## Year             1.01 16           1.000
```



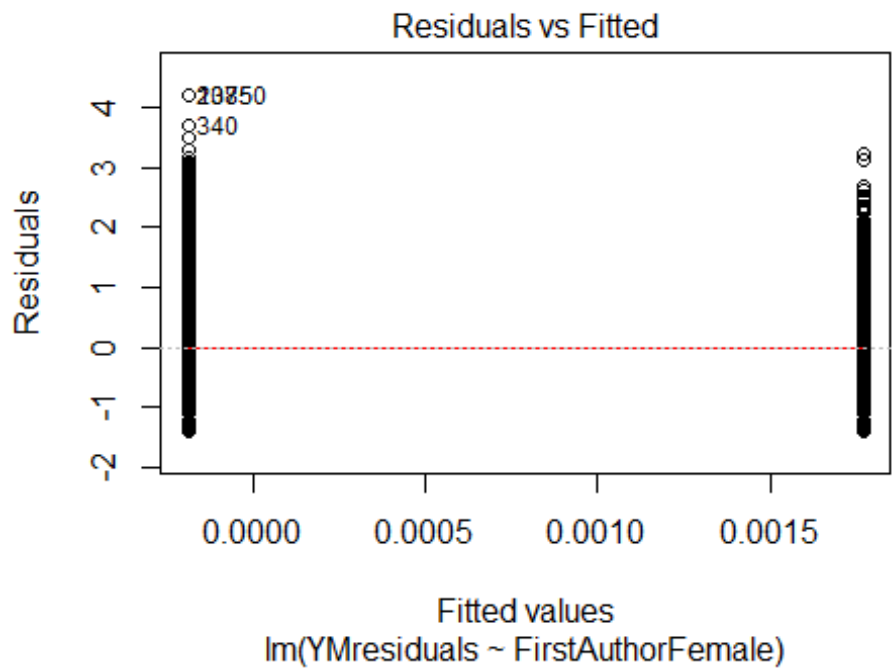

```
## [1] "List of  10 outliers with residuals above  2.5"
##         ScopusId  NLCS Year OneField Fields residuals
## 13    0029679044 3.953 1996     1702      2     2.609
## 100   0030211964 4.574 1996     1702      2     3.270
## 120   0030261036 4.407 1996     2207      2     3.103
## 243   0030142764 3.833 1996     1702      4     2.529
## 2908  0032308533 4.056 1998     2207      2     2.736
## 2976  0032047505 3.905 1998     1706      3     2.585
## 3085  0032204063 4.140 1998     1702      4     2.820
## 5601  4243148480 4.090 1999     1702      5     2.754
## 11511 0141607824 4.104 2003     1702      4     2.814
## 12457 4644244041 4.109 2004     1706      3     2.731
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.4779 -0.4954  0.0269  0.5000  3.2675
```

```
##
## Coefficients:
##                   Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.3065     0.0397   32.88  < 2e-16 ***
## LastAuthorFemale1  -0.0522     0.0309   -1.69  0.09159 .
## Year1997           -0.0492     0.0629   -0.78  0.43448
## Year1998            0.0160     0.0550    0.29  0.77063
## Year1999            0.0320     0.0528    0.61  0.54452
## Year2000            0.0135     0.0491    0.27  0.78367
## Year2001            0.1714     0.0516    3.32  0.00089 ***
## Year2002            0.0945     0.0510    1.85  0.06398 .
## Year2003           -0.0142     0.0535   -0.26  0.79108
## Year2004            0.0749     0.0513    1.46  0.14471
## Year2005            0.1052     0.0520    2.02  0.04291 *
## Year2006            0.1016     0.0497    2.05  0.04072 *
## Year2007            0.0444     0.0490    0.91  0.36489
## Year2008            0.0108     0.0490    0.22  0.82599
## Year2009           -0.0405     0.0472   -0.86  0.39095
## Year2010           -0.0660     0.0484   -1.36  0.17295
## Year2011           -0.0454     0.0488   -0.93  0.35263
## Year2012           -0.0798     0.0496   -1.61  0.10811
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.732
## Multiple R-squared:  0.00865,    Adjusted R-squared:  0.00686
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
##  777 weights are ~= 1. The remaining 8686 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0084  0.8670  0.9500  0.9080  0.9860  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.06e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Sample size for the above analysis:  9463"
## [1] ""
```

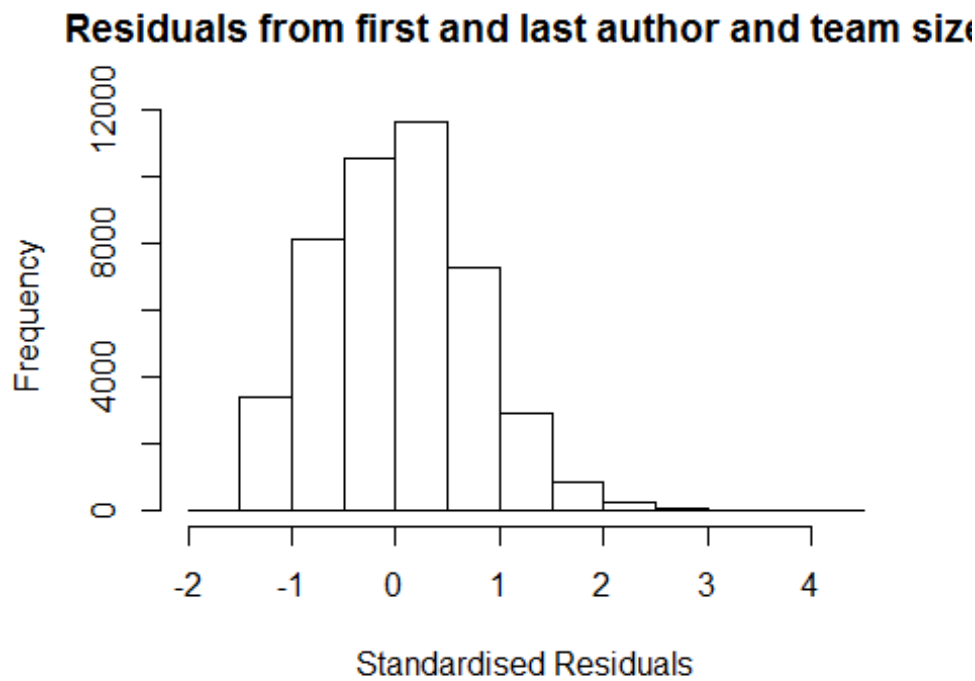
```
## [1] ""
## [1] "####################################"
## [1] "Analysis of AJSC 2208"
## [1] "####################################"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 7401 7647 7060  858 7933 9394 8834 5199 5408 6158 7545 7250 6825 6502 6187
## 2011 2012
## 6437 6385
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 3612 3609 3445  358 3971 4502 4728 2679 2735 3167 3695 3588 3393 3382 3116
## 2011 2012
## 3269 3268
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 2974 2977 2857  296 3232 3558 3761 2131 2171 2492 2879 2739 2656 2644 2479
## 2011 2012
## 2532 2581
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
##  Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 1600, df = 16, p-value <2e-16
```




```
##
##  Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 14, df = 1, p-value = 2e-04
```




```
## [1] "Female first author team size 2018 geometric mean: 3.38355713918163"
## [1] "Male first author team size 2018 geometric mean: 3.05642902070959"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 50000, p-value = 0.1
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 3.66967147878159"
## [1] "Male last author team size 2018 geometric mean: 3.02855331305212"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 50000, p-value = 6e-04
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.040  1           1.020
## LastAuthorFemale  1.034  1           1.017
## UniqueAuthors     1.044  4           1.005
## Year              1.045 16           1.001
```

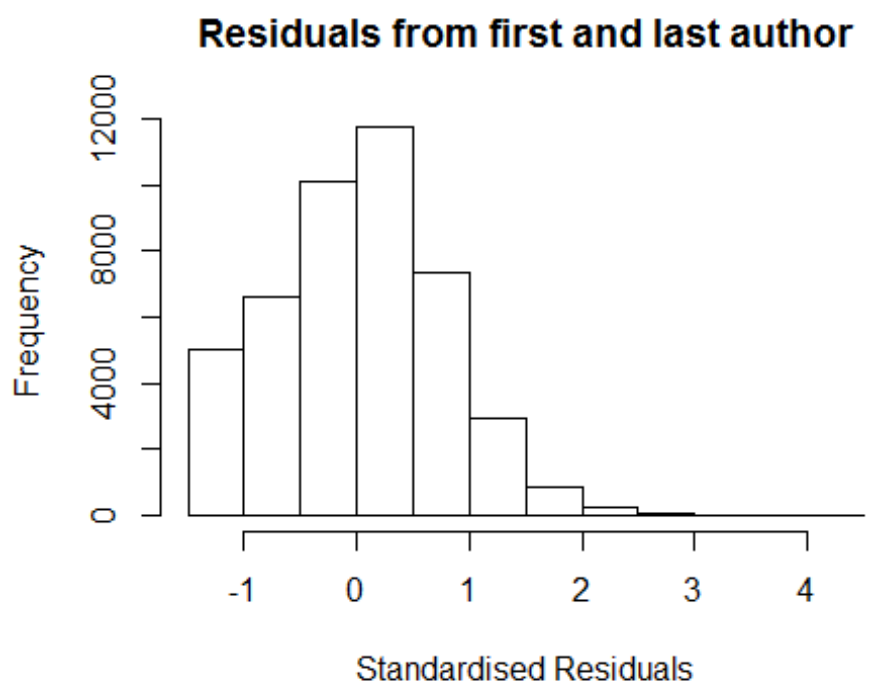



```
## [1] "List of  67 outliers with residuals above  2.5"
##           ScopusId  NLCS Year OneField Fields residuals
## 224     0030151353 3.787 1996     2208      1     2.813
## 340     0030234863 4.961 1996     2208      1     3.987
## 737     0029777083 3.495 1996     1710      2     2.521
## 1385    0030173121 5.461 1996     2208      2     4.125
## 2021    0030261036 4.407 1996     2207      2     3.022
## 2228    0030384174 3.534 1996     2208      2     2.560
## 2620    0030107810 3.576 1996     1706      3     2.602
## 4202    0030262865 4.271 1996     1705      2     3.297
## 4576    0030232761 3.847 1996     1502      6     2.511
## 5055    0030170353 3.905 1996     1705      2     2.569
## 5059    0030171479 3.606 1996     1705      2     2.632
## 5628    0030126485 3.949 1996     2208      3     3.010
## 5756    0030105412 3.759 1996     2208      1     2.785
## 6298    0030290680 3.728 1996     2208      1     2.754
## 7404    0030287048 3.862 1996     1711      3     2.888
## 8225    0030735427 3.883 1997     2208      1     2.526
## 8231    0030735959 4.307 1997     2208      1     2.950
## 8284    0031070220 4.375 1997     2208      1     3.381
## 8300    0031075691 3.577 1997     2208      1     2.583
## 8558    0031233424 4.026 1997     2208      1     3.032
## 8575    0031236835 4.282 1997     2208      1     3.288
## 8595    0031257246 4.420 1997     2208      1     3.460
## 8679    0031338376 4.599 1997     2208      1     3.242
## 9671    0031165386 3.657 1997     1711      2     2.663
## 10250   0031249402 3.505 1997     2208      2     2.511
```

```
## 10793   0030643777 3.593 1997     2207      3     2.599
## 12053   0031383380 3.875 1997     1705      4     2.518
## 13488   0031166167 3.576 1997     2208      1     2.582
## 13718   0031147079 4.121 1997     2208      1     2.764
## 14709   0031078854 4.109 1997     1711      2     2.705
## 17008   0032138896 4.051 1998     2208      1     2.655
## 17089   0032202775 3.489 1998     2208      1     2.503
## 17249  33747433684 3.922 1998     2208      1     2.574
## 17721   0032049914 3.632 1998     2207      2     2.681
## 18241   0032136715 3.940 1998     2208      2     2.592
## 18731   0032308533 4.056 1998     2207      2     2.708
## 19247   0032047505 3.905 1998     1706      3     2.919
## 19775   0032204063 4.140 1998     1702      4     2.744
## 20750   0032183752 5.467 1998     1705      2     4.481
## 22953   0031999108 3.530 1998     1705      2     2.544
## 26013   0000894702 3.678 2000     2208      1     2.746
## 26076   0034156558 3.887 2000     2208      1     2.955
## 31752   0033884858 3.916 2000     1711      2     2.573
## 34945   0035119053 4.046 2001     2208      1     3.131
## 35072  33646900503 3.909 2001     2208      1     2.567
## 36386   0035481826 4.359 2001     2208      1     3.444
## 36725   0035493267 3.845 2001     2208      3     2.503
## 38425   0035396555 3.489 2001     2208      2     2.574
## 38991   0035364996 3.796 2001     2208      2     2.881
## 42632  33845588040 3.436 2001     2202      2     2.521
## 46143  33645307063 3.797 2002     2208      1     2.502
## 46610  33646589837 4.485 2002     1706      3     3.193
## 50846   0036544252 3.391 2002     2202      2     2.506
## 51783   0036476080 3.674 2002     2208      2     2.789
## 54477   0036608987 3.709 2002     2207      3     2.824
## 54533   0036688074 4.740 2002     1705      3     3.445
## 57609   1642401353 3.922 2003     2202      2     2.894
## 63628  10944266504 4.134 2004     1705      3     2.801
## 64295   4644244041 4.109 2004     1706      3     2.745
## 67847   0742290133 4.216 2004     2208      2     2.804
## 69908  33846118079 3.662 2005     1708      3     2.621
## 82611  33644997478 3.797 2006     1706      3     2.793
## 86201  34347379461 3.514 2007     2202      2     2.566
## 92163  64149119332 4.203 2007     2208      1     2.844
## 92889  85032751965 3.652 2007     1711      3     2.704
## 101063 85032750937 4.080 2008     1711      3     2.792
## 104546 67449128189 3.630 2009     2208      1     2.718
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
UniqueAuthors +
##     Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##     k.max = 1000))
##  \--> method = "MM"
```

```
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.5534 -0.5159  0.0184  0.5011  4.4814
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         0.97385    0.01817   53.61  < 2e-16 ***
## FirstAuthorFemale1 -0.00412    0.01199   -0.34  0.73118
## LastAuthorFemale1  -0.03059    0.01304   -2.35  0.01902 *
## UniqueAuthors2      0.36245    0.01086   33.36  < 2e-16 ***
## UniqueAuthors3      0.41078    0.01170   35.11  < 2e-16 ***
## UniqueAuthors4      0.40932    0.01310   31.24  < 2e-16 ***
## UniqueAuthors5      0.42666    0.01213   35.19  < 2e-16 ***
## Year1997            0.02046    0.02391    0.86  0.39210
## Year1998            0.01177    0.02370    0.50  0.61965
## Year1999            0.15289    0.05093    3.00  0.00268 **
## Year2000           -0.04178    0.02281   -1.83  0.06701 .
## Year2001           -0.05880    0.02335   -2.52  0.01180 *
## Year2002           -0.08852    0.02211   -4.00  6.2e-05 ***
## Year2003            0.05365    0.02332    2.30  0.02140 *
## Year2004            0.02750    0.02288    1.20  0.22939
## Year2005            0.06735    0.02183    3.08  0.00204 **
## Year2006            0.02979    0.02112    1.41  0.15845
## Year2007           -0.02555    0.02092   -1.22  0.22189
## Year2008           -0.04826    0.02110   -2.29  0.02218 *
## Year2009           -0.06215    0.02123   -2.93  0.00342 **
## Year2010           -0.07426    0.02150   -3.45  0.00055 ***
## Year2011           -0.10607    0.02109   -5.03  4.9e-07 ***
## Year2012           -0.11257    0.02124   -5.30  1.2e-07 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.747
## Multiple R-squared:  0.051,  Adjusted R-squared:  0.0506
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
##  3 observations c(167,596,7473) are outliers with |weight| = 0 ( < 2.2e-
06);
##  3790 weights are ~= 1. The remaining 41166 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0005  0.8660  0.9500  0.9100  0.9860  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          2.22e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
```

```
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.026  1           1.013
## LastAuthorFemale  1.022  1           1.011
## Year              1.007 16           1.000
```

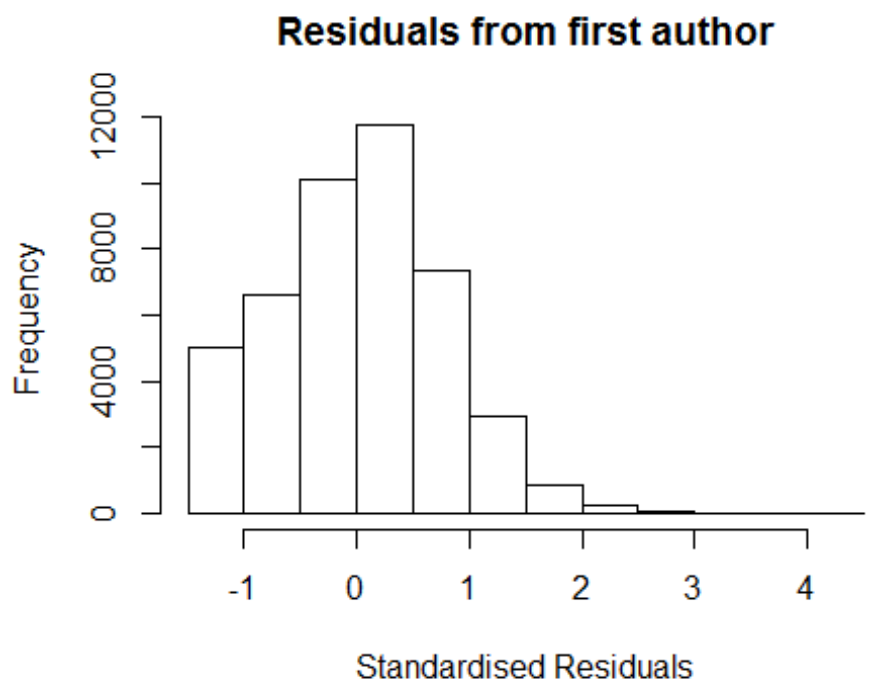



```
## [1] "List of  56 outliers with residuals above  2.5"
##           ScopusId  NLCS Year OneField Fields residuals
## 224     0030151353 3.787 1996     2208      1     2.565
## 340     0030234863 4.961 1996     2208      1     3.739
## 1385    0030173121 5.461 1996     2208      2     4.239
## 2021    0030261036 4.407 1996     2207      2     3.185
## 4202    0030262865 4.271 1996     1705      2     3.049
## 4576    0030232761 3.847 1996     1502      6     2.625
## 5035    0030166739 3.759 1996     1705      2     2.537
## 5055    0030170353 3.905 1996     1705      2     2.683
## 5628    0030126485 3.949 1996     2208      3     2.733
## 5756    0030105412 3.759 1996     2208      1     2.537
```

```
## 6298    0030290680 3.728 1996     2208      1     2.506
## 7404    0030287048 3.862 1996     1711      3     2.640
## 8225    0030735427 3.883 1997     2208      1     2.633
## 8231    0030735959 4.307 1997     2208      1     3.057
## 8271    0030846969 3.814 1997     2208      1     2.564
## 8284    0031070220 4.375 1997     2208      1     3.125
## 8558    0031233424 4.026 1997     2208      1     2.776
## 8575    0031236835 4.282 1997     2208      1     3.032
## 8595    0031257246 4.420 1997     2208      1     3.176
## 8679    0031338376 4.599 1997     2208      1     3.349
## 9312    0031102203 3.829 1997     1711      2     2.559
## 12053   0031383380 3.875 1997     1705      4     2.625
## 13718   0031147079 4.121 1997     2208      1     2.871
## 14655   0030643536 3.766 1997     2201      2     2.516
## 14709   0031078854 4.109 1997     1711      2     2.859
## 15694   0031246031 3.912 1997     1705      3     2.662
## 17008   0032138896 4.051 1998     2208      1     2.802
## 17249  33747433684 3.922 1998     2208      1     2.673
## 18241   0032136715 3.940 1998     2208      2     2.691
## 18731   0032308533 4.056 1998     2207      2     2.807
## 18930  33747921774 3.844 1998     1900      2     2.595
## 19247   0032047505 3.905 1998     1706      3     2.656
## 19775   0032204063 4.140 1998     1702      4     2.891
## 20750   0032183752 5.467 1998     1705      2     4.218
## 21975   0032074579 3.796 1998     1705      2     2.547
## 26076   0034156558 3.887 2000     2208      1     2.675
## 31752   0033884858 3.916 2000     1711      2     2.704
## 34945   0035119053 4.046 2001     2208      1     2.805
## 35072  33646900503 3.909 2001     2208      1     2.668
## 36386   0035481826 4.359 2001     2208      1     3.118
## 36725   0035493267 3.845 2001     2208      3     2.604
## 38991   0035364996 3.796 2001     2208      2     2.555
## 46143  33645307063 3.797 2002     2208      1     2.591
## 46610  33646589837 4.485 2002     1706      3     3.258
## 47973   0036747248 3.735 2002     1705      2     2.508
## 48115   0036738266 3.707 2002     1705      3     2.501
## 51612   0036474829 3.755 2002     2208      1     2.549
## 54477   0036608987 3.709 2002     2207      3     2.503
## 54533   0036688074 4.740 2002     1705      3     3.534
## 57609   1642401353 3.922 2003     2202      2     2.600
## 63628  10944266504 4.134 2004     1705      3     2.850
## 64295   4644244041 4.109 2004     1706      3     2.799
## 67847   0742290133 4.216 2004     2208      2     2.906
## 92163  64149119332 4.203 2007     2208      1     2.930
## 92165  64149124758 3.845 2007     2208      1     2.572
## 101063 85032750937 4.080 2008     1711      3     2.833
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
```

```
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##    Min     1Q Median     3Q    Max
## -1.397 -0.520  0.031  0.511  4.239
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.22154    0.01751   69.78  < 2e-16 ***
## FirstAuthorFemale1  0.02055    0.01217    1.69  0.09133 .
## LastAuthorFemale1  -0.02626    0.01333   -1.97  0.04874 *
## Year1997            0.02797    0.02458    1.14  0.25521
## Year1998            0.02724    0.02452    1.11  0.26660
## Year1999            0.15532    0.05070    3.06  0.00219 **
## Year2000           -0.00936    0.02322   -0.40  0.68698
## Year2001            0.01913    0.02369    0.81  0.41943
## Year2002           -0.01513    0.02243   -0.67  0.49996
## Year2003            0.10015    0.02430    4.12  3.8e-05 ***
## Year2004            0.08880    0.02356    3.77  0.00016 ***
## Year2005            0.13277    0.02251    5.90  3.7e-09 ***
## Year2006            0.10176    0.02170    4.69  2.8e-06 ***
## Year2007            0.05149    0.02162    2.38  0.01722 *
## Year2008            0.02514    0.02184    1.15  0.24974
## Year2009            0.01544    0.02198    0.70  0.48236
## Year2010            0.00681    0.02225    0.31  0.75950
## Year2011           -0.02805    0.02179   -1.29  0.19797
## Year2012           -0.02217    0.02184   -1.01  0.31019
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.763
## Multiple R-squared:  0.00391,    Adjusted R-squared:  0.00351
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
##  3 observations c(167,596,7473) are outliers with |weight| = 0 ( < 2.2e-
06);
##  3787 weights are ~= 1. The remaining 41169 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0005  0.8710  0.9500  0.9100  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          2.22e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
```

```
##           0        1000           0
##                   psi           subsampling                    cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.005  1           1.003
## Year              1.005 16           1.000
```

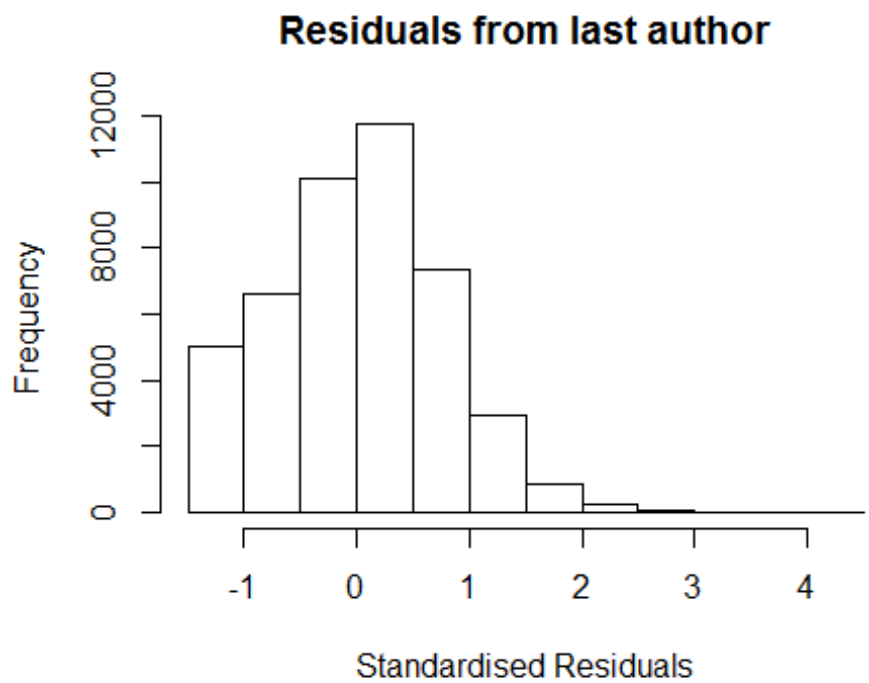



```
## [1] "List of  56 outliers with residuals above  2.5"
##           ScopusId  NLCS Year OneField Fields residuals
## 224     0030151353 3.787 1996     2208      1     2.565
## 340     0030234863 4.961 1996     2208      1     3.739
## 1385    0030173121 5.461 1996     2208      2     4.239
## 2021    0030261036 4.407 1996     2207      2     3.185
## 4202    0030262865 4.271 1996     1705      2     3.049
## 4576    0030232761 3.847 1996     1502      6     2.625
## 5035    0030166739 3.759 1996     1705      2     2.537
## 5055    0030170353 3.905 1996     1705      2     2.683
## 5628    0030126485 3.949 1996     2208      3     2.733
## 5756    0030105412 3.759 1996     2208      1     2.537
## 6298    0030290680 3.728 1996     2208      1     2.506
## 7404    0030287048 3.862 1996     1711      3     2.640
## 8225    0030735427 3.883 1997     2208      1     2.633
## 8231    0030735959 4.307 1997     2208      1     3.057
```

```
## 8271    0030846969 3.814 1997     2208      1     2.564
## 8284    0031070220 4.375 1997     2208      1     3.125
## 8558    0031233424 4.026 1997     2208      1     2.776
## 8575    0031236835 4.282 1997     2208      1     3.032
## 8595    0031257246 4.420 1997     2208      1     3.176
## 8679    0031338376 4.599 1997     2208      1     3.349
## 9312    0031102203 3.829 1997     1711      2     2.559
## 12053   0031383380 3.875 1997     1705      4     2.625
## 13718   0031147079 4.121 1997     2208      1     2.871
## 14655   0030643536 3.766 1997     2201      2     2.516
## 14709   0031078854 4.109 1997     1711      2     2.859
## 15694   0031246031 3.912 1997     1705      3     2.662
## 17008   0032138896 4.051 1998     2208      1     2.802
## 17249  33747433684 3.922 1998     2208      1     2.673
## 18241   0032136715 3.940 1998     2208      2     2.691
## 18731   0032308533 4.056 1998     2207      2     2.807
## 18930  33747921774 3.844 1998     1900      2     2.595
## 19247   0032047505 3.905 1998     1706      3     2.656
## 19775   0032204063 4.140 1998     1702      4     2.891
## 20750   0032183752 5.467 1998     1705      2     4.218
## 21975   0032074579 3.796 1998     1705      2     2.547
## 26076   0034156558 3.887 2000     2208      1     2.675
## 31752   0033884858 3.916 2000     1711      2     2.704
## 34945   0035119053 4.046 2001     2208      1     2.805
## 35072  33646900503 3.909 2001     2208      1     2.668
## 36386   0035481826 4.359 2001     2208      1     3.118
## 36725   0035493267 3.845 2001     2208      3     2.604
## 38991   0035364996 3.796 2001     2208      2     2.555
## 46143  33645307063 3.797 2002     2208      1     2.591
## 46610  33646589837 4.485 2002     1706      3     3.258
## 47973   0036747248 3.735 2002     1705      2     2.508
## 48115   0036738266 3.707 2002     1705      3     2.501
## 51612   0036474829 3.755 2002     2208      1     2.549
## 54477   0036608987 3.709 2002     2207      3     2.503
## 54533   0036688074 4.740 2002     1705      3     3.534
## 57609   1642401353 3.922 2003     2202      2     2.600
## 63628  10944266504 4.134 2004     1705      3     2.850
## 64295   4644244041 4.109 2004     1706      3     2.799
## 67847   0742290133 4.216 2004     2208      2     2.906
## 92163  64149119332 4.203 2007     2208      1     2.930
## 92165  64149124758 3.845 2007     2208      1     2.572
## 101063 85032750937 4.080 2008     1711      3     2.833
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
```

```
## -1.3905 -0.5198  0.0311  0.5104  4.2408
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.22021    0.01749   69.78  < 2e-16 ***
## FirstAuthorFemale1  0.01518    0.01212    1.25  0.21024
## Year1997            0.02814    0.02458    1.14  0.25229
## Year1998            0.02721    0.02452    1.11  0.26725
## Year1999            0.15513    0.05069    3.06  0.00221 **
## Year2000           -0.00949    0.02323   -0.41  0.68279
## Year2001            0.01869    0.02369    0.79  0.43014
## Year2002           -0.01555    0.02243   -0.69  0.48807
## Year2003            0.09983    0.02430    4.11  4.0e-05 ***
## Year2004            0.08865    0.02355    3.76  0.00017 ***
## Year2005            0.13240    0.02251    5.88  4.1e-09 ***
## Year2006            0.10139    0.02170    4.67  3.0e-06 ***
## Year2007            0.05134    0.02162    2.37  0.01757 *
## Year2008            0.02472    0.02184    1.13  0.25765
## Year2009            0.01519    0.02199    0.69  0.48969
## Year2010            0.00641    0.02225    0.29  0.77330
## Year2011           -0.02844    0.02179   -1.31  0.19188
## Year2012           -0.02287    0.02185   -1.05  0.29514
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.763
## Multiple R-squared:  0.00383,    Adjusted R-squared:  0.00346
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
##  3 observations c(167,596,7473) are outliers with |weight| = 0 ( < 2.2e-
06);
##  3798 weights are ~= 1. The remaining 41158 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0005  0.8710  0.9500  0.9100  0.9860  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          2.22e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
```

```
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.001  1           1.001
## Year             1.001 16           1.000
```

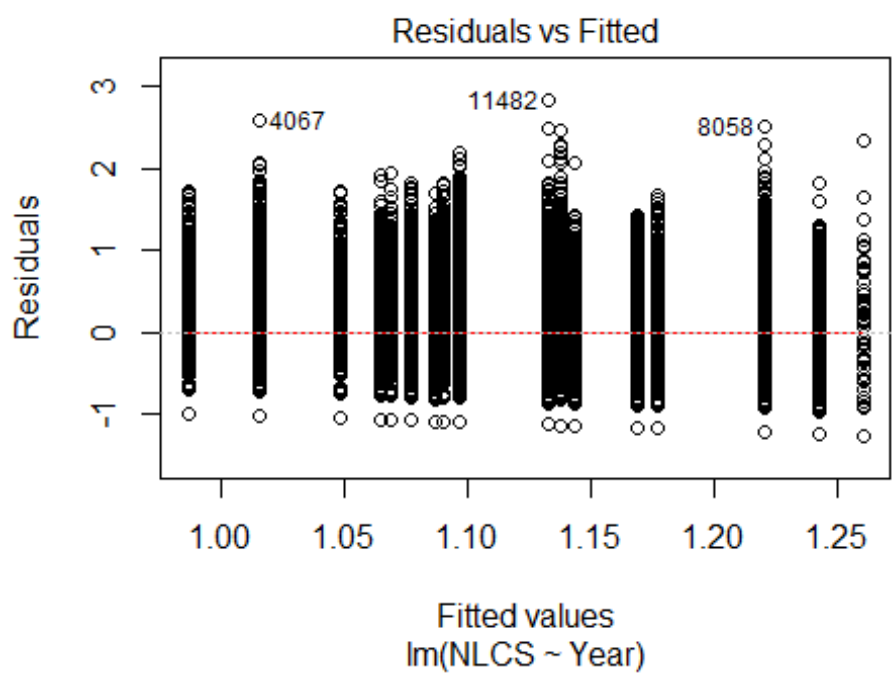



```
## [1] "List of  56 outliers with residuals above  2.5"
##           ScopusId  NLCS Year OneField Fields residuals
## 224     0030151353 3.787 1996     2208      1     2.565
## 340     0030234863 4.961 1996     2208      1     3.739
## 1385    0030173121 5.461 1996     2208      2     4.239
## 2021    0030261036 4.407 1996     2207      2     3.185
## 4202    0030262865 4.271 1996     1705      2     3.049
## 4576    0030232761 3.847 1996     1502      6     2.625
## 5035    0030166739 3.759 1996     1705      2     2.537
## 5055    0030170353 3.905 1996     1705      2     2.683
## 5628    0030126485 3.949 1996     2208      3     2.733
## 5756    0030105412 3.759 1996     2208      1     2.537
## 6298    0030290680 3.728 1996     2208      1     2.506
## 7404    0030287048 3.862 1996     1711      3     2.640
## 8225    0030735427 3.883 1997     2208      1     2.633
## 8231    0030735959 4.307 1997     2208      1     3.057
## 8271    0030846969 3.814 1997     2208      1     2.564
## 8284    0031070220 4.375 1997     2208      1     3.125
## 8558    0031233424 4.026 1997     2208      1     2.776
## 8575    0031236835 4.282 1997     2208      1     3.032
## 8595    0031257246 4.420 1997     2208      1     3.176
```

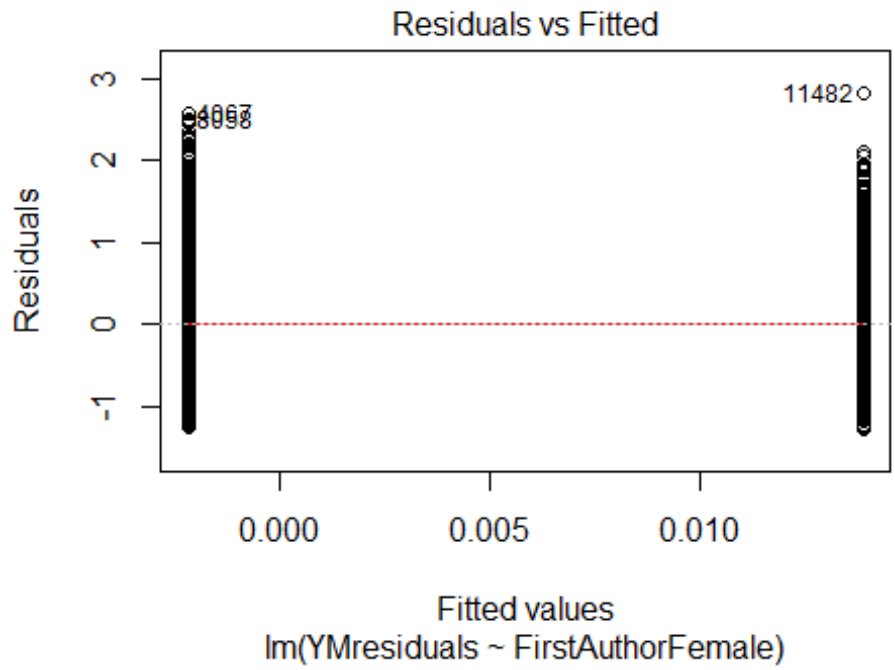
```
## 8679    0031338376 4.599 1997     2208      1     3.349
## 9312    0031102203 3.829 1997     1711      2     2.559
## 12053   0031383380 3.875 1997     1705      4     2.625
## 13718   0031147079 4.121 1997     2208      1     2.871
## 14655   0030643536 3.766 1997     2201      2     2.516
## 14709   0031078854 4.109 1997     1711      2     2.859
## 15694   0031246031 3.912 1997     1705      3     2.662
## 17008   0032138896 4.051 1998     2208      1     2.802
## 17249  33747433684 3.922 1998     2208      1     2.673
## 18241   0032136715 3.940 1998     2208      2     2.691
## 18731   0032308533 4.056 1998     2207      2     2.807
## 18930  33747921774 3.844 1998     1900      2     2.595
## 19247   0032047505 3.905 1998     1706      3     2.656
## 19775   0032204063 4.140 1998     1702      4     2.891
## 20750   0032183752 5.467 1998     1705      2     4.218
## 21975   0032074579 3.796 1998     1705      2     2.547
## 26076   0034156558 3.887 2000     2208      1     2.675
## 31752   0033884858 3.916 2000     1711      2     2.704
## 34945   0035119053 4.046 2001     2208      1     2.805
## 35072  33646900503 3.909 2001     2208      1     2.668
## 36386   0035481826 4.359 2001     2208      1     3.118
## 36725   0035493267 3.845 2001     2208      3     2.604
## 38991   0035364996 3.796 2001     2208      2     2.555
## 46143  33645307063 3.797 2002     2208      1     2.591
## 46610  33646589837 4.485 2002     1706      3     3.258
## 47973   0036747248 3.735 2002     1705      2     2.508
## 48115   0036738266 3.707 2002     1705      3     2.501
## 51612   0036474829 3.755 2002     2208      1     2.549
## 54477   0036608987 3.709 2002     2207      3     2.503
## 54533   0036688074 4.740 2002     1705      3     3.534
## 57609   1642401353 3.922 2003     2202      2     2.600
## 63628  10944266504 4.134 2004     1705      3     2.850
## 64295   4644244041 4.109 2004     1706      3     2.799
## 67847   0742290133 4.216 2004     2208      2     2.906
## 92163  64149119332 4.203 2007     2208      1     2.930
## 92165  64149124758 3.845 2007     2208      1     2.572
## 101063 85032750937 4.080 2008     1711      3     2.833
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##    Min     1Q Median     3Q    Max
## -1.378 -0.521  0.031  0.510  4.238
##
## Coefficients:
##                   Estimate Std. Error t value Pr(>|t|)
## (Intercept)        1.22270    0.01749   69.91  < 2e-16 ***
```

```
## LastAuthorFemale1 -0.02128    0.01321   -1.61  0.10722
## Year1997           0.02785    0.02458    1.13  0.25723
## Year1998           0.02732    0.02452    1.11  0.26532
## Year1999           0.15552    0.05070    3.07  0.00216 **
## Year2000          -0.00907    0.02322   -0.39  0.69592
## Year2001           0.01953    0.02369    0.82  0.40980
## Year2002          -0.01472    0.02242   -0.66  0.51143
## Year2003           0.10026    0.02430    4.13  3.7e-05 ***
## Year2004           0.08918    0.02355    3.79  0.00015 ***
## Year2005           0.13312    0.02250    5.92  3.3e-09 ***
## Year2006           0.10229    0.02169    4.72  2.4e-06 ***
## Year2007           0.05202    0.02162    2.41  0.01611 *
## Year2008           0.02590    0.02183    1.19  0.23543
## Year2009           0.01615    0.02198    0.73  0.46239
## Year2010           0.00777    0.02224    0.35  0.72664
## Year2011          -0.02750    0.02178   -1.26  0.20673
## Year2012          -0.02105    0.02183   -0.96  0.33478
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.763
## Multiple R-squared:  0.00386,    Adjusted R-squared:  0.00348
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
##  3 observations c(167,596,7473) are outliers with |weight| = 0 ( < 2.2e-
06);
##  3788 weights are ~= 1. The remaining 41168 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0006  0.8710  0.9500  0.9100  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          2.22e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Sample size for the above analysis:  44959"
## [1] ""
## [1] ""
## [1] "##################################"
```

```
## [1] "Analysis of AJSC 2209"
## [1] "####################################"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  112 1759 1764 1849 1780 1882 1584 1330 1215 1205 1394 1464 1516 1838 1600
## 2011 2012
## 1556 1516
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##   62  950 1046  980 1003  832  901  726  697  621  729  767  883 1093  953
## 2011 2012
##  910  910
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##   50  802  907  820  839  678  749  607  565  475  609  629  712  886  774
## 2011 2012
##  712  708
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
##  Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 330, df = 16, p-value <2e-16
```

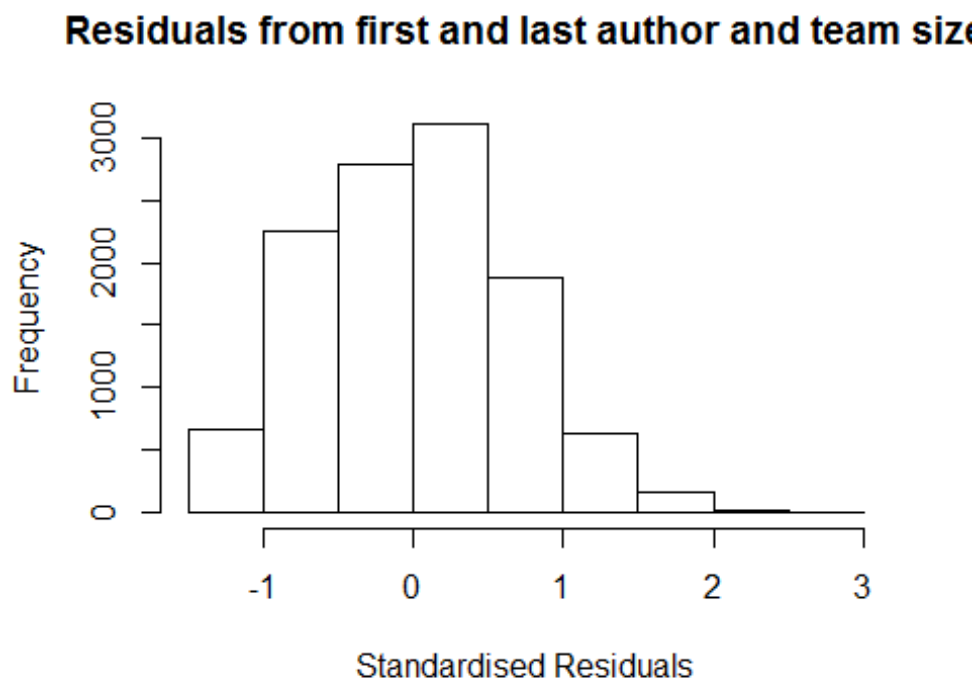


```
##
##  Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 0.85, df = 1, p-value = 0.4
```




```
## [1] "Female first author team size 2018 geometric mean: 2.76110577398757"
## [1] "Male first author team size 2018 geometric mean: 2.67182971171014"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 57000, p-value = 0.4
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 2.4950926538325"
## [1] "Male last author team size 2018 geometric mean: 2.72626347842684"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 42000, p-value = 0.1
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
```

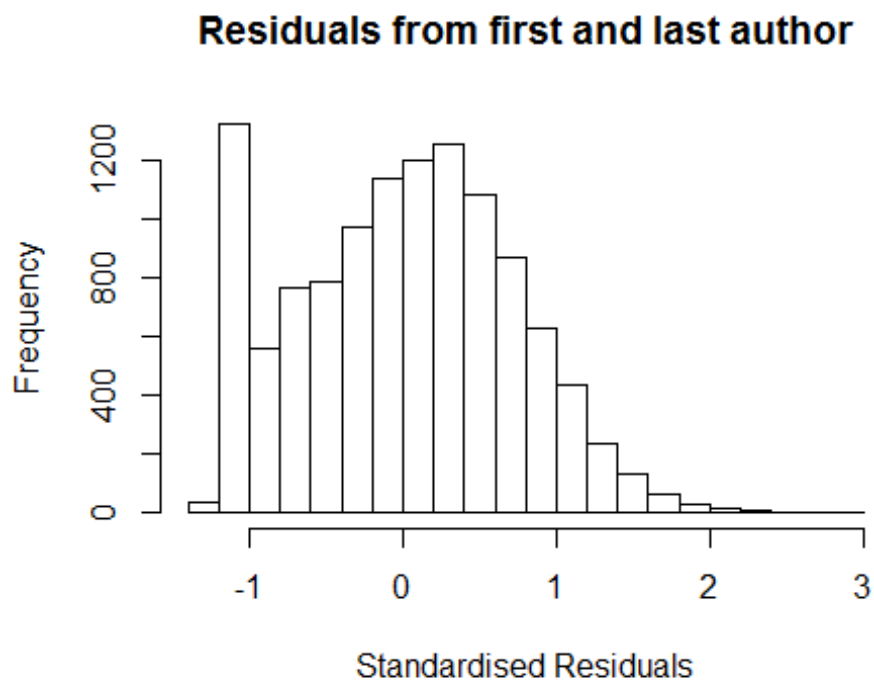
```
## FirstAuthorFemale 1.107  1           1.052
## LastAuthorFemale  1.098  1           1.048
## UniqueAuthors     1.085  4           1.010
## Year              1.088 16           1.003
```




```
## [1] "List of  6 outliers with residuals above  2.5"
##          ScopusId  NLCS Year OneField Fields residuals
## 240   24044543390 3.438 1997     2209      2     2.654
## 255    0030643777 3.593 1997     2207      3     2.809
## 313    0031139613 3.326 1997     2207      3     2.542
## 2529   0032136666 3.207 1998     1400      4     2.532
## 9442   0035081861 3.351 2001     2207      3     2.551
## 11482  0036504271 3.965 2002     2205      4     2.736
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
UniqueAuthors +
##     Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##     k.max = 1000))
##  \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.39746 -0.50961  0.00852  0.46403  2.80860
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
```

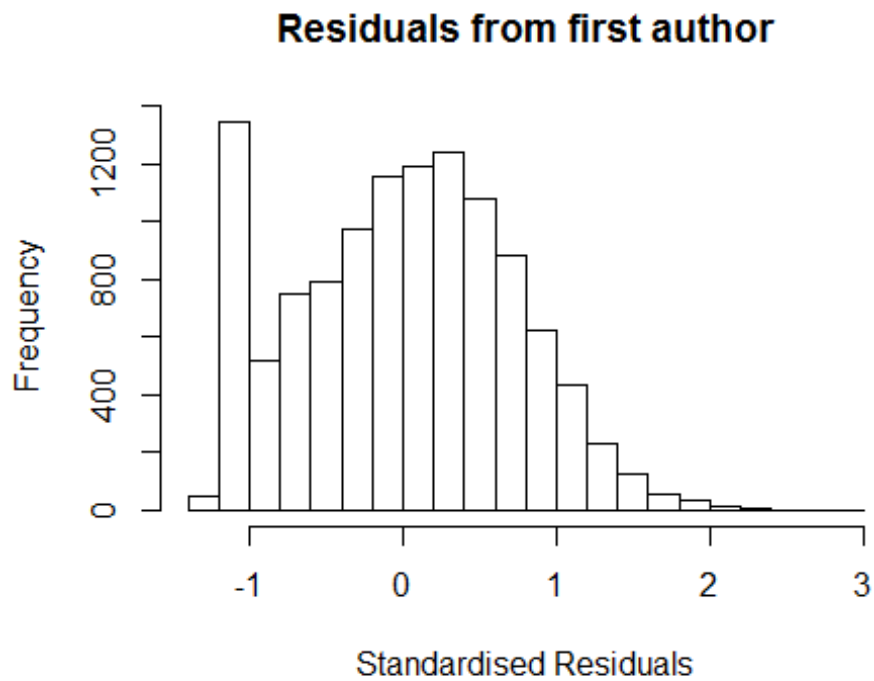
```
## (Intercept)          0.9066     0.1184    7.66    2e-14 ***
## FirstAuthorFemale1   0.0218     0.0193    1.13   0.2588
## LastAuthorFemale1   -0.0669     0.0219   -3.05   0.0023 **
## UniqueAuthors2       0.4945     0.0173   28.57   <2e-16 ***
## UniqueAuthors3       0.5756     0.0191   30.08   <2e-16 ***
## UniqueAuthors4       0.5915     0.0246   24.09   <2e-16 ***
## UniqueAuthors5       0.5694     0.0277   20.53   <2e-16 ***
## Year1997            -0.1222     0.1211   -1.01   0.3132
## Year1998            -0.1867     0.1211   -1.54   0.1233
## Year1999            -0.2633     0.1213   -2.17   0.0300 *
## Year2000            -0.2052     0.1208   -1.70   0.0894 .
## Year2001            -0.1065     0.1218   -0.87   0.3822
## Year2002            -0.2014     0.1206   -1.67   0.0950 .
## Year2003            -0.1352     0.1205   -1.12   0.2621
## Year2004            -0.1472     0.1211   -1.22   0.2243
## Year2005            -0.1127     0.1205   -0.94   0.3496
## Year2006            -0.1841     0.1207   -1.52   0.1273
## Year2007            -0.2354     0.1205   -1.95   0.0507 .
## Year2008            -0.2294     0.1204   -1.90   0.0568 .
## Year2009            -0.3429     0.1197   -2.86   0.0042 **
## Year2010            -0.2445     0.1200   -2.04   0.0415 *
## Year2011            -0.2764     0.1198   -2.31   0.0210 *
## Year2012            -0.3299     0.1199   -2.75   0.0059 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.659
## Multiple R-squared:  0.132,  Adjusted R-squared:  0.131
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
##  879 weights are ~= 1. The remaining 10643 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0302  0.8750  0.9420  0.9070  0.9840  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          8.68e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                    psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
```

```
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##                     GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.063  1           1.031
## LastAuthorFemale  1.059  1           1.029
## Year              1.014 16           1.000
```



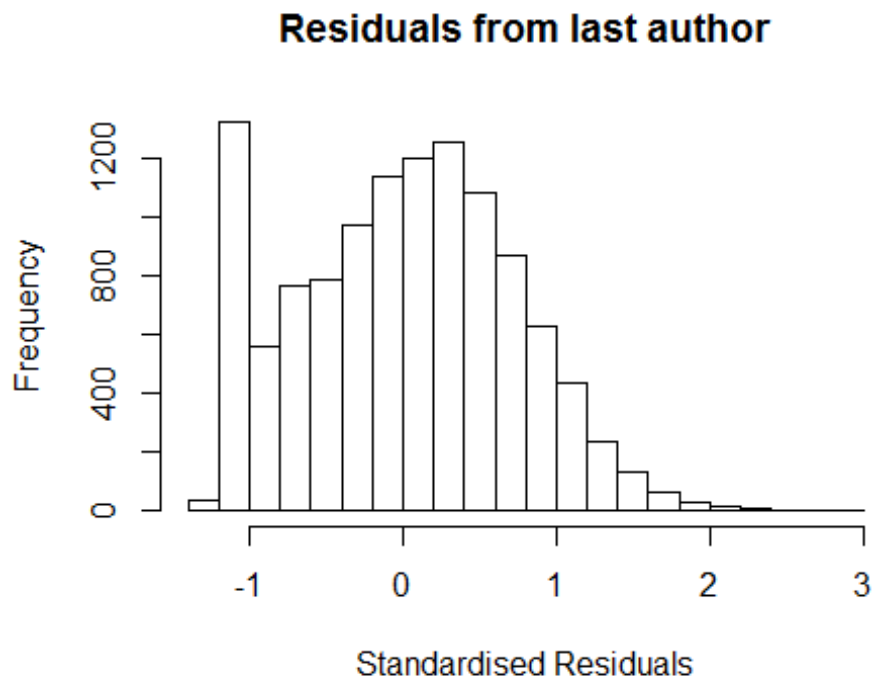

```
## [1] "List of  4 outliers with residuals above  2.5"
##         ScopusId  NLCS Year OneField Fields residuals
## 255   0030643777 3.593 1997     2207      3     2.512
## 4067  0001923938 3.599 1999     2207      3     2.654
## 10389 0036763450 3.614 2002     2205      4     2.521
## 11482 0036504271 3.965 2002     2205      4     2.885
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.2430 -0.5387  0.0295  0.5044  2.8846
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)          1.21469    0.12909    9.41  < 2e-16 ***
## FirstAuthorFemale1  0.06679    0.02017    3.31  0.00093 ***
```

```
## LastAuthorFemale1  -0.07903    0.02330   -3.39  0.00070 ***
## Year1997           -0.13400    0.13277   -1.01  0.31288
## Year1998           -0.19271    0.13269   -1.45  0.14645
## Year1999           -0.26966    0.13271   -2.03  0.04219 *
## Year2000           -0.16528    0.13185   -1.25  0.21004
## Year2001           -0.03852    0.13305   -0.29  0.77218
## Year2002           -0.12209    0.13204   -0.92  0.35517
## Year2003           -0.07622    0.13197   -0.58  0.56358
## Year2004           -0.05695    0.13240   -0.43  0.66708
## Year2005           -0.00276    0.13178   -0.02  0.98329
## Year2006           -0.08650    0.13189   -0.66  0.51192
## Year2007           -0.16633    0.13199   -1.26  0.20763
## Year2008           -0.16165    0.13204   -1.22  0.22087
## Year2009           -0.27150    0.13153   -2.06  0.03903 *
## Year2010           -0.18312    0.13174   -1.39  0.16455
## Year2011           -0.17415    0.13125   -1.33  0.18458
## Year2012           -0.20909    0.13146   -1.59  0.11172
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.759
## Multiple R-squared:  0.0115, Adjusted R-squared:  0.00994
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
##  957 weights are ~= 1. The remaining 10565 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.117   0.864   0.949   0.920   0.985   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          8.68e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.009  1           1.005
## Year              1.009 16           1.000
```




```
## [1] "List of  4 outliers with residuals above  2.5"
##         ScopusId  NLCS Year OneField Fields residuals
## 255   0030643777 3.593 1997     2207      3     2.512
## 4067  0001923938 3.599 1999     2207      3     2.654
## 10389 0036763450 3.614 2002     2205      4     2.521
## 11482 0036504271 3.965 2002     2205      4     2.885
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.2537 -0.5405  0.0316  0.5050  2.8350
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.21150    0.12965    9.34   <2e-16 ***
## FirstAuthorFemale1  0.04216    0.02013    2.09    0.036 *
## Year1997           -0.13628    0.13331   -1.02    0.307
## Year1998           -0.19420    0.13322   -1.46    0.145
## Year1999           -0.27196    0.13327   -2.04    0.041 *
## Year2000           -0.16761    0.13238   -1.27    0.205
## Year2001           -0.03867    0.13359   -0.29    0.772
## Year2002           -0.12364    0.13258   -0.93    0.351
## Year2003           -0.07762    0.13250   -0.59    0.558
```

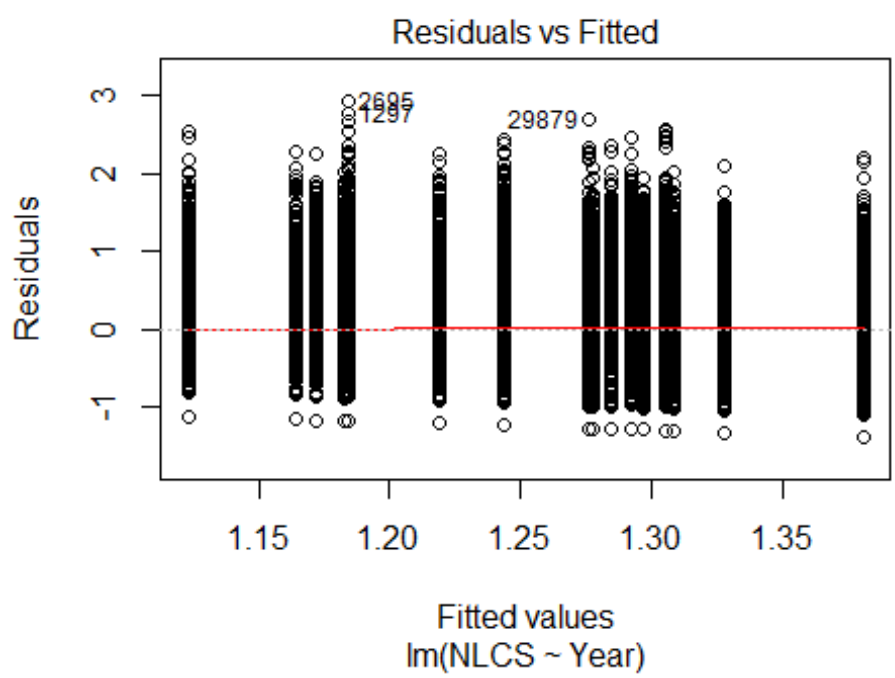
```
## Year2004           -0.05847    0.13292   -0.44    0.660
## Year2005           -0.00492    0.13231   -0.04    0.970
## Year2006           -0.08913    0.13241   -0.67    0.501
## Year2007           -0.16967    0.13251   -1.28    0.200
## Year2008           -0.16422    0.13257   -1.24    0.215
## Year2009           -0.27275    0.13207   -2.07    0.039 *
## Year2010           -0.18548    0.13228   -1.40    0.161
## Year2011           -0.17807    0.13179   -1.35    0.177
## Year2012           -0.21088    0.13199   -1.60    0.110
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.759
## Multiple R-squared:  0.0105, Adjusted R-squared:  0.00905
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
##  958 weights are ~= 1. The remaining 10564 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.133   0.865   0.948   0.920   0.985   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          8.68e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##                   GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.006  1           1.003
## Year             1.006 16           1.000
```



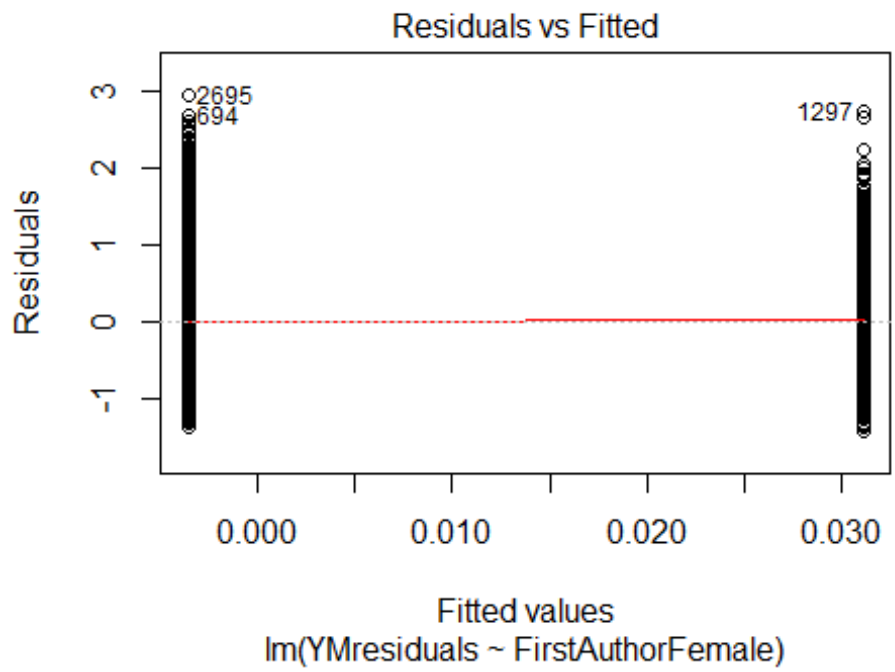

```
## [1] "List of  4 outliers with residuals above  2.5"
##         ScopusId  NLCS Year OneField Fields residuals
## 255   0030643777 3.593 1997     2207      3     2.512
## 4067  0001923938 3.599 1999     2207      3     2.654
## 10389 0036763450 3.614 2002     2205      4     2.521
## 11482 0036504271 3.965 2002     2205      4     2.885
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.2209 -0.5403  0.0303  0.5042  2.9192
##
## Coefficients:
##                   Estimate Std. Error t value Pr(>|t|)
## (Intercept)        1.22088    0.12972    9.41   <2e-16 ***
## LastAuthorFemale1 -0.05361    0.02255   -2.38    0.017 *
## Year1997          -0.13477    0.13336   -1.01    0.312
## Year1998          -0.19515    0.13329   -1.46    0.143
## Year1999          -0.27149    0.13331   -2.04    0.042 *
## Year2000          -0.16719    0.13246   -1.26    0.207
## Year2001          -0.03661    0.13364   -0.27    0.784
## Year2002          -0.12146    0.13262   -0.92    0.360
## Year2003          -0.07598    0.13256   -0.57    0.567
```

```
## Year2004          -0.05778    0.13299   -0.43    0.664
## Year2005          -0.00307    0.13238   -0.02    0.982
## Year2006          -0.08630    0.13248   -0.65    0.515
## Year2007          -0.16678    0.13259   -1.26    0.208
## Year2008          -0.15880    0.13262   -1.20    0.231
## Year2009          -0.26916    0.13212   -2.04    0.042 *
## Year2010          -0.18203    0.13233   -1.38    0.169
## Year2011          -0.17322    0.13184   -1.31    0.189
## Year2012          -0.20754    0.13203   -1.57    0.116
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.76
## Multiple R-squared:  0.0107, Adjusted R-squared:  0.0092
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
##  924 weights are ~= 1. The remaining 10598 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.107   0.863   0.949   0.920   0.985   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          8.68e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Sample size for the above analysis:  11522"
## [1] ""
## [1] ""
## [1] "####################################"
## [1] "Analysis of AJSC 2210"
## [1] "####################################"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 4775 4653 4381 2631 2889 4532 4197 3687 3826 3880 4192 4483 4690 5263 4727
## 2011 2012
## 4973 4676
##
```

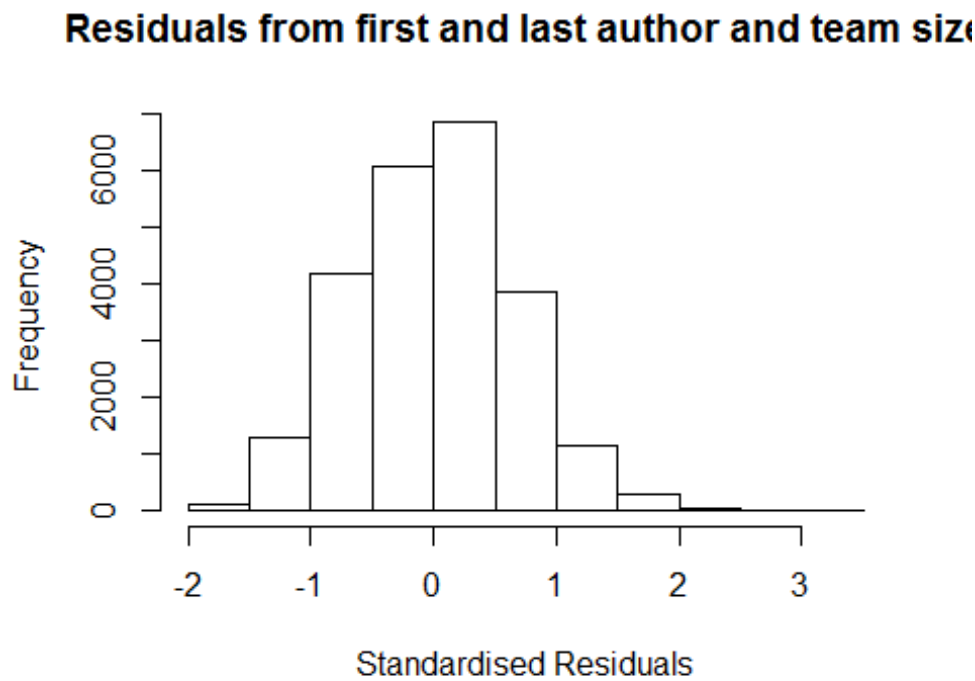
```
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 1458 1389 1416 1063 1049 1365 1680 1498 1599 1639 1798 2075 2145 2567 2265
## 2011 2012
## 2502 2332
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 1211 1146 1162  885  887 1088 1387 1190 1285 1291 1418 1648 1701 2025 1798
## 2011 2012
## 1938 1854
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
##  Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 420, df = 16, p-value <2e-16
```




```
##
##  Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 0.96, df = 1, p-value = 0.3
```



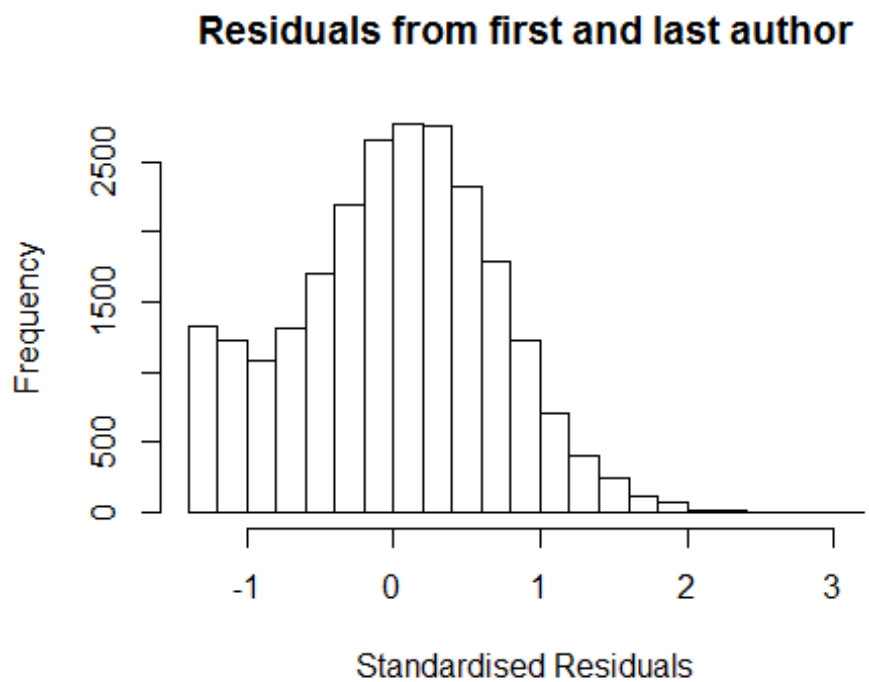

```
## [1] "Female first author team size 2018 geometric mean: 3.16264851972005"
## [1] "Male first author team size 2018 geometric mean: 2.94373925365379"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 270000, p-value = 0.09
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 3.08408854760063"
## [1] "Male last author team size 2018 geometric mean: 2.95967159148615"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 230000, p-value = 0.3
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.035  1           1.018
## LastAuthorFemale  1.028  1           1.014
## UniqueAuthors     1.076  4           1.009
## Year              1.071 16           1.002
```




```
## [1] "List of  10 outliers with residuals above  2.5"
##          ScopusId  NLCS Year OneField Fields residuals
## 694    0030232761 3.847 1996     1502      6     2.622
## 1297   0030126485 3.949 1996     2208      3     3.185
## 1601   0029724421 3.734 1996     2210      1     2.509
## 2695   0029777844 4.113 1996     2210      3     3.312
## 3830   0030290066 3.443 1996     2210      3     2.642
## 5309   0031097861 3.492 1997     2205      4     2.601
## 12233  0031700973 3.543 1998     2210      3     2.608
## 16890  0032632334 3.598 1999     2210      1     2.848
## 28040  0036591982 3.537 2002     1500      5     2.666
## 47366 33645634748 3.428 2006     2210      4     2.529
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
UniqueAuthors +
##     Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##     k.max = 1000))
##  \--> method = "MM"
## Residuals:
##    Min     1Q Median     3Q    Max
## -1.656 -0.456  0.019  0.450  3.312
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)          0.80071    0.02524   31.72  < 2e-16 ***
```

```
## FirstAuthorFemale1 -0.01732    0.01462   -1.18  0.23627
## LastAuthorFemale1  -0.01968    0.01639   -1.20  0.22993
## UniqueAuthors2      0.42386    0.01326   31.96  < 2e-16 ***
## UniqueAuthors3      0.50487    0.01415   35.69  < 2e-16 ***
## UniqueAuthors4      0.54579    0.01661   32.85  < 2e-16 ***
## UniqueAuthors5      0.71300    0.01662   42.90  < 2e-16 ***
## Year1997            0.09011    0.03321    2.71  0.00667 **
## Year1998            0.13406    0.03351    4.00  6.3e-05 ***
## Year1999           -0.05031    0.03482   -1.45  0.14846
## Year2000            0.09496    0.03451    2.75  0.00594 **
## Year2001            0.09012    0.03659    2.46  0.01379 *
## Year2002            0.07060    0.03143    2.25  0.02468 *
## Year2003            0.10861    0.03077    3.53  0.00042 ***
## Year2004            0.08291    0.03007    2.76  0.00583 **
## Year2005            0.14247    0.02880    4.95  7.6e-07 ***
## Year2006            0.09798    0.02882    3.40  0.00068 ***
## Year2007            0.02470    0.02811    0.88  0.37968
## Year2008            0.00538    0.02822    0.19  0.84879
## Year2009           -0.07020    0.02784   -2.52  0.01169 *
## Year2010           -0.04346    0.02804   -1.55  0.12121
## Year2011           -0.07527    0.02781   -2.71  0.00680 **
## Year2012           -0.10099    0.02793   -3.62  0.00030 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.66
## Multiple R-squared:  0.115,  Adjusted R-squared:  0.114
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
##  2 observations c(342,715) are outliers with |weight| = 0 ( < 4.2e-06);
##  2012 weights are ~= 1. The remaining 21900 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0233  0.8690  0.9490  0.9080  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          4.18e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                    psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
```

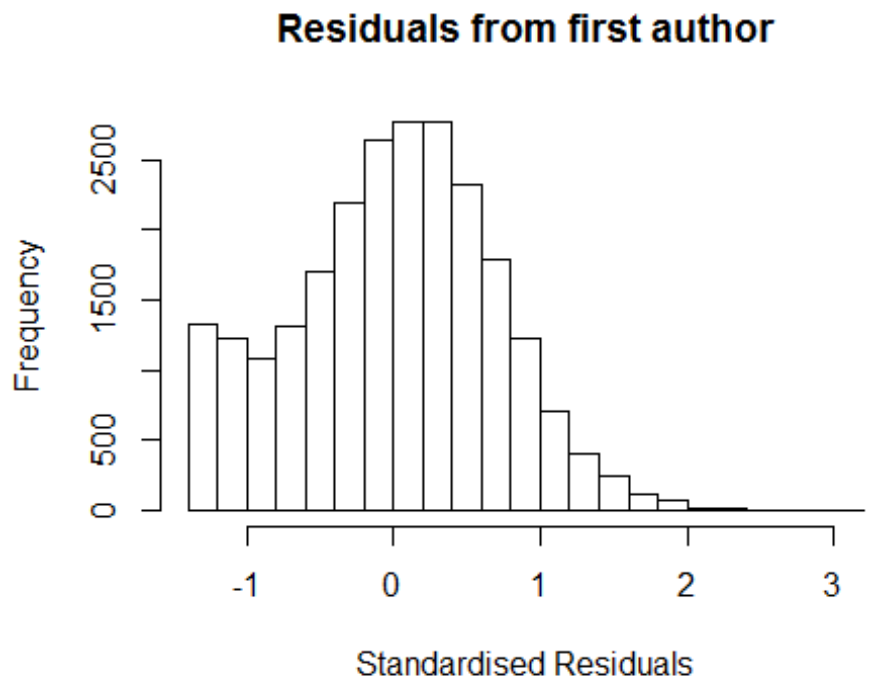
```
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##                     GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.010  1           1.005
## LastAuthorFemale  1.010  1           1.005
## Year              1.006 16           1.000
```




```
## [1] "List of  7 outliers with residuals above  2.5"
##         ScopusId  NLCS Year OneField Fields residuals
## 694   0030232761 3.847 1996     1502      6     2.736
## 1297  0030126485 3.949 1996     2208      3     2.791
## 1601  0029724421 3.734 1996     2210      1     2.623
## 2695  0029777844 4.113 1996     2210      3     3.002
## 16890 0032632334 3.598 1999     2210      1     2.529
## 22070 0035493267 3.845 2001     2208      3     2.578
## 29879 0036504271 3.965 2002     2205      4     2.662
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.3919 -0.4729  0.0355  0.4758  3.0022
##
## Coefficients:
```

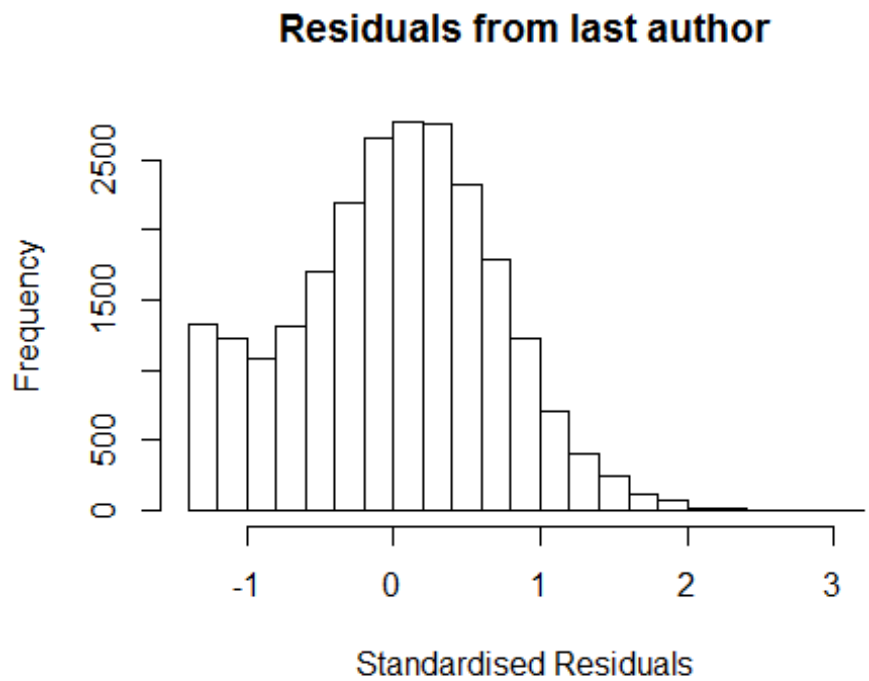
```
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)           1.1108     0.0249   44.60  < 2e-16 ***
## FirstAuthorFemale1    0.0419     0.0153    2.73  0.00632 **
## LastAuthorFemale1     0.0048     0.0173    0.28  0.78108
## Year1997              0.0900     0.0348    2.59  0.00961 **
## Year1998              0.1525     0.0348    4.39  1.2e-05 ***
## Year1999             -0.0421     0.0367   -1.15  0.25134
## Year2000              0.1464     0.0366    4.00  6.4e-05 ***
## Year2001              0.1559     0.0371    4.20  2.7e-05 ***
## Year2002              0.1454     0.0324    4.49  7.3e-06 ***
## Year2003              0.1686     0.0323    5.22  1.8e-07 ***
## Year2004              0.1523     0.0321    4.74  2.1e-06 ***
## Year2005              0.2343     0.0302    7.76  8.9e-15 ***
## Year2006              0.1983     0.0301    6.58  4.8e-11 ***
## Year2007              0.1357     0.0294    4.61  4.0e-06 ***
## Year2008              0.1074     0.0298    3.61  0.00031 ***
## Year2009              0.0276     0.0295    0.94  0.34952
## Year2010              0.0773     0.0297    2.60  0.00920 **
## Year2011              0.0494     0.0293    1.69  0.09102 .
## Year2012              0.0289     0.0296    0.98  0.32815
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.697
## Multiple R-squared:  0.00997,    Adjusted R-squared:  0.00922
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
##  2032 weights are ~= 1. The remaining 21882 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.024   0.864   0.949   0.908   0.985   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          4.18e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##                     GVIF Df GVIF^(1/(2*Df))
```

```
## FirstAuthorFemale 1.003  1           1.002
## Year              1.003 16           1.000
```




```
## [1] "List of  7 outliers with residuals above  2.5"
##         ScopusId  NLCS Year OneField Fields residuals
## 694   0030232761 3.847 1996     1502      6     2.736
## 1297  0030126485 3.949 1996     2208      3     2.791
## 1601  0029724421 3.734 1996     2210      1     2.623
## 2695  0029777844 4.113 1996     2210      3     3.002
## 16890 0032632334 3.598 1999     2210      1     2.529
## 22070 0035493267 3.845 2001     2208      3     2.578
## 29879 0036504271 3.965 2002     2205      4     2.662
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.3882 -0.4727  0.0355  0.4758  3.0019
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)          1.1111     0.0249   44.66  < 2e-16 ***
## FirstAuthorFemale1   0.0428     0.0154    2.78  0.00549 **
## Year1997             0.0899     0.0347    2.59  0.00965 **
```

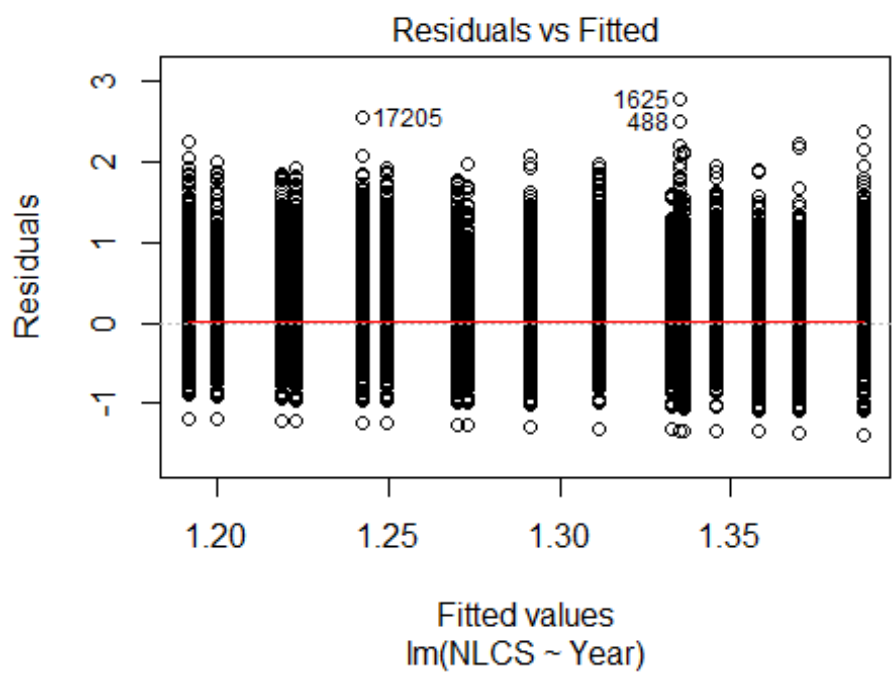
```
## Year1998             0.1524     0.0348    4.38  1.2e-05 ***
## Year1999            -0.0420     0.0367   -1.14  0.25261
## Year2000             0.1464     0.0366    4.00  6.4e-05 ***
## Year2001             0.1559     0.0371    4.20  2.7e-05 ***
## Year2002             0.1454     0.0324    4.49  7.2e-06 ***
## Year2003             0.1686     0.0323    5.22  1.8e-07 ***
## Year2004             0.1523     0.0321    4.74  2.1e-06 ***
## Year2005             0.2343     0.0302    7.76  8.8e-15 ***
## Year2006             0.1983     0.0301    6.58  4.7e-11 ***
## Year2007             0.1357     0.0294    4.61  4.0e-06 ***
## Year2008             0.1074     0.0298    3.60  0.00031 ***
## Year2009             0.0276     0.0295    0.94  0.34885
## Year2010             0.0773     0.0297    2.60  0.00923 **
## Year2011             0.0495     0.0293    1.69  0.09063 .
## Year2012             0.0291     0.0296    0.98  0.32553
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.697
## Multiple R-squared:  0.00996,    Adjusted R-squared:  0.00926
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
##  2025 weights are ~= 1. The remaining 21889 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0241  0.8640  0.9490  0.9080  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          4.18e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##                   GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.004  1           1.002
## Year             1.004 16           1.000
```



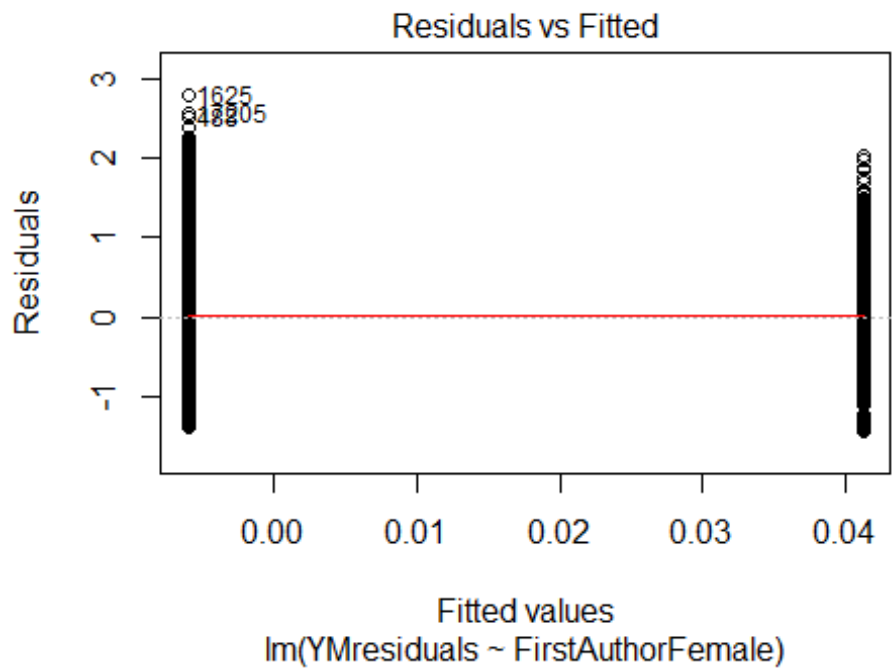

```
## [1] "List of  7 outliers with residuals above  2.5"
##         ScopusId  NLCS Year OneField Fields residuals
## 694   0030232761 3.847 1996     1502      6     2.736
## 1297  0030126485 3.949 1996     2208      3     2.791
## 1601  0029724421 3.734 1996     2210      1     2.623
## 2695  0029777844 4.113 1996     2210      3     3.002
## 16890 0032632334 3.598 1999     2210      1     2.529
## 22070 0035493267 3.845 2001     2208      3     2.578
## 29879 0036504271 3.965 2002     2205      4     2.662
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.3627 -0.4732  0.0358  0.4759  2.9991
##
## Coefficients:
##                   Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.1139     0.0249   44.81  < 2e-16 ***
## LastAuthorFemale1   0.0140     0.0172    0.81  0.41715
## Year1997            0.0897     0.0347    2.58  0.00989 **
## Year1998            0.1527     0.0348    4.39  1.1e-05 ***
## Year1999           -0.0418     0.0367   -1.14  0.25484
## Year2000            0.1463     0.0366    3.99  6.5e-05 ***
```

```
## Year2001            0.1562     0.0371    4.20  2.6e-05 ***
## Year2002            0.1460     0.0324    4.51  6.6e-06 ***
## Year2003            0.1690     0.0323    5.23  1.7e-07 ***
## Year2004            0.1519     0.0321    4.73  2.3e-06 ***
## Year2005            0.2348     0.0302    7.77  8.2e-15 ***
## Year2006            0.1986     0.0301    6.59  4.6e-11 ***
## Year2007            0.1357     0.0294    4.61  4.0e-06 ***
## Year2008            0.1081     0.0298    3.63  0.00029 ***
## Year2009            0.0283     0.0295    0.96  0.33843
## Year2010            0.0779     0.0297    2.62  0.00871 **
## Year2011            0.0503     0.0293    1.72  0.08569 .
## Year2012            0.0300     0.0296    1.01  0.31062
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.697
## Multiple R-squared:  0.00965,    Adjusted R-squared:  0.00895
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
##  2039 weights are ~= 1. The remaining 21875 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0248  0.8640  0.9490  0.9080  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          4.18e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Sample size for the above analysis:  23914"
## [1] ""
## [1] ""
## [1] "##################################"
## [1] "Analysis of AJSC 2211"
## [1] "##################################"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 3305 3417 3330 1558 2124 3494 3032 2761 2902 2851 3160 3150 3239 3732 3533
```

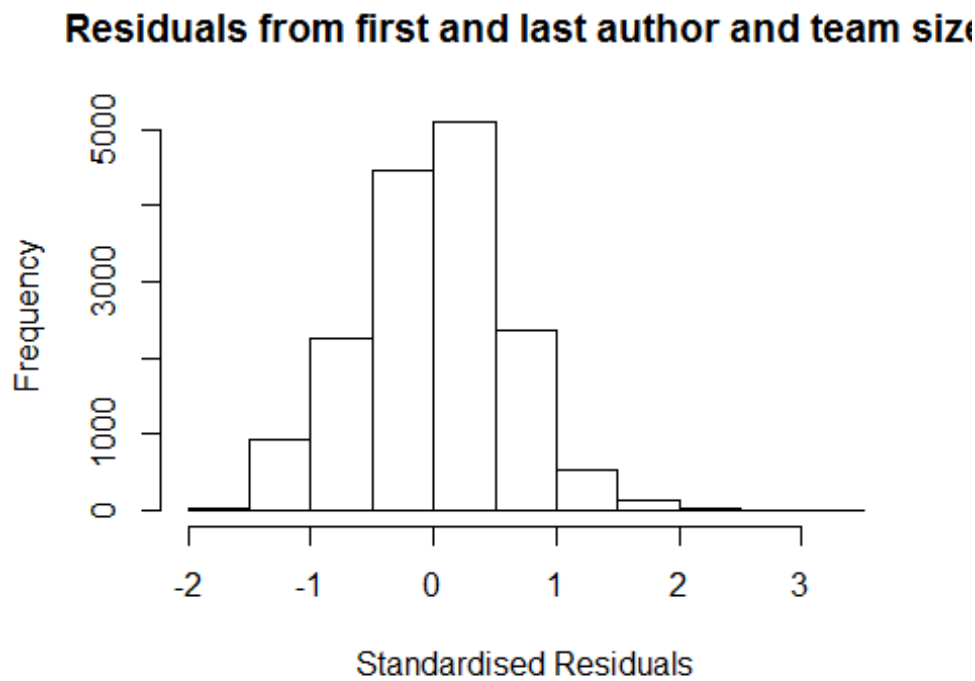
```
## 2011 2012
## 3742 3434
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  901  945  956  512  705  961 1097 1007 1075 1121 1289 1402 1365 1739 1599
## 2011 2012
## 1717 1637
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  730  762  778  409  574  767  884  805  841  872  990 1107 1060 1347 1252
## 2011 2012
## 1315 1278
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
##  Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 320, df = 16, p-value <2e-16
```




```
##
##  Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 13, df = 1, p-value = 3e-04
```

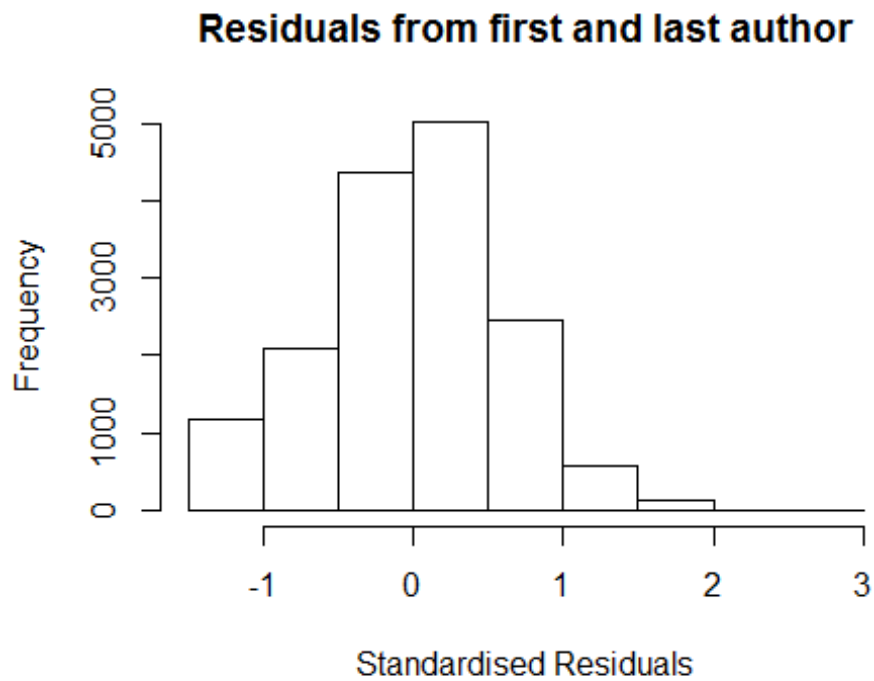



```
## [1] "Female first author team size 2018 geometric mean: 3.5319912138151"
## [1] "Male first author team size 2018 geometric mean: 3.07929775314843"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 140000, p-value = 0.005
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 3.6824829792554"
## [1] "Male last author team size 2018 geometric mean: 3.0645471164641"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 140000, p-value = 1e-04
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##                     GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.056  1           1.028
## LastAuthorFemale  1.029  1           1.014
## UniqueAuthors     1.127  4           1.015
## Year              1.117 16           1.003
```



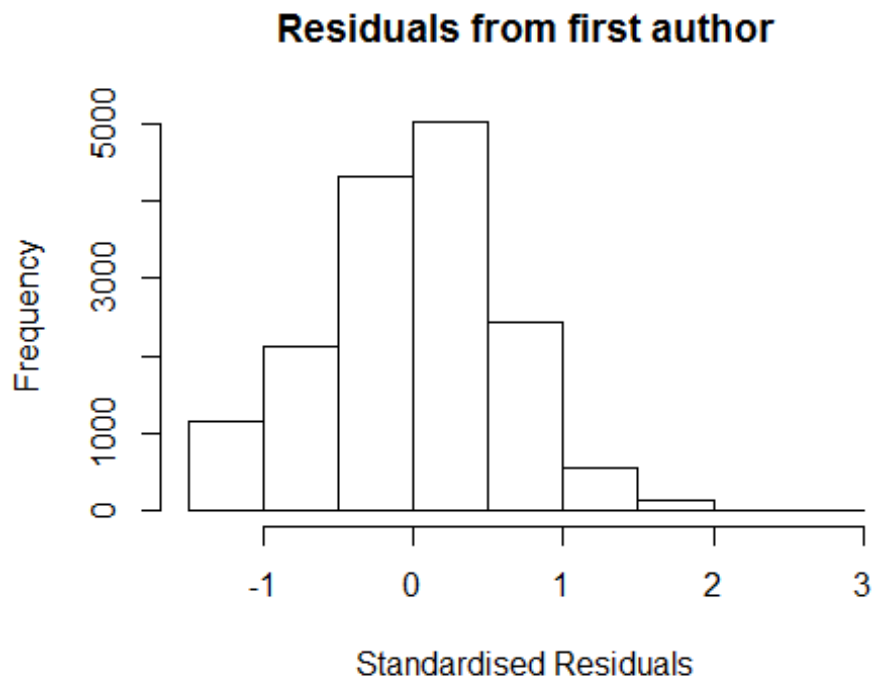

```
## [1] "List of  2 outliers with residuals above  2.5"
##        ScopusId  NLCS Year OneField Fields residuals
## 488  0030232761 3.847 1996     1502      6     2.510
## 1625 0029777844 4.113 1996     2210      3     3.066
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
UniqueAuthors +
##     Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##     k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.6265 -0.3928  0.0204  0.3972  3.0662
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.04676    0.03081   33.97  < 2e-16 ***
## FirstAuthorFemale1  0.01231    0.01424    0.86   0.3873
## LastAuthorFemale1   0.01519    0.01683    0.90   0.3668
## UniqueAuthors2      0.29025    0.01608   18.05  < 2e-16 ***
## UniqueAuthors3      0.34797    0.01692   20.57  < 2e-16 ***
## UniqueAuthors4      0.38565    0.01927   20.01  < 2e-16 ***
## UniqueAuthors5      0.50190    0.01898   26.44  < 2e-16 ***
## Year1997            0.01288    0.03884    0.33   0.7402
## Year1998            0.07787    0.03790    2.05   0.0399 *
```

```
## Year1999            0.04219    0.04231    1.00   0.3187
## Year2000            0.00807    0.03973    0.20   0.8390
## Year2001           -0.08422    0.04278   -1.97   0.0490 *
## Year2002           -0.11417    0.03746   -3.05   0.0023 **
## Year2003           -0.06006    0.03579   -1.68   0.0933 .
## Year2004           -0.00343    0.03430   -0.10   0.9203
## Year2005            0.00708    0.03300    0.21   0.8301
## Year2006           -0.03056    0.03329   -0.92   0.3586
## Year2007           -0.06674    0.03299   -2.02   0.0431 *
## Year2008           -0.09190    0.03322   -2.77   0.0057 **
## Year2009           -0.15133    0.03283   -4.61  4.1e-06 ***
## Year2010           -0.13968    0.03243   -4.31  1.7e-05 ***
## Year2011           -0.18605    0.03266   -5.70  1.2e-08 ***
## Year2012           -0.18968    0.03287   -5.77  8.0e-09 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.584
## Multiple R-squared:  0.07,   Adjusted R-squared:  0.0687
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
##  observation 328 is an outlier with |weight| = 0 ( < 6.3e-06);
##  1346 weights are ~= 1. The remaining 14424 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0249  0.8650  0.9500  0.9010  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          6.34e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##                     GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.027  1           1.013
## LastAuthorFemale  1.023  1           1.011
## Year              1.024 16           1.001
```



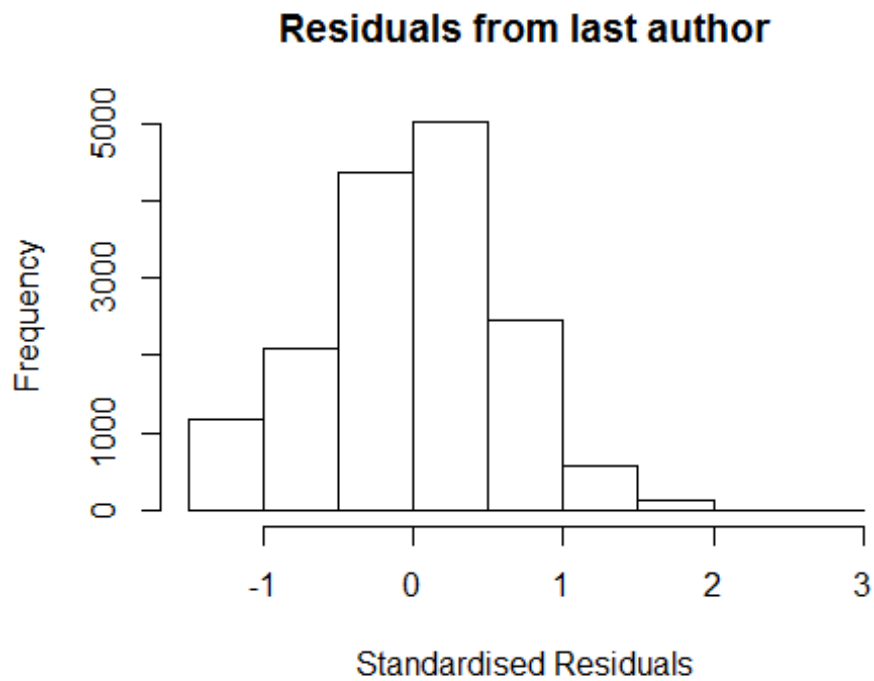

```
## [1] "List of  2 outliers with residuals above  2.5"
##        ScopusId  NLCS Year OneField Fields residuals
## 488  0030232761 3.847 1996     1502      6     2.555
## 1625 0029777844 4.113 1996     2210      3     2.821
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.4611 -0.4004  0.0268  0.4084  2.8213
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)          1.29172    0.02951   43.78  < 2e-16 ***
## FirstAuthorFemale1   0.06340    0.01445    4.39  1.2e-05 ***
## LastAuthorFemale1    0.03298    0.01735    1.90  0.05739 .
## Year1997             0.00506    0.03999    0.13  0.89940
## Year1998             0.07297    0.03909    1.87  0.06194 .
## Year1999             0.05560    0.04409    1.26  0.20733
## Year2000             0.03484    0.04085    0.85  0.39371
## Year2001            -0.06104    0.04265   -1.43  0.15237
## Year2002            -0.08091    0.03834   -2.11  0.03484 *
## Year2003            -0.04522    0.03698   -1.22  0.22135
## Year2004             0.03486    0.03574    0.98  0.32947
```

```
## Year2005            0.05516    0.03427    1.61  0.10757
## Year2006            0.02192    0.03448    0.64  0.52501
## Year2007           -0.01074    0.03432   -0.31  0.75435
## Year2008           -0.03352    0.03443   -0.97  0.33025
## Year2009           -0.08911    0.03420   -2.61  0.00919 **
## Year2010           -0.07009    0.03371   -2.08  0.03759 *
## Year2011           -0.11013    0.03393   -3.25  0.00117 **
## Year2012           -0.11390    0.03427   -3.32  0.00089 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.598
## Multiple R-squared:  0.0103, Adjusted R-squared:  0.00919
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
##  observation 328 is an outlier with |weight| = 0 ( < 6.3e-06);
##  1314 weights are ~= 1. The remaining 14456 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0285  0.8650  0.9500  0.9010  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          6.34e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.015  1           1.008
## Year              1.015 16           1.000
```




```
## [1] "List of  2 outliers with residuals above  2.5"
##        ScopusId  NLCS Year OneField Fields residuals
## 488  0030232761 3.847 1996     1502      6     2.555
## 1625 0029777844 4.113 1996     2210      3     2.821
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.4349 -0.4008  0.0261  0.4092  2.8201
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)          1.29291    0.02948   43.85   <2e-16 ***
## FirstAuthorFemale1   0.06848    0.01439    4.76    2e-06 ***
## Year1997             0.00584    0.03997    0.15   0.8838
## Year1998             0.07352    0.03906    1.88   0.0598 .
## Year1999             0.05613    0.04408    1.27   0.2028
## Year2000             0.03524    0.04085    0.86   0.3884
## Year2001            -0.05997    0.04263   -1.41   0.1595
## Year2002            -0.07971    0.03830   -2.08   0.0374 *
## Year2003            -0.04466    0.03696   -1.21   0.2270
## Year2004             0.03627    0.03570    1.02   0.3097
## Year2005             0.05589    0.03425    1.63   0.1027
```

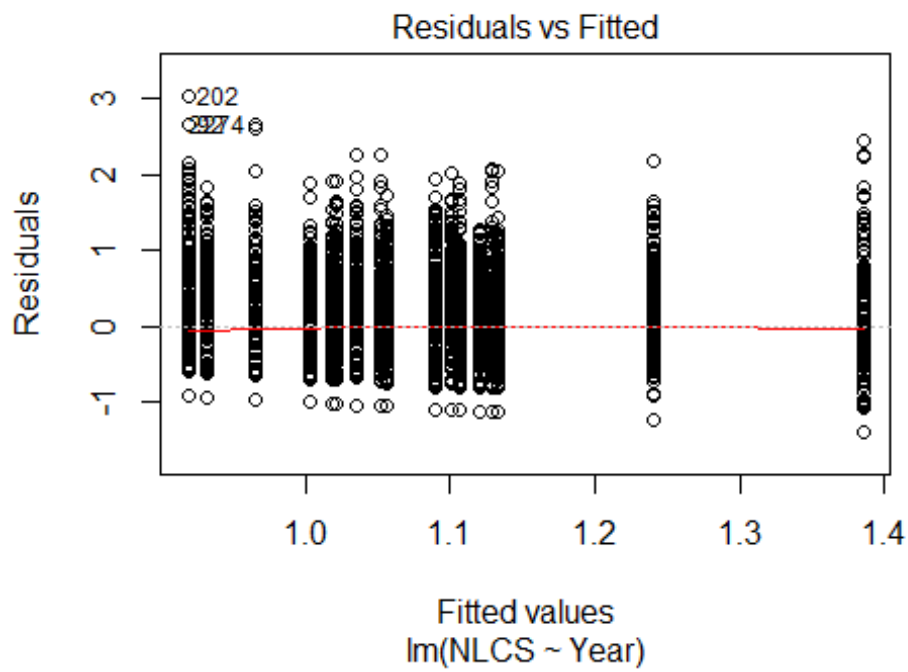
```
## Year2006            0.02316    0.03444    0.67   0.5014
## Year2007           -0.00942    0.03429   -0.27   0.7836
## Year2008           -0.03231    0.03439   -0.94   0.3475
## Year2009           -0.08725    0.03415   -2.56   0.0106 *
## Year2010           -0.06911    0.03368   -2.05   0.0402 *
## Year2011           -0.10831    0.03388   -3.20   0.0014 **
## Year2012           -0.11168    0.03421   -3.26   0.0011 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.598
## Multiple R-squared:  0.0101, Adjusted R-squared:  0.009
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
##  observation 328 is an outlier with |weight| = 0 ( < 6.3e-06);
##  1329 weights are ~= 1. The remaining 14441 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0289  0.8650  0.9500  0.9010  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          6.34e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##                   GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.011  1           1.006
## Year             1.011 16           1.000
```



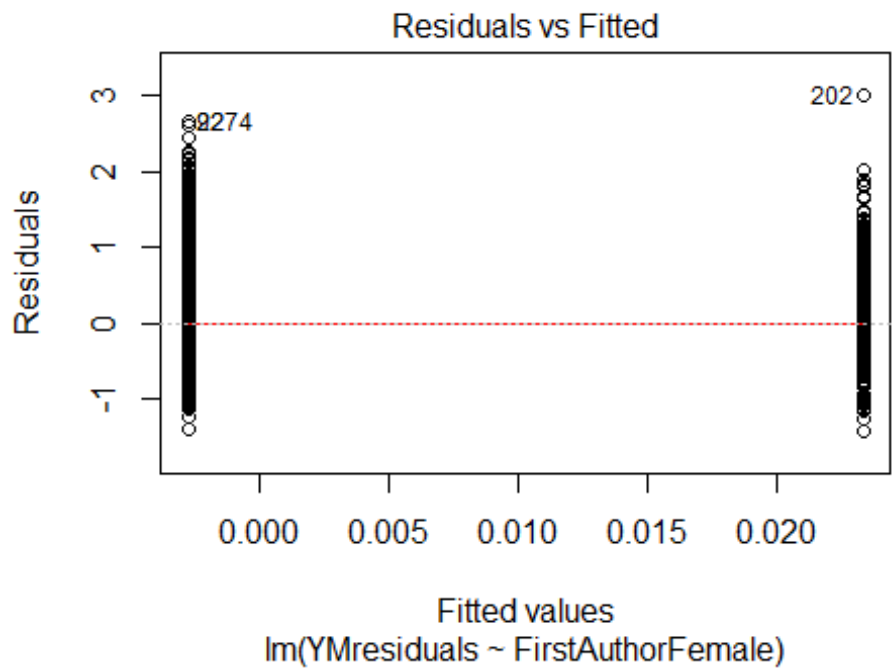

```
## [1] "List of  2 outliers with residuals above  2.5"
##        ScopusId  NLCS Year OneField Fields residuals
## 488  0030232761 3.847 1996     1502      6     2.555
## 1625 0029777844 4.113 1996     2210      3     2.821
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##    Min     1Q Median     3Q    Max
## -1.416 -0.401  0.028  0.410  2.815
##
## Coefficients:
##                   Estimate Std. Error t value Pr(>|t|)
## (Intercept)        1.29810    0.02941   44.14   <2e-16 ***
## LastAuthorFemale1  0.04621    0.01723    2.68   0.0073 **
## Year1997           0.00363    0.03997    0.09   0.9277
## Year1998           0.07187    0.03906    1.84   0.0658 .
## Year1999           0.05467    0.04406    1.24   0.2147
## Year2000           0.03584    0.04088    0.88   0.3807
## Year2001          -0.06112    0.04262   -1.43   0.1516
## Year2002          -0.07992    0.03828   -2.09   0.0368 *
## Year2003          -0.04599    0.03696   -1.24   0.2134
## Year2004           0.03443    0.03573    0.96   0.3353
## Year2005           0.05546    0.03429    1.62   0.1058
```

```
## Year2006           0.02270    0.03446    0.66   0.5101
## Year2007          -0.01088    0.03431   -0.32   0.7513
## Year2008          -0.03392    0.03443   -0.99   0.3245
## Year2009          -0.08798    0.03420   -2.57   0.0101 *
## Year2010          -0.06717    0.03369   -1.99   0.0462 *
## Year2011          -0.10771    0.03393   -3.17   0.0015 **
## Year2012          -0.11132    0.03428   -3.25   0.0012 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.599
## Multiple R-squared:  0.00915,    Adjusted R-squared:  0.00808
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
##  observation 328 is an outlier with |weight| = 0 ( < 6.3e-06);
##  1331 weights are ~= 1. The remaining 14439 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0302  0.8660  0.9500  0.9010  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          6.34e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                  psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Sample size for the above analysis:  15771"
## [1] ""
## [1] ""
## [1] "##################################"
## [1] "Analysis of AJSC 2212"
## [1] "##################################"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  470  468  391  440  365  406  335  357  363  391  365  375  398  417  429
## 2011 2012
##  402  396
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
```

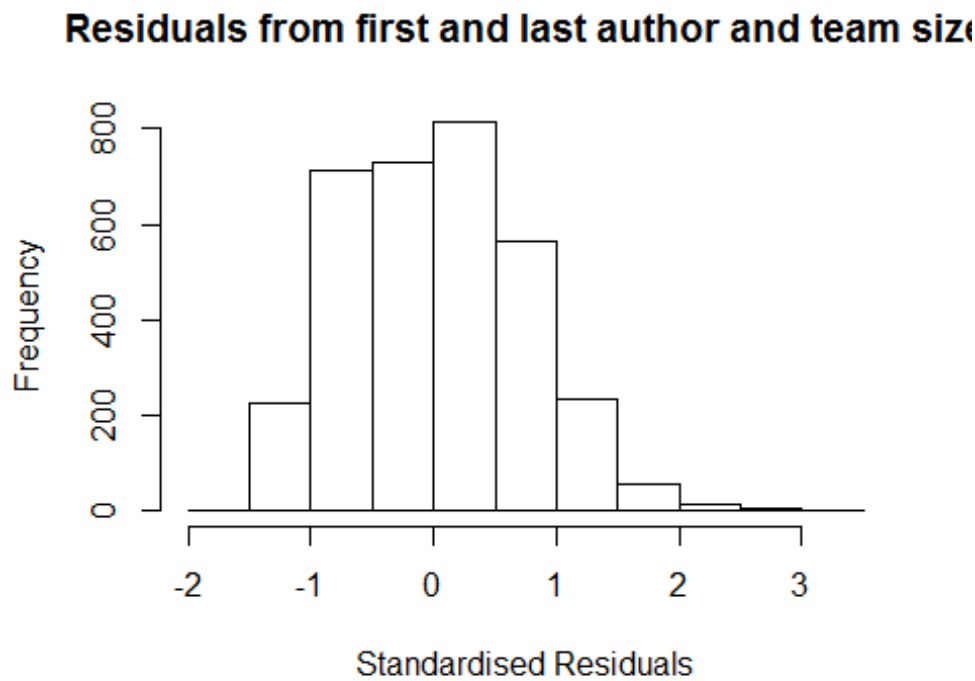
```
##  273  242  197  266  160  156  225  215  256  258  214  257  226  246  262
## 2011 2012
##  254  245
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  255  219  176  233  138  133  194  175  213  214  174  207  185  211  223
## 2011 2012
##  203  202
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
##  Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 79, df = 16, p-value = 2e-10
```




```
##
##  Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 0.56, df = 1, p-value = 0.5
```

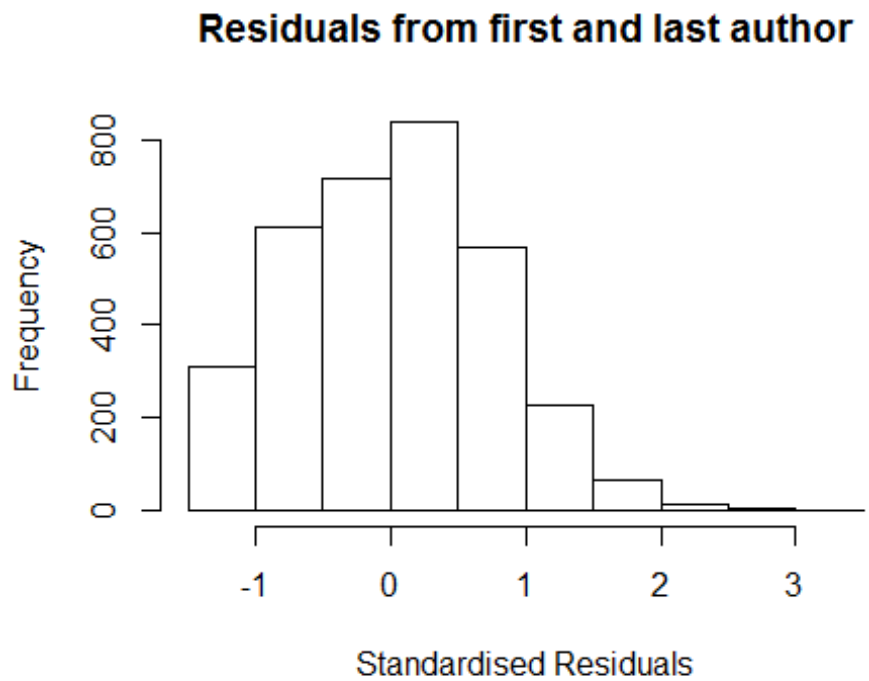



```
## [1] "Female first author team size 2018 geometric mean: 3.81285282199054"
## [1] "Male first author team size 2018 geometric mean: 2.79182538499607"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 10000, p-value = 1e-04
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 4.0610895437424"
## [1] "Male last author team size 2018 geometric mean: 2.86007067049505"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 7600, p-value = 2e-04
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.132  1           1.064
## LastAuthorFemale  1.129  1           1.063
## UniqueAuthors     1.101  4           1.012
## Year              1.131 16           1.004
```



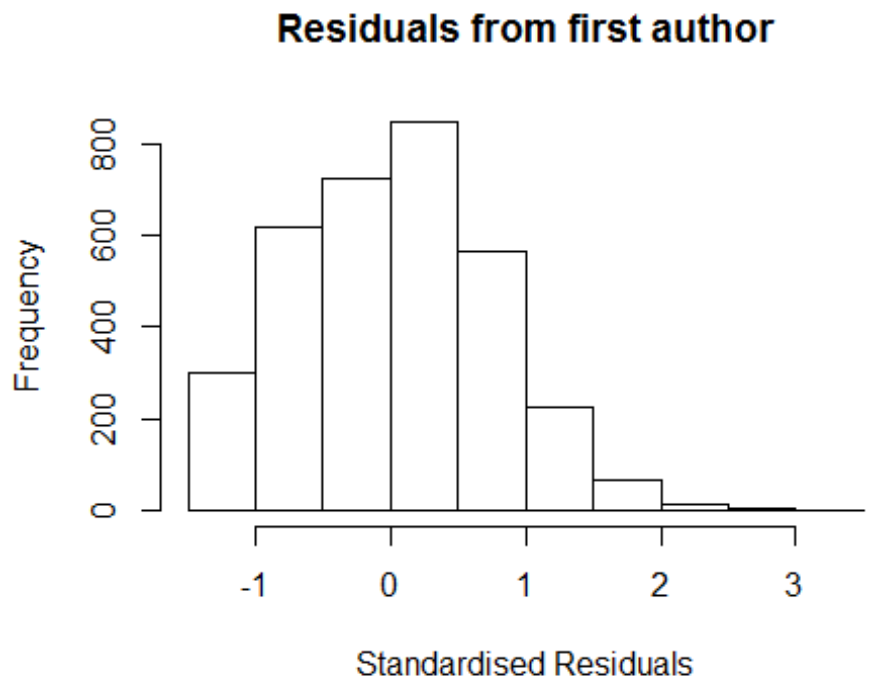

```
## [1] "List of  3 outliers with residuals above  2.5"
##        ScopusId  NLCS Year OneField Fields residuals
## 92   0030291231 3.572 1996     2212      1     2.661
## 202  0030126485 3.949 1996     2208      3     3.428
## 2274 0033907614 3.633 2000     2208      3     2.670
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
UniqueAuthors +
##     Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##     k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.5847 -0.5853  0.0025  0.5176  3.4281
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         0.60035    0.05818   10.32  < 2e-16 ***
## FirstAuthorFemale1 -0.00813    0.04485   -0.18   0.8562
## LastAuthorFemale1  -0.07133    0.04700   -1.52   0.1292
## UniqueAuthors2      0.31109    0.03459    8.99  < 2e-16 ***
## UniqueAuthors3      0.44134    0.03760   11.74  < 2e-16 ***
## UniqueAuthors4      0.45814    0.04647    9.86  < 2e-16 ***
## UniqueAuthors5      0.59783    0.04939   12.10  < 2e-16 ***
## Year1997            0.12510    0.08051    1.55   0.1203
```

```
## Year1998            0.21589    0.08263    2.61   0.0090 **
## Year1999            0.02198    0.07537    0.29   0.7706
## Year2000            0.05193    0.09201    0.56   0.5726
## Year2001            0.39469    0.09374    4.21  2.6e-05 ***
## Year2002            0.31941    0.07721    4.14  3.6e-05 ***
## Year2003            0.15267    0.07522    2.03   0.0425 *
## Year2004            0.21636    0.07324    2.95   0.0032 **
## Year2005            0.11714    0.07449    1.57   0.1159
## Year2006            0.07948    0.07377    1.08   0.2814
## Year2007            0.18282    0.07379    2.48   0.0133 *
## Year2008            0.10928    0.07869    1.39   0.1650
## Year2009            0.19082    0.06938    2.75   0.0060 **
## Year2010            0.20453    0.07292    2.80   0.0051 **
## Year2011            0.06436    0.07229    0.89   0.3734
## Year2012            0.06442    0.07576    0.85   0.3953
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.738
## Multiple R-squared:  0.0906, Adjusted R-squared:  0.0846
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
##  254 weights are ~= 1. The remaining 3101 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0003  0.8780  0.9410  0.9130  0.9840  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          2.98e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.090  1           1.044
## LastAuthorFemale  1.086  1           1.042
## Year              1.035 16           1.001
```



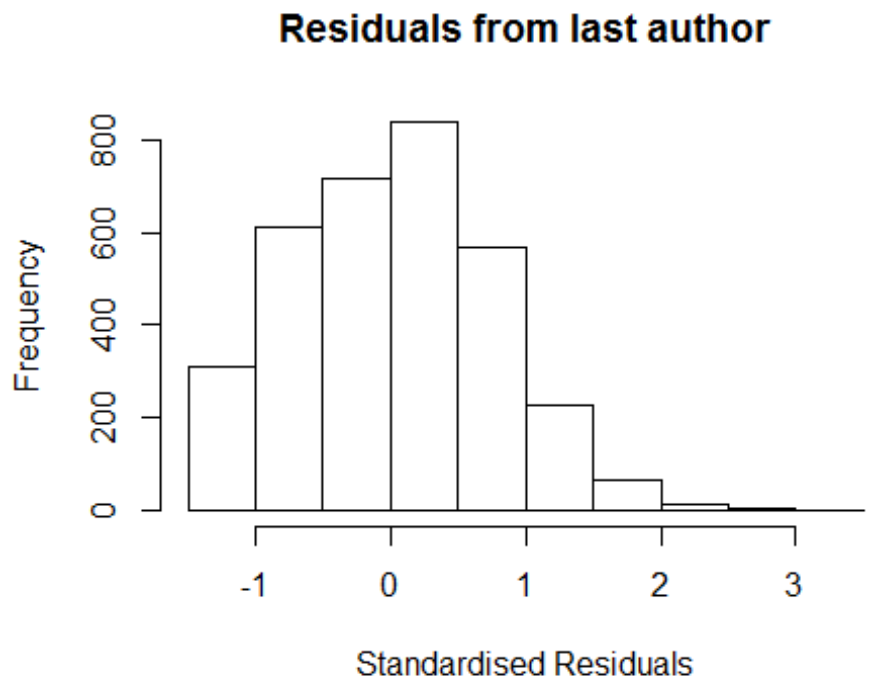

```
## [1] "List of  5 outliers with residuals above  2.5"
##        ScopusId  NLCS Year OneField Fields residuals
## 92   0030291231 3.572 1996     2212      1     2.727
## 202  0030126485 3.949 1996     2208      3     3.128
## 2269 0033882824 3.566 2000     2208      3     2.673
## 2274 0033907614 3.633 2000     2208      3     2.740
## 2459 0035493267 3.845 2001     2208      3     2.556
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##    Min     1Q Median     3Q    Max
## -1.311 -0.570  0.022  0.532  3.128
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)          0.8453     0.0563   15.02  < 2e-16 ***
## FirstAuthorFemale1   0.0218     0.0457    0.48  0.63429
## LastAuthorFemale1   -0.0463     0.0485   -0.96  0.33952
## Year1997             0.1388     0.0824    1.68  0.09225 .
## Year1998             0.2057     0.0851    2.42  0.01567 *
## Year1999             0.0139     0.0764    0.18  0.85573
## Year2000             0.0480     0.0925    0.52  0.60388
## Year2001             0.4440     0.0961    4.62  4.0e-06 ***
```

```
## Year2002             0.3474     0.0826    4.21  2.7e-05 ***
## Year2003             0.1634     0.0786    2.08  0.03770 *
## Year2004             0.2431     0.0753    3.23  0.00125 **
## Year2005             0.1788     0.0754    2.37  0.01783 *
## Year2006             0.1155     0.0757    1.53  0.12725
## Year2007             0.2422     0.0758    3.20  0.00140 **
## Year2008             0.1995     0.0822    2.43  0.01528 *
## Year2009             0.2507     0.0709    3.54  0.00041 ***
## Year2010             0.2983     0.0748    3.99  6.8e-05 ***
## Year2011             0.1591     0.0748    2.13  0.03348 *
## Year2012             0.1324     0.0788    1.68  0.09299 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.8
## Multiple R-squared:  0.02,   Adjusted R-squared:  0.0147
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  269 weights are ~= 1. The remaining 3086 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0916  0.8730  0.9470  0.9200  0.9840  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          2.98e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##                   GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.02  1           1.010
## Year              1.02 16           1.001
```




```
## [1] "List of  5 outliers with residuals above  2.5"
##        ScopusId  NLCS Year OneField Fields residuals
## 92   0030291231 3.572 1996     2212      1     2.727
## 202  0030126485 3.949 1996     2208      3     3.128
## 2269 0033882824 3.566 2000     2208      3     2.673
## 2274 0033907614 3.633 2000     2208      3     2.740
## 2459 0035493267 3.845 2001     2208      3     2.556
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.2953 -0.5670  0.0242  0.5318  3.0978
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         0.84224    0.05629   14.96  < 2e-16 ***
## FirstAuthorFemale1  0.00893    0.04458    0.20   0.8413
## Year1997            0.13972    0.08245    1.69   0.0903 .
## Year1998            0.20435    0.08521    2.40   0.0165 *
## Year1999            0.01452    0.07644    0.19   0.8494
## Year2000            0.04950    0.09257    0.53   0.5929
## Year2001            0.44412    0.09620    4.62  4.0e-06 ***
## Year2002            0.34830    0.08259    4.22  2.5e-05 ***
```

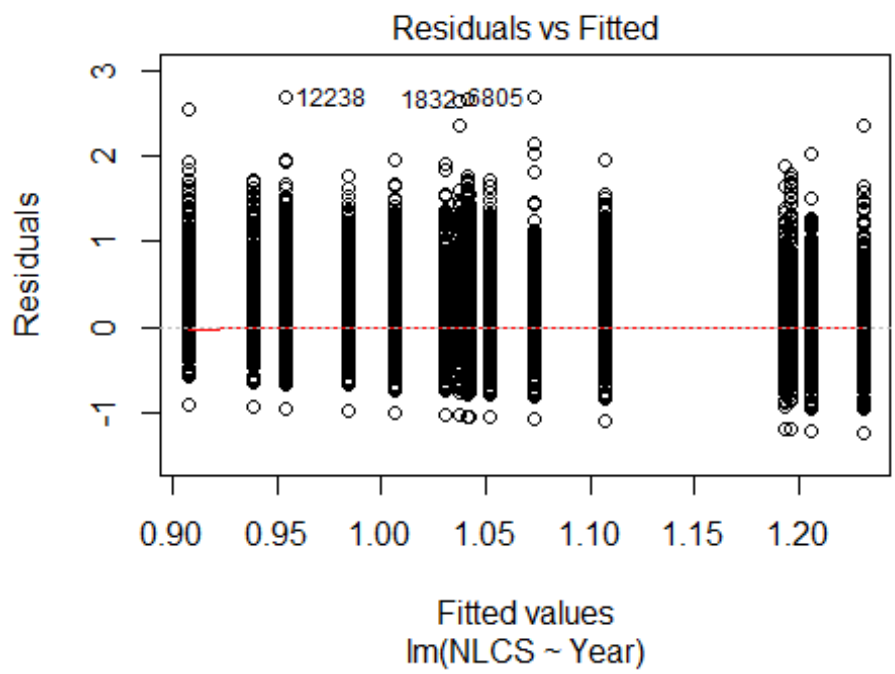
```
## Year2003            0.16376    0.07866    2.08   0.0374 *
## Year2004            0.24304    0.07528    3.23   0.0013 **
## Year2005            0.17777    0.07550    2.35   0.0186 *
## Year2006            0.11675    0.07571    1.54   0.1231
## Year2007            0.23976    0.07581    3.16   0.0016 **
## Year2008            0.20102    0.08232    2.44   0.0147 *
## Year2009            0.25153    0.07094    3.55   0.0004 ***
## Year2010            0.29746    0.07492    3.97  7.3e-05 ***
## Year2011            0.15682    0.07492    2.09   0.0364 *
## Year2012            0.13100    0.07893    1.66   0.0971 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.8
## Multiple R-squared:  0.0197, Adjusted R-squared:  0.0147
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  272 weights are ~= 1. The remaining 3083 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0999  0.8710  0.9470  0.9200  0.9840  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          2.98e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                    psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##                   GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.015  1           1.008
## Year             1.015 16           1.000
```



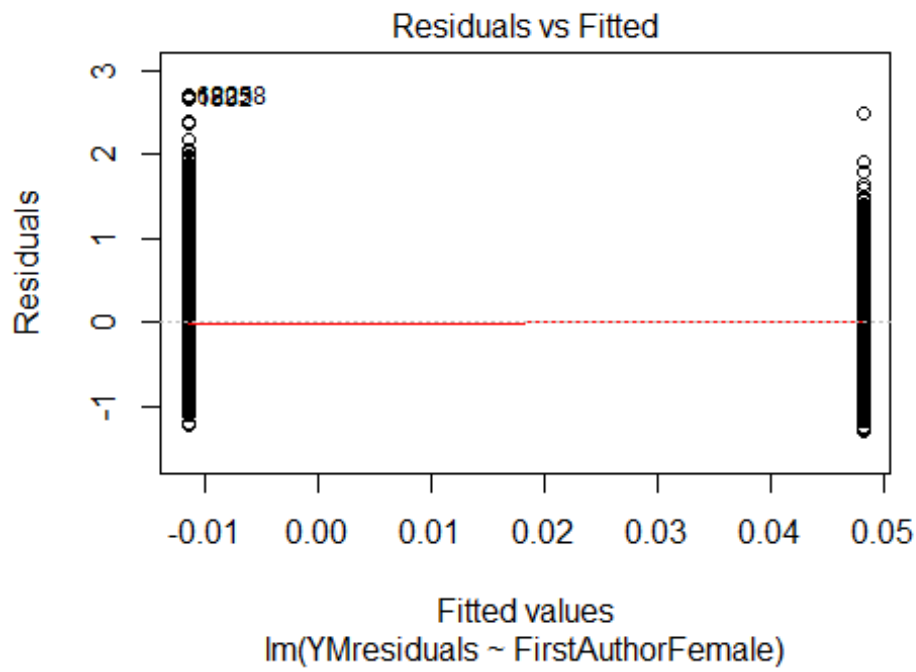

```
## [1] "List of  5 outliers with residuals above  2.5"
##        ScopusId  NLCS Year OneField Fields residuals
## 92   0030291231 3.572 1996     2212      1     2.727
## 202  0030126485 3.949 1996     2208      3     3.128
## 2269 0033882824 3.566 2000     2208      3     2.673
## 2274 0033907614 3.633 2000     2208      3     2.740
## 2459 0035493267 3.845 2001     2208      3     2.556
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.2906 -0.5704  0.0233  0.5341  3.1422
##
## Coefficients:
##                   Estimate Std. Error t value Pr(>|t|)
## (Intercept)         0.8466     0.0563   15.05  < 2e-16 ***
## LastAuthorFemale1  -0.0399     0.0468   -0.85  0.39418
## Year1997            0.1392     0.0827    1.68  0.09261 .
## Year1998            0.2050     0.0851    2.41  0.01601 *
## Year1999            0.0141     0.0763    0.18  0.85377
## Year2000            0.0476     0.0924    0.52  0.60622
## Year2001            0.4439     0.0960    4.62  3.9e-06 ***
## Year2002            0.3470     0.0825    4.21  2.7e-05 ***
```

```
## Year2003            0.1638     0.0786    2.08  0.03723 *
## Year2004            0.2425     0.0753    3.22  0.00129 **
## Year2005            0.1798     0.0754    2.38  0.01719 *
## Year2006            0.1158     0.0757    1.53  0.12610
## Year2007            0.2427     0.0758    3.20  0.00138 **
## Year2008            0.2009     0.0821    2.45  0.01448 *
## Year2009            0.2521     0.0708    3.56  0.00037 ***
## Year2010            0.2988     0.0748    4.00  6.6e-05 ***
## Year2011            0.1592     0.0747    2.13  0.03326 *
## Year2012            0.1337     0.0788    1.70  0.08965 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.8
## Multiple R-squared:  0.0199, Adjusted R-squared:  0.0149
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  274 weights are ~= 1. The remaining 3081 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0884  0.8720  0.9470  0.9200  0.9840  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          2.98e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Sample size for the above analysis:  3355"
## [1] ""
## [1] ""
## [1] "##################################"
## [1] "Analysis of AJSC 2213"
## [1] "##################################"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  630  564  612  599  583  652  536  438  493  597  748  807  860 1092  959
## 2011 2012
## 1028 1048
```

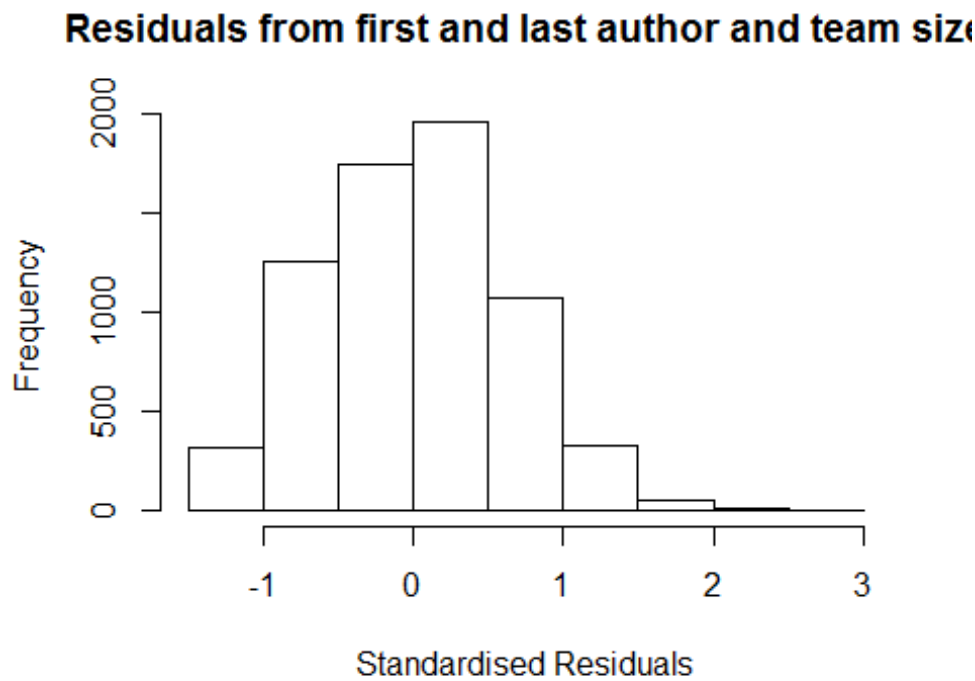
```
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  358  318  349  319  372  366  348  267  310  384  441  524  566  737  646
## 2011 2012
##  699  748
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  316  283  309  285  338  319  293  229  270  338  379  459  489  644  551
## 2011 2012
##  595  636
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
##  Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 28, df = 16, p-value = 0.04
```




```
##
##  Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 3, df = 1, p-value = 0.08
```

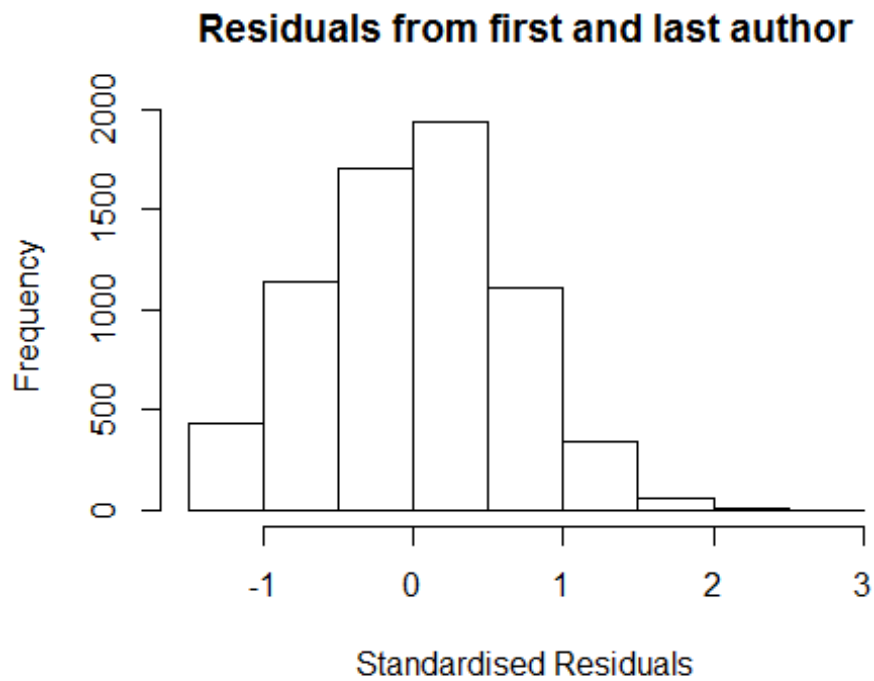



```
## [1] "Female first author team size 2018 geometric mean: 2.86996068944352"
## [1] "Male first author team size 2018 geometric mean: 2.54002690216393"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 40000, p-value = 0.02
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 2.70231417020666"
## [1] "Male last author team size 2018 geometric mean: 2.59334079720806"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 36000, p-value = 0.5
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.185  1           1.089
## LastAuthorFemale  1.168  1           1.081
## UniqueAuthors     1.069  4           1.008
## Year              1.066 16           1.002
```



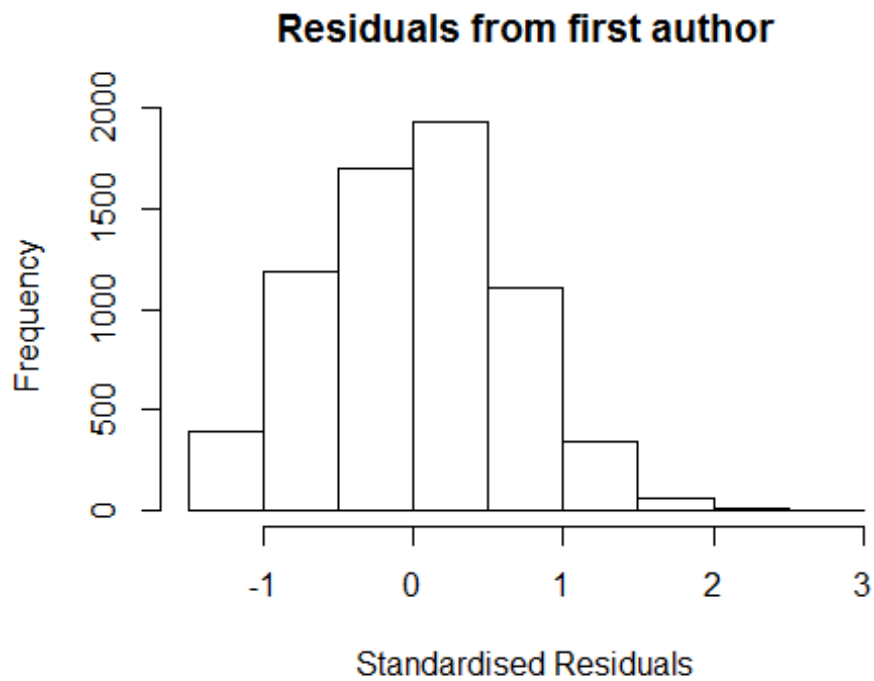

```
## [1] "List of  4 outliers with residuals above  2.5"
##          ScopusId  NLCS Year OneField Fields residuals
## 537    0030370740 3.688 1996     1705      4     2.500
## 1832   0032313923 3.699 1998     1705      4     2.554
## 6805  29144523061 3.760 2006     1700      7     2.695
## 12238 79551514315 3.634 2011     1408      4     2.623
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
UniqueAuthors +
##     Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##     k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.4473 -0.4660  0.0142  0.4360  2.6947
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)           0.8674     0.0424   20.47  < 2e-16 ***
## FirstAuthorFemale1    0.0479     0.0211    2.27  0.02328 *
## LastAuthorFemale1    -0.0209     0.0221   -0.94  0.34498
## UniqueAuthors2        0.2489     0.0210   11.87  < 2e-16 ***
## UniqueAuthors3        0.3205     0.0233   13.77  < 2e-16 ***
## UniqueAuthors4        0.3720     0.0277   13.45  < 2e-16 ***
## UniqueAuthors5        0.3893     0.0278   14.02  < 2e-16 ***
```

```
## Year1997            -0.0183     0.0597   -0.31  0.75956
## Year1998            -0.0427     0.0609   -0.70  0.48297
## Year1999            -0.0296     0.0606   -0.49  0.62566
## Year2000            -0.0750     0.0531   -1.41  0.15797
## Year2001             0.1427     0.0562    2.54  0.01114 *
## Year2002             0.0967     0.0571    1.69  0.09037 .
## Year2003             0.0743     0.0601    1.24  0.21634
## Year2004             0.1425     0.0565    2.52  0.01166 *
## Year2005             0.0197     0.0532    0.37  0.71111
## Year2006            -0.0509     0.0531   -0.96  0.33729
## Year2007            -0.0838     0.0508   -1.65  0.09942 .
## Year2008            -0.0635     0.0500   -1.27  0.20435
## Year2009            -0.1659     0.0487   -3.41  0.00067 ***
## Year2010            -0.1315     0.0489   -2.69  0.00721 **
## Year2011            -0.1770     0.0493   -3.59  0.00033 ***
## Year2012            -0.2266     0.0489   -4.64  3.6e-06 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.654
## Multiple R-squared:  0.079,  Adjusted R-squared:  0.076
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  543 weights are ~= 1. The remaining 6190 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0512  0.8760  0.9500  0.9150  0.9860  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.49e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.135  1           1.065
## LastAuthorFemale  1.127  1           1.062
## Year              1.018 16           1.001
```



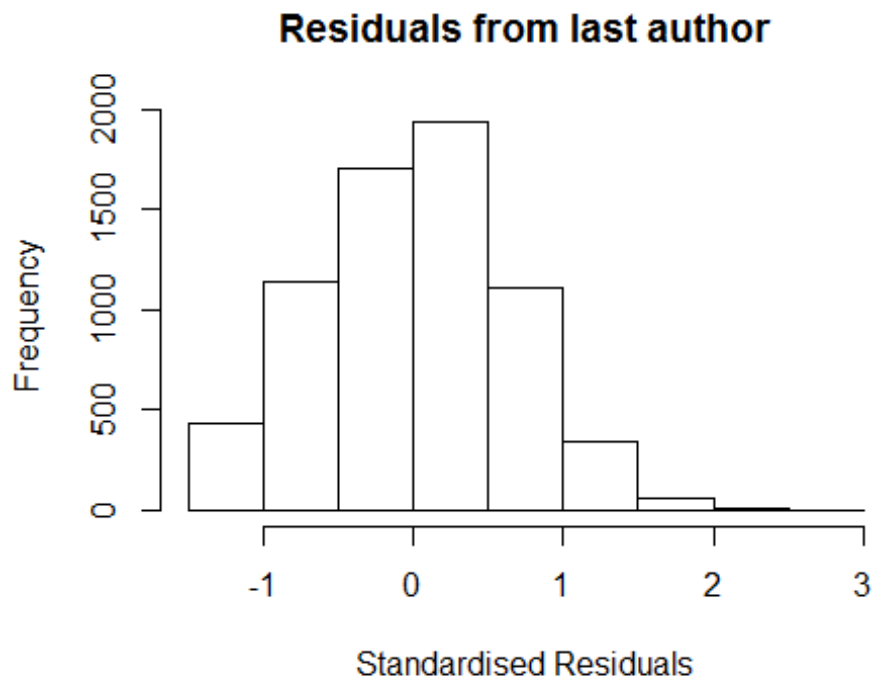

```
## [1] "List of  4 outliers with residuals above  2.5"
##          ScopusId  NLCS Year OneField Fields residuals
## 537    0030370740 3.688 1996     1705      4     2.667
## 1832   0032313923 3.699 1998     1705      4     2.696
## 6805  29144523061 3.760 2006     1700      7     2.741
## 12238 79551514315 3.634 2011     1408      4     2.746
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.2942 -0.4719  0.0218  0.4495  2.7463
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.020532   0.041814   24.41  < 2e-16 ***
## FirstAuthorFemale1  0.081815   0.021666    3.78  0.00016 ***
## LastAuthorFemale1  -0.006697   0.023031   -0.29  0.77124
## Year1997            0.018640   0.059956    0.31  0.75589
## Year1998           -0.017924   0.060909   -0.29  0.76856
## Year1999           -0.000797   0.060596   -0.01  0.98951
## Year2000           -0.035615   0.053868   -0.66  0.50854
## Year2001            0.191877   0.056091    3.42  0.00063 ***
## Year2002            0.147489   0.057436    2.57  0.01025 *
```

```
## Year2003            0.122990   0.059933    2.05  0.04020 *
## Year2004            0.189197   0.057691    3.28  0.00105 **
## Year2005            0.054798   0.055159    0.99  0.32053
## Year2006           -0.001040   0.054129   -0.02  0.98467
## Year2007           -0.045216   0.051512   -0.88  0.38010
## Year2008           -0.019883   0.050737   -0.39  0.69515
## Year2009           -0.120123   0.049399   -2.43  0.01505 *
## Year2010           -0.081439   0.050133   -1.62  0.10433
## Year2011           -0.132832   0.050328   -2.64  0.00833 **
## Year2012           -0.184606   0.049555   -3.73  0.00020 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.675
## Multiple R-squared:  0.027,  Adjusted R-squared:  0.0244
## Convergence in 10 IRWLS iterations
##
## Robustness weights:
##  560 weights are ~= 1. The remaining 6173 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0607  0.8650  0.9500  0.9160  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.49e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                    psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##                     GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.012  1           1.006
## Year              1.012 16           1.000
```




```
## [1] "List of  4 outliers with residuals above  2.5"
##          ScopusId  NLCS Year OneField Fields residuals
## 537    0030370740 3.688 1996     1705      4     2.667
## 1832   0032313923 3.699 1998     1705      4     2.696
## 6805  29144523061 3.760 2006     1700      7     2.741
## 12238 79551514315 3.634 2011     1408      4     2.746
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.2911 -0.4718  0.0225  0.4496  2.7470
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.02019    0.04180   24.41  < 2e-16 ***
## FirstAuthorFemale1  0.07946    0.02051    3.87  0.00011 ***
## Year1997            0.01828    0.05996    0.30  0.76049
## Year1998           -0.01835    0.06087   -0.30  0.76308
## Year1999           -0.00106    0.06058   -0.02  0.98606
## Year2000           -0.03585    0.05386   -0.67  0.50566
## Year2001            0.19145    0.05607    3.41  0.00064 ***
## Year2002            0.14707    0.05743    2.56  0.01046 *
## Year2003            0.12243    0.05991    2.04  0.04106 *
```

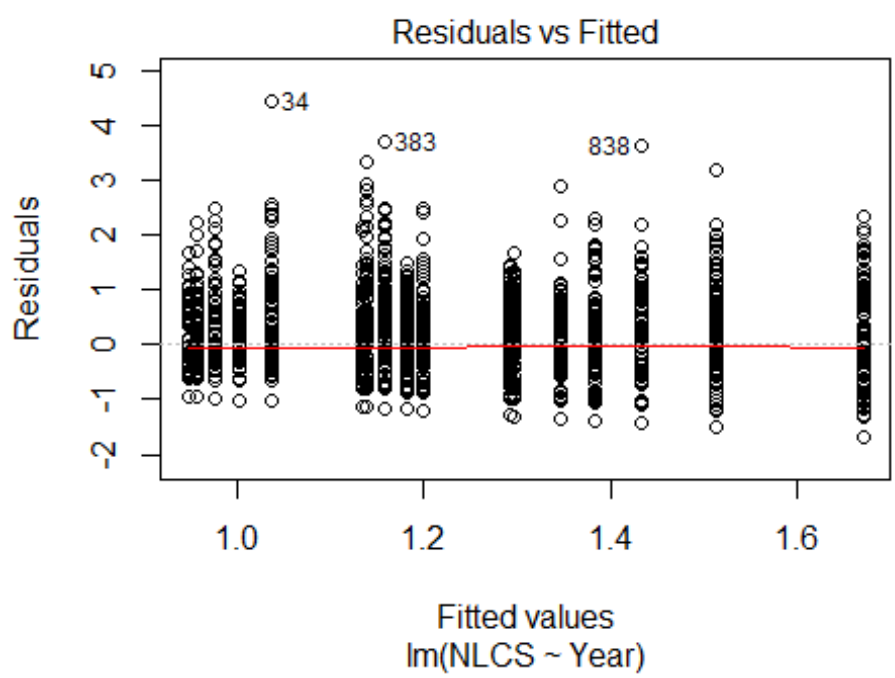
```
## Year2004            0.18898    0.05769    3.28  0.00106 **
## Year2005            0.05432    0.05517    0.98  0.32490
## Year2006           -0.00139    0.05412   -0.03  0.97953
## Year2007           -0.04579    0.05148   -0.89  0.37376
## Year2008           -0.02031    0.05071   -0.40  0.68885
## Year2009           -0.12030    0.04940   -2.44  0.01492 *
## Year2010           -0.08173    0.05013   -1.63  0.10305
## Year2011           -0.13322    0.05032   -2.65  0.00813 **
## Year2012           -0.18503    0.04954   -3.73  0.00019 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.675
## Multiple R-squared:  0.027,  Adjusted R-squared:  0.0245
## Convergence in 10 IRWLS iterations
##
## Robustness weights:
##  559 weights are ~= 1. The remaining 6174 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0605  0.8650  0.9500  0.9160  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.49e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##                   GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.005  1           1.003
## Year             1.005 16           1.000
```




```
## [1] "List of  4 outliers with residuals above  2.5"
##          ScopusId  NLCS Year OneField Fields residuals
## 537    0030370740 3.688 1996     1705      4     2.667
## 1832   0032313923 3.699 1998     1705      4     2.696
## 6805  29144523061 3.760 2006     1700      7     2.741
## 12238 79551514315 3.634 2011     1408      4     2.746
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.2460 -0.4723  0.0223  0.4517  2.7355
##
## Coefficients:
##                   Estimate Std. Error t value Pr(>|t|)
## (Intercept)        1.02875    0.04197   24.51  < 2e-16 ***
## LastAuthorFemale1  0.02489    0.02155    1.15  0.24819
## Year1997           0.01525    0.06017    0.25  0.79990
## Year1998          -0.02274    0.06105   -0.37  0.70953
## Year1999          -0.00218    0.06070   -0.04  0.97135
## Year2000          -0.03397    0.05407   -0.63  0.52984
## Year2001           0.19236    0.05628    3.42  0.00064 ***
## Year2002           0.14913    0.05746    2.60  0.00947 **
## Year2003           0.12672    0.06006    2.11  0.03490 *
```

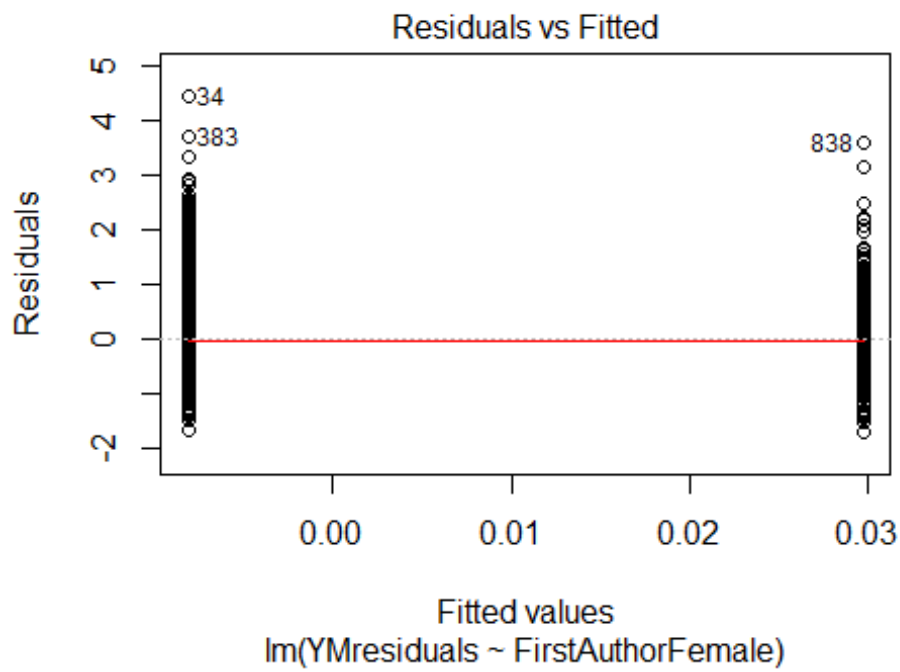
```
## Year2004           0.19208    0.05785    3.32  0.00090 ***
## Year2005           0.05757    0.05515    1.04  0.29660
## Year2006           0.00202    0.05430    0.04  0.97034
## Year2007          -0.04337    0.05164   -0.84  0.40102
## Year2008          -0.01675    0.05086   -0.33  0.74185
## Year2009          -0.11503    0.04955   -2.32  0.02030 *
## Year2010          -0.07667    0.05037   -1.52  0.12803
## Year2011          -0.13030    0.05052   -2.58  0.00992 **
## Year2012          -0.18001    0.04972   -3.62  0.00030 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.677
## Multiple R-squared:  0.025,  Adjusted R-squared:  0.0225
## Convergence in 10 IRWLS iterations
##
## Robustness weights:
##  577 weights are ~= 1. The remaining 6156 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0653  0.8620  0.9510  0.9150  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.49e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Sample size for the above analysis:  6733"
## [1] ""
## [1] ""
## [1] "#####################################"
## [1] "Analysis of AJSC 2214"
## [1] "#####################################"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  202  222  190  200  207  194  183  140  137  197  256  262  252  187  133
## 2011 2012
##  130  127
##
```

```
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  145  138  115   83  128   98   90   84   77   98  136  146  144  105   71
## 2011 2012
##   82   74
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  131  117  104   69   97   80   73   71   59   79  116  117  113   91   53
## 2011 2012
##   68   61
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
##  Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 100, df = 16, p-value = 4e-15
```



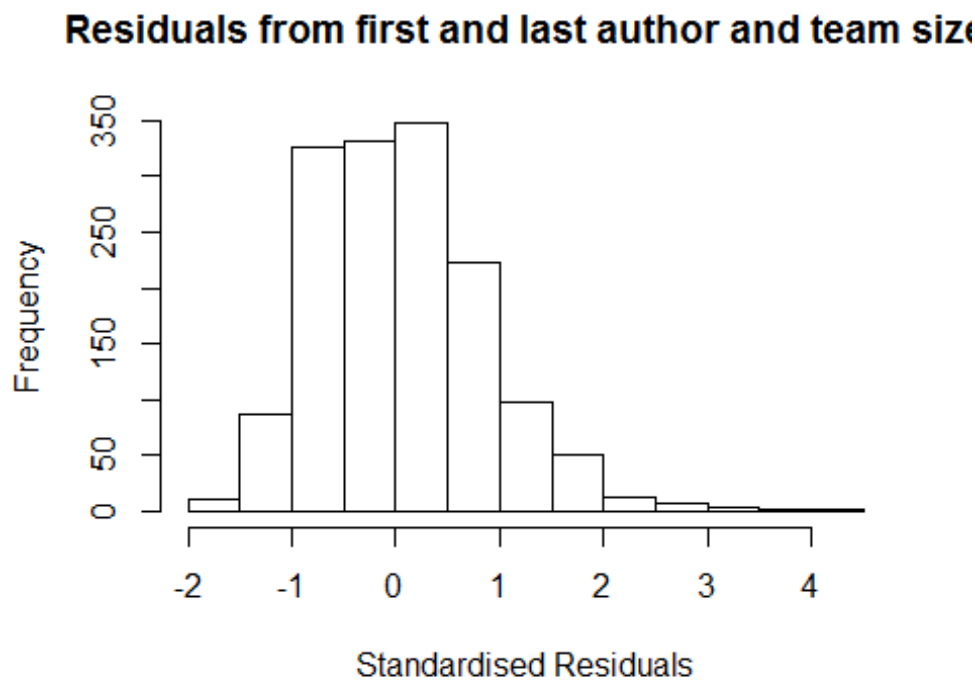

```
##
##  Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 3, df = 1, p-value = 0.08

## [1] "Female first author team size 2018 geometric mean: 1.84756137378954"
## [1] "Male first author team size 2018 geometric mean: 1.74961845240595"
```

```
## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties

##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 690, p-value = 0.7
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 1.76500726747574"
## [1] "Male last author team size 2018 geometric mean: 1.80088484189588"

## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties
```




```
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 600, p-value = 0.8
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##                     GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.497  1           1.223
```

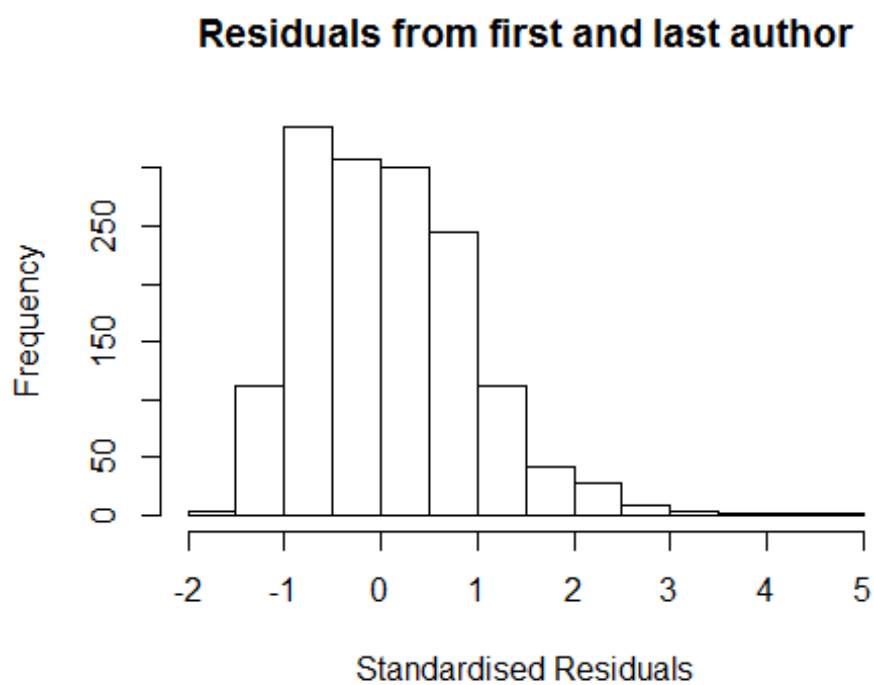
```
## LastAuthorFemale  1.480  1           1.216
## UniqueAuthors     1.296  4           1.033
## Year              1.482 16           1.012
```




```
## [1] "List of  12 outliers with residuals above  2.5"
##        ScopusId  NLCS Year OneField Fields residuals
## 2    0030403552 3.421 1996     2214      1     2.897
## 34   0030173121 5.461 1996     2208      2     4.233
## 106  0030213167 3.312 1996     2214      1     2.788
## 145  0030082306 3.589 1996     2214      1     3.065
## 383  0031075621 4.855 1997     2214      1     3.456
## 505  0032136715 3.940 1998     2208      2     2.632
## 598  0032099560 4.067 1998     2214      1     3.463
## 619  0031998522 4.476 1998     2214      1     3.872
## 960  0034291601 4.678 2000     2214      1     2.940
## 1245 0035364996 3.796 2001     2208      2     2.756
## 1527 0036476080 3.674 2002     2208      2     2.894
## 1915 0742290133 4.216 2004     2208      2     2.835
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
UniqueAuthors +
##     Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##     k.max = 1000))
##  \--> method = "MM"
## Residuals:
```

```
##      Min       1Q   Median       3Q      Max
## -1.80173 -0.52473 -0.00787  0.53185  4.23321
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)          0.5236     0.0834    6.28  4.6e-10 ***
## FirstAuthorFemale1   0.1387     0.0569    2.44   0.0149 *
## LastAuthorFemale1   -0.0373     0.0547   -0.68   0.4955
## UniqueAuthors2       0.7042     0.0527   13.37  < 2e-16 ***
## UniqueAuthors3       0.7617     0.0620   12.28  < 2e-16 ***
## UniqueAuthors4       0.8599     0.0827   10.40  < 2e-16 ***
## UniqueAuthors5       0.7466     0.1124    6.64  4.4e-11 ***
## Year1997             0.0152     0.1105    0.14   0.8906
## Year1998             0.0800     0.1249    0.64   0.5222
## Year1999             0.4187     0.1366    3.06   0.0022 **
## Year2000             0.3512     0.1266    2.78   0.0056 **
## Year2001             0.5165     0.1388    3.72   0.0002 ***
## Year2002             0.2561     0.1271    2.02   0.0440 *
## Year2003            -0.0347     0.1233   -0.28   0.7783
## Year2004             0.0957     0.1176    0.81   0.4162
## Year2005             0.2874     0.1169    2.46   0.0141 *
## Year2006             0.1707     0.1088    1.57   0.1167
## Year2007             0.2116     0.1036    2.04   0.0414 *
## Year2008             0.0945     0.1015    0.93   0.3519
## Year2009            -0.1117     0.1074   -1.04   0.2985
## Year2010             0.0452     0.1232    0.37   0.7138
## Year2011             0.1003     0.1116    0.90   0.3688
## Year2012            -0.0495     0.1201   -0.41   0.6802
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.723
## Multiple R-squared:  0.241,  Adjusted R-squared:  0.23
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
##  4 observations c(18,201,296,305)
##   are outliers with |weight| = 0 ( < 6.7e-05);
##  112 weights are ~= 1. The remaining 1383 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0329  0.8760  0.9470  0.8980  0.9810  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          6.67e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
```

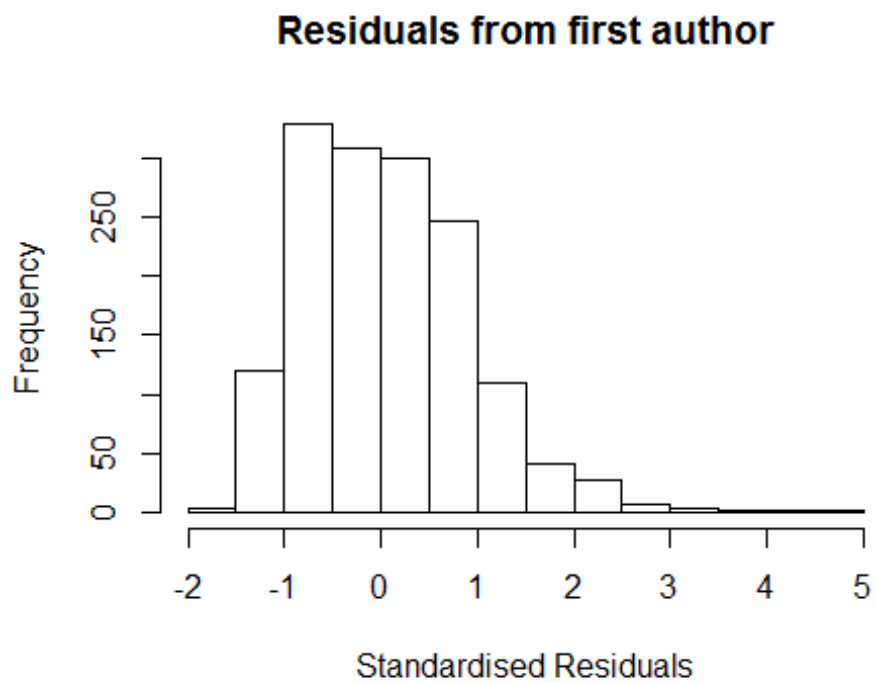
```
##   trace.lev         mts  compute.rd
##           0        1000           0
##                    psi           subsampling                    cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                   "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##                     GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.470  1           1.212
## LastAuthorFemale  1.428  1           1.195
## Year              1.136 16           1.004
```




```
## [1] "List of  14 outliers with residuals above  2.5"
##        ScopusId  NLCS Year OneField Fields residuals
## 2    0030403552 3.421 1996     2214      1     2.567
## 34   0030173121 5.461 1996     2208      2     4.607
## 145  0030082306 3.589 1996     2214      1     2.735
## 147  0030083686 3.376 1996     2214      1     2.522
## 383  0031075621 4.855 1997     2214      1     4.005
## 486  0032314844 3.756 1998     2214      1     2.810
## 505  0032136715 3.940 1998     2208      2     2.994
## 566  0032178780 3.469 1998     2214      1     2.523
## 597  0032099120 3.674 1998     2214      1     2.728
## 598  0032099560 4.067 1998     2214      1     3.121
## 619  0031998522 4.476 1998     2214      1     3.530
```

```
## 960  0034291601 4.678 2000     2214      1     3.269
## 1714 0037282477 3.441 2003     2208      2     2.606
## 1915 0742290133 4.216 2004     2208      2     3.010
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.5122 -0.6796 -0.0151  0.6013  4.6069
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         0.85409    0.08940    9.55  < 2e-16 ***
## FirstAuthorFemale1  0.11807    0.06149    1.92  0.05500 .
## LastAuthorFemale1  -0.09134    0.06005   -1.52  0.12844
## Year1997           -0.00438    0.12751   -0.03  0.97258
## Year1998            0.09188    0.14346    0.64  0.52200
## Year1999            0.44784    0.15615    2.87  0.00419 **
## Year2000            0.52798    0.13225    3.99  6.9e-05 ***
## Year2001            0.63137    0.15346    4.11  4.1e-05 ***
## Year2002            0.35852    0.14493    2.47  0.01348 *
## Year2003           -0.01945    0.13227   -0.15  0.88309
## Year2004            0.35187    0.13421    2.62  0.00883 **
## Year2005            0.39351    0.13498    2.92  0.00361 **
## Year2006            0.30774    0.11105    2.77  0.00566 **
## Year2007            0.40753    0.11558    3.53  0.00043 ***
## Year2008            0.22264    0.11499    1.94  0.05305 .
## Year2009            0.00296    0.11567    0.03  0.97958
## Year2010            0.13879    0.12693    1.09  0.27438
## Year2011            0.16632    0.12013    1.38  0.16639
## Year2012           -0.03713    0.12924   -0.29  0.77391
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.861
## Multiple R-squared:  0.056,  Adjusted R-squared:  0.0445
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
##  observation 18 is an outlier with |weight| = 0 ( < 6.7e-05);
##  97 weights are ~= 1. The remaining 1401 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0002  0.8910  0.9400  0.9100  0.9830  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
```

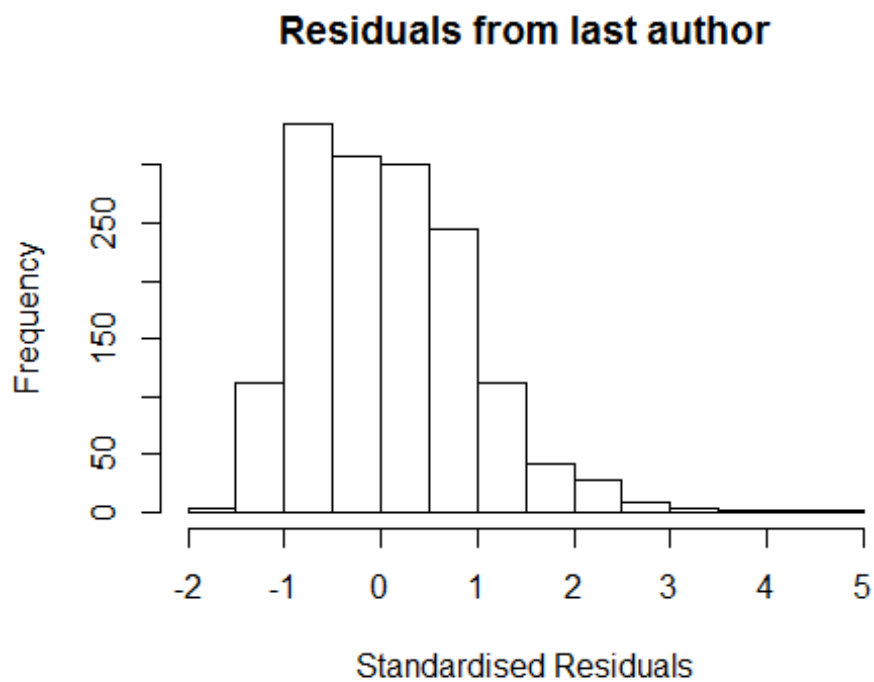
```
##          1.00e-07          1.00e-07          6.67e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.089  1           1.044
## Year              1.089 16           1.003
```




```
## [1] "List of  14 outliers with residuals above  2.5"
##        ScopusId  NLCS Year OneField Fields residuals
## 2    0030403552 3.421 1996     2214      1     2.567
## 34   0030173121 5.461 1996     2208      2     4.607
## 145  0030082306 3.589 1996     2214      1     2.735
## 147  0030083686 3.376 1996     2214      1     2.522
## 383  0031075621 4.855 1997     2214      1     4.005
## 486  0032314844 3.756 1998     2214      1     2.810
## 505  0032136715 3.940 1998     2208      2     2.994
## 566  0032178780 3.469 1998     2214      1     2.523
```

```
## 597  0032099120 3.674 1998     2214      1     2.728
## 598  0032099560 4.067 1998     2214      1     3.121
## 619  0031998522 4.476 1998     2214      1     3.530
## 960  0034291601 4.678 2000     2214      1     3.269
## 1714 0037282477 3.441 2003     2208      2     2.606
## 1915 0742290133 4.216 2004     2208      2     3.010
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.5369 -0.6940 -0.0205  0.6164  4.6154
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         0.84564    0.08890    9.51  < 2e-16 ***
## FirstAuthorFemale1  0.06693    0.05369    1.25  0.21276
## Year1997           -0.00271    0.12740   -0.02  0.98304
## Year1998            0.09333    0.14327    0.65  0.51488
## Year1999            0.45133    0.15629    2.89  0.00394 **
## Year2000            0.53073    0.13208    4.02  6.2e-05 ***
## Year2001            0.62435    0.15361    4.06  5.1e-05 ***
## Year2002            0.36392    0.14466    2.52  0.01199 *
## Year2003           -0.01711    0.13281   -0.13  0.89754
## Year2004            0.35724    0.13451    2.66  0.00800 **
## Year2005            0.39813    0.13543    2.94  0.00334 **
## Year2006            0.30682    0.11121    2.76  0.00587 **
## Year2007            0.40407    0.11510    3.51  0.00046 ***
## Year2008            0.22333    0.11467    1.95  0.05166 .
## Year2009           -0.00785    0.11570   -0.07  0.94589
## Year2010            0.14070    0.12766    1.10  0.27060
## Year2011            0.16236    0.12009    1.35  0.17657
## Year2012           -0.04223    0.12853   -0.33  0.74251
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.86
## Multiple R-squared:  0.0549, Adjusted R-squared:  0.044
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
##  observation 18 is an outlier with |weight| = 0 ( < 6.7e-05);
##  102 weights are ~= 1. The remaining 1396 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0001  0.8890  0.9410  0.9090  0.9830  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
```

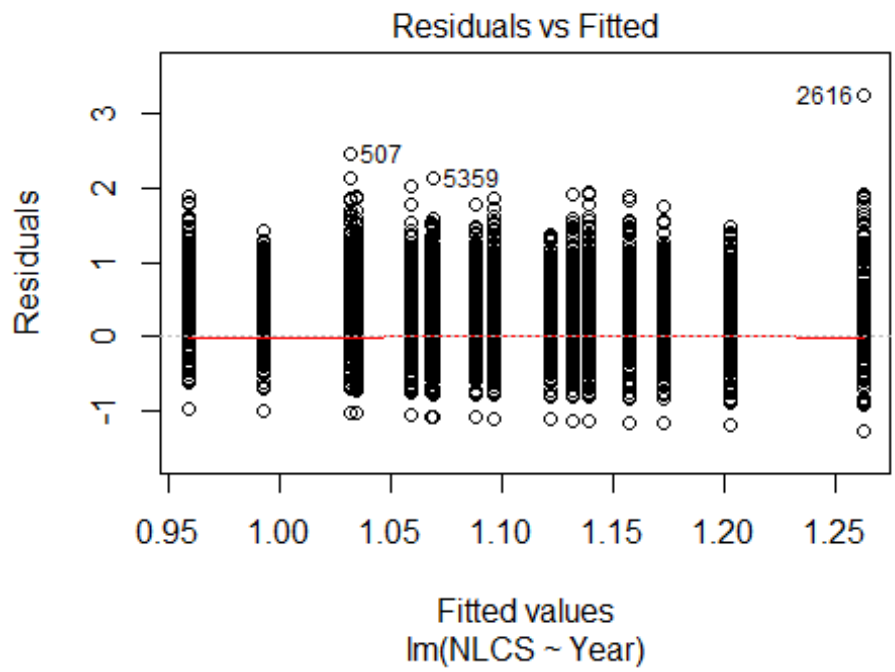
```
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          6.67e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##                   GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.064  1           1.031
## Year             1.064 16           1.002
```



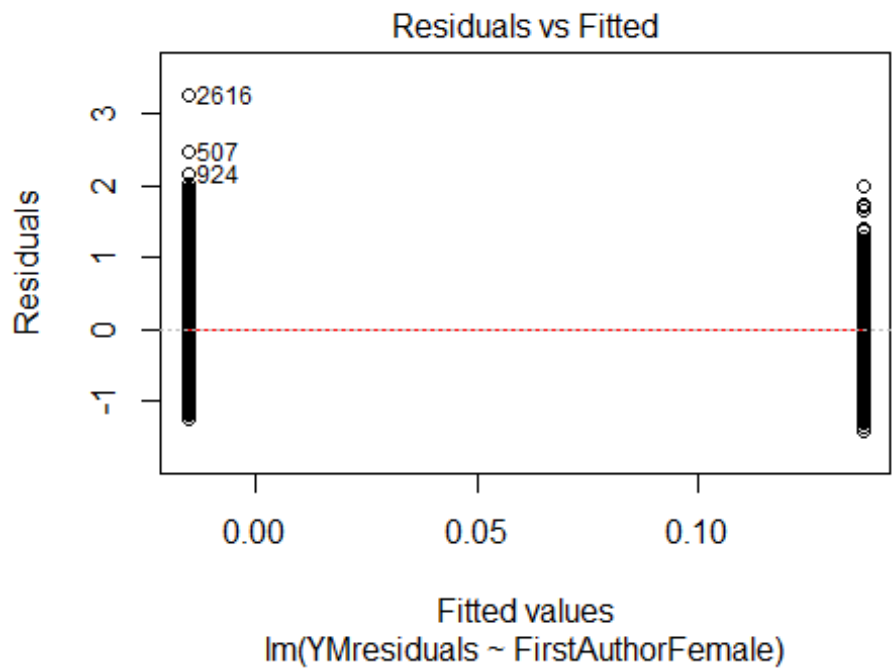

```
## [1] "List of  14 outliers with residuals above  2.5"
##        ScopusId  NLCS Year OneField Fields residuals
## 2    0030403552 3.421 1996     2214      1     2.567
## 34   0030173121 5.461 1996     2208      2     4.607
## 145  0030082306 3.589 1996     2214      1     2.735
## 147  0030083686 3.376 1996     2214      1     2.522
## 383  0031075621 4.855 1997     2214      1     4.005
## 486  0032314844 3.756 1998     2214      1     2.810
```

```
## 505  0032136715 3.940 1998     2208      2     2.994
## 566  0032178780 3.469 1998     2214      1     2.523
## 597  0032099120 3.674 1998     2214      1     2.728
## 598  0032099560 4.067 1998     2214      1     3.121
## 619  0031998522 4.476 1998     2214      1     3.530
## 960  0034291601 4.678 2000     2214      1     3.269
## 1714 0037282477 3.441 2003     2208      2     2.606
## 1915 0742290133 4.216 2004     2208      2     3.010
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.4910 -0.6842 -0.0174  0.6235  4.6013
##
## Coefficients:
##                   Estimate Std. Error t value Pr(>|t|)
## (Intercept)        0.85971    0.08980    9.57  < 2e-16 ***
## LastAuthorFemale1 -0.02388    0.05118   -0.47  0.64087
## Year1997          -0.00204    0.12732   -0.02  0.98724
## Year1998           0.10103    0.14378    0.70  0.48237
## Year1999           0.45194    0.15677    2.88  0.00400 **
## Year2000           0.52812    0.13306    3.97  7.6e-05 ***
## Year2001           0.63131    0.15368    4.11  4.2e-05 ***
## Year2002           0.36878    0.14547    2.54  0.01134 *
## Year2003          -0.01406    0.13422   -0.10  0.91657
## Year2004           0.36424    0.13515    2.70  0.00712 **
## Year2005           0.39985    0.13492    2.96  0.00309 **
## Year2006           0.31566    0.11136    2.83  0.00465 **
## Year2007           0.40745    0.11549    3.53  0.00043 ***
## Year2008           0.23393    0.11491    2.04  0.04195 *
## Year2009           0.00414    0.11617    0.04  0.97156
## Year2010           0.15631    0.12821    1.22  0.22298
## Year2011           0.17849    0.12001    1.49  0.13717
## Year2012          -0.02040    0.12826   -0.16  0.87365
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.86
## Multiple R-squared:  0.0539, Adjusted R-squared:  0.043
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
##  observation 18 is an outlier with |weight| = 0 ( < 6.7e-05);
##  99 weights are ~= 1. The remaining 1399 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0003  0.8900  0.9400  0.9090  0.9830  0.9990
```

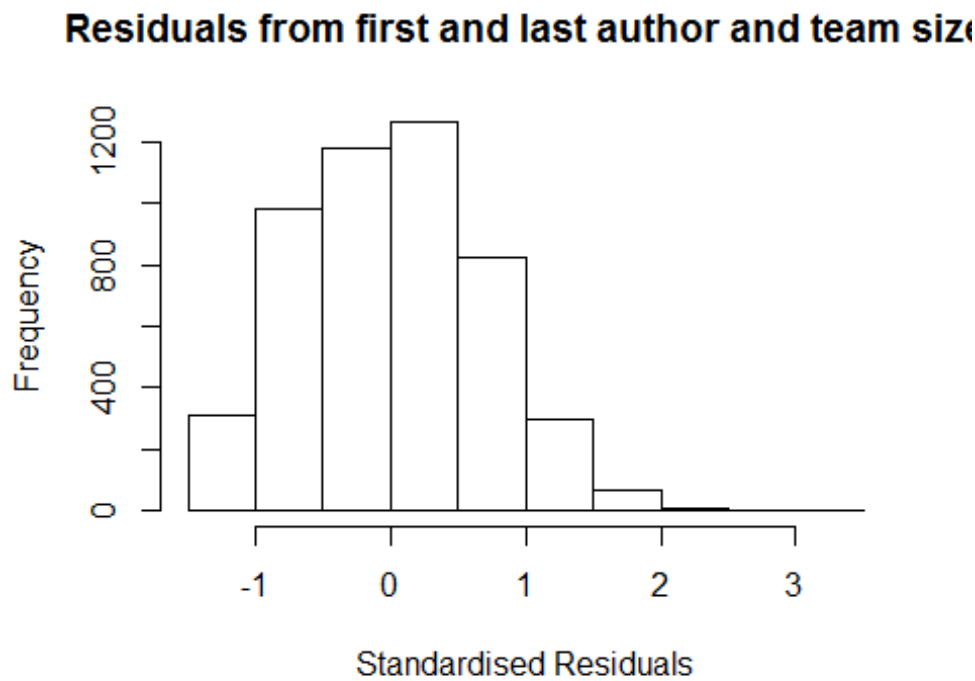
```
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          6.67e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Sample size for the above analysis:  1499"
## [1] ""
## [1] ""
## [1] "####################################"
## [1] "Analysis of AJSC 2215"
## [1] "####################################"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  450  461  461  457  451  533  466  448  475  527  595  642  593  653  793
## 2011 2012
##  844  824
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  279  304  304  274  269  208  274  279  315  326  373  411  392  386  471
## 2011 2012
##  514  497
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  249  261  266  239  219  173  231  228  259  277  319  361  326  308  388
## 2011 2012
##  421  424
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
##  Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 110, df = 16, p-value = 2e-15
```




```
##
##  Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 2, df = 1, p-value = 0.2
```

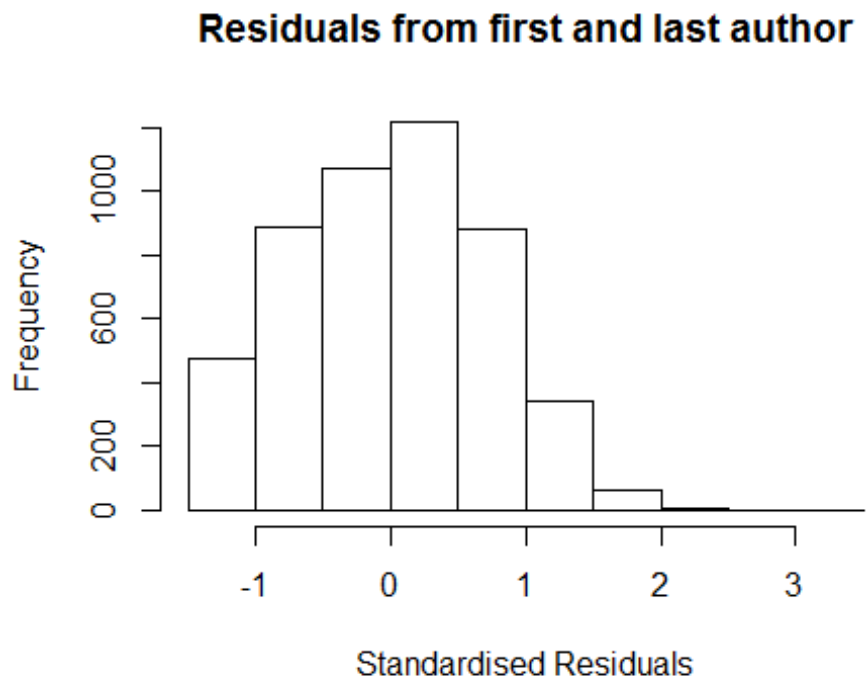



```
## [1] "Female first author team size 2018 geometric mean: 2.97491373841808"
## [1] "Male first author team size 2018 geometric mean: 2.58012095192425"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 27000, p-value = 0.07
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 2.76655945086227"
## [1] "Male last author team size 2018 geometric mean: 2.61803928975874"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 22000, p-value = 0.8
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.054  1           1.027
## LastAuthorFemale  1.054  1           1.027
## UniqueAuthors     1.128  4           1.015
## Year              1.108 16           1.003
```



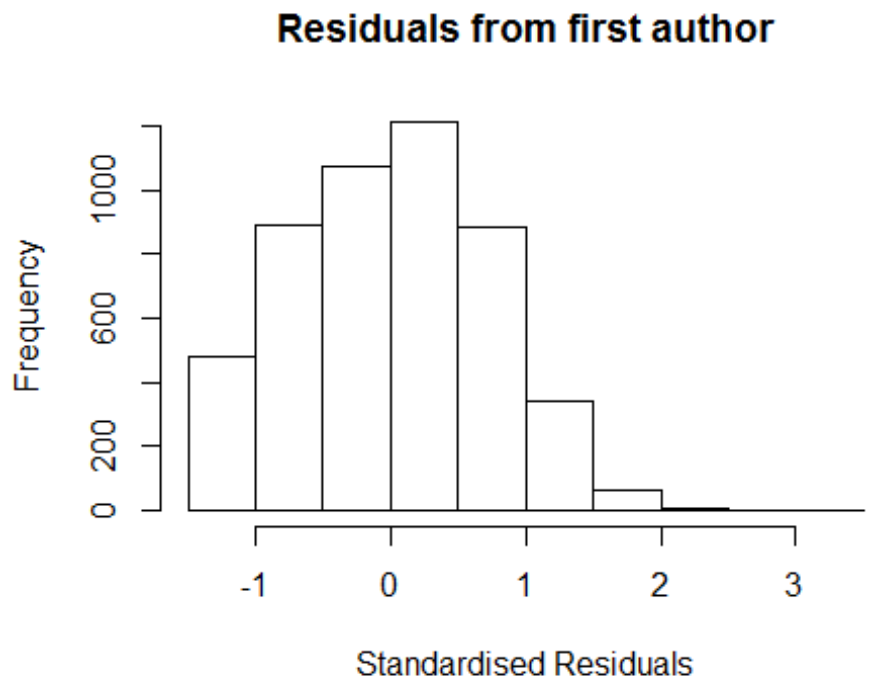

```
## [1] "List of  2 outliers with residuals above  2.5"
##        ScopusId  NLCS Year OneField Fields residuals
## 507  0031097861 3.492 1997     2205      4     2.838
## 2616 0035492228 4.506 2001     1909      2     3.153
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
UniqueAuthors +
##     Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##     k.max = 1000))
##  \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.43009 -0.54621 -0.00163  0.48438  3.15327
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         0.63474    0.04870   13.03   <2e-16 ***
## FirstAuthorFemale1  0.09921    0.03531    2.81   0.0050 **
## LastAuthorFemale1  -0.07376    0.03610   -2.04   0.0411 *
## UniqueAuthors2      0.49486    0.02732   18.12   <2e-16 ***
## UniqueAuthors3      0.52872    0.02881   18.35   <2e-16 ***
## UniqueAuthors4      0.55445    0.03750   14.79   <2e-16 ***
## UniqueAuthors5      0.49580    0.04653   10.65   <2e-16 ***
## Year1997            0.01889    0.07050    0.27   0.7888
## Year1998            0.11395    0.06963    1.64   0.1018
```

```
## Year1999            0.20128    0.06979    2.88   0.0039 **
## Year2000            0.14057    0.07198    1.95   0.0509 .
## Year2001            0.16355    0.08746    1.87   0.0616 .
## Year2002            0.08716    0.07062    1.23   0.2172
## Year2003            0.08450    0.06961    1.21   0.2248
## Year2004            0.04259    0.06343    0.67   0.5019
## Year2005            0.02750    0.06269    0.44   0.6609
## Year2006            0.03486    0.06026    0.58   0.5629
## Year2007           -0.02427    0.05842   -0.42   0.6779
## Year2008           -0.03438    0.05940   -0.58   0.5628
## Year2009            0.03455    0.06052    0.57   0.5681
## Year2010            0.00889    0.05966    0.15   0.8815
## Year2011           -0.02630    0.05784   -0.45   0.6493
## Year2012           -0.01503    0.05729   -0.26   0.7930
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.683
## Multiple R-squared:  0.117,  Adjusted R-squared:  0.113
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
##  359 weights are ~= 1. The remaining 4590 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0009  0.8790  0.9410  0.9110  0.9830  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          2.02e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.029  1           1.014
## LastAuthorFemale  1.023  1           1.012
## Year              1.034 16           1.001
```



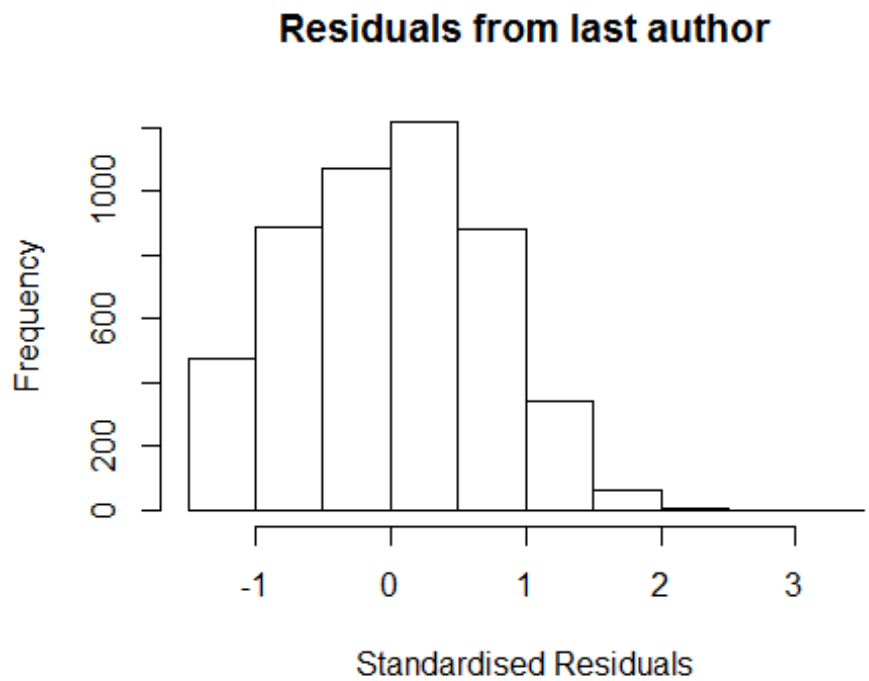

```
## [1] "List of  2 outliers with residuals above  2.5"
##        ScopusId  NLCS Year OneField Fields residuals
## 507  0031097861 3.492 1997     2205      4     2.546
## 2616 0035492228 4.506 2001     1909      2     3.370
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.3064 -0.5771  0.0196  0.5255  3.3703
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)           0.8719     0.0513   17.01  < 2e-16 ***
## FirstAuthorFemale1    0.1659     0.0353    4.71  2.6e-06 ***
## LastAuthorFemale1    -0.0142     0.0357   -0.40  0.69217
## Year1997              0.0742     0.0741    1.00  0.31638
## Year1998              0.1893     0.0739    2.56  0.01041 *
## Year1999              0.2568     0.0730    3.52  0.00044 ***
## Year2000              0.2686     0.0709    3.79  0.00015 ***
## Year2001              0.2638     0.0905    2.92  0.00356 **
## Year2002              0.1900     0.0741    2.56  0.01035 *
## Year2003              0.1541     0.0755    2.04  0.04135 *
## Year2004              0.1493     0.0690    2.16  0.03061 *
```

```
## Year2005             0.1121     0.0683    1.64  0.10081
## Year2006             0.1565     0.0638    2.45  0.01420 *
## Year2007             0.0786     0.0627    1.25  0.20957
## Year2008             0.0656     0.0642    1.02  0.30675
## Year2009             0.1873     0.0635    2.95  0.00319 **
## Year2010             0.1585     0.0635    2.49  0.01263 *
## Year2011             0.1082     0.0616    1.76  0.07896 .
## Year2012             0.1401     0.0618    2.27  0.02347 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.785
## Multiple R-squared:  0.0123, Adjusted R-squared:  0.0087
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  390 weights are ~= 1. The remaining 4559 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0256  0.8720  0.9480  0.9240  0.9840  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          2.02e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.021  1           1.011
## Year              1.021 16           1.001
```




```
## [1] "List of  2 outliers with residuals above  2.5"
##        ScopusId  NLCS Year OneField Fields residuals
## 507  0031097861 3.492 1997     2205      4     2.546
## 2616 0035492228 4.506 2001     1909      2     3.370
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.3025 -0.5781  0.0193  0.5267  3.3709
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)           0.8713     0.0513   16.99  < 2e-16 ***
## FirstAuthorFemale1    0.1634     0.0353    4.63  3.8e-06 ***
## Year1997              0.0740     0.0741    1.00  0.31770
## Year1998              0.1894     0.0739    2.56  0.01039 *
## Year1999              0.2565     0.0731    3.51  0.00045 ***
## Year2000              0.2679     0.0709    3.78  0.00016 ***
## Year2001              0.2638     0.0905    2.91  0.00359 **
## Year2002              0.1897     0.0741    2.56  0.01052 *
## Year2003              0.1540     0.0755    2.04  0.04149 *
## Year2004              0.1492     0.0690    2.16  0.03068 *
## Year2005              0.1120     0.0683    1.64  0.10121
```

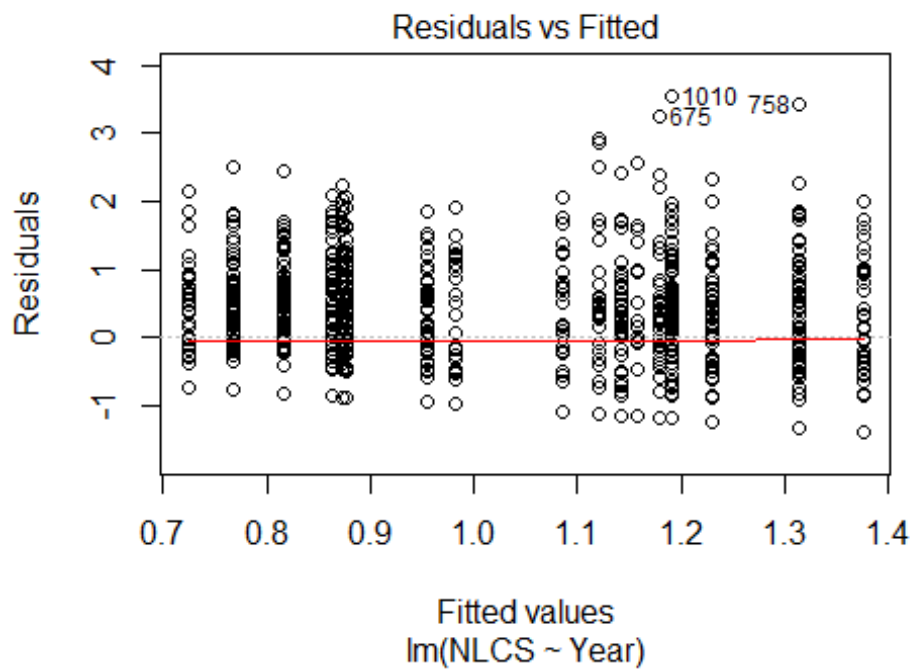
```
## Year2006             0.1557     0.0638    2.44  0.01467 *
## Year2007             0.0780     0.0626    1.25  0.21302
## Year2008             0.0652     0.0642    1.02  0.30944
## Year2009             0.1872     0.0635    2.95  0.00323 **
## Year2010             0.1580     0.0635    2.49  0.01289 *
## Year2011             0.1077     0.0616    1.75  0.08011 .
## Year2012             0.1396     0.0618    2.26  0.02395 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.785
## Multiple R-squared:  0.0123, Adjusted R-squared:  0.00888
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  397 weights are ~= 1. The remaining 4552 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0256  0.8720  0.9480  0.9240  0.9840  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          2.02e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##                   GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.015  1           1.007
## Year             1.015 16           1.000
```



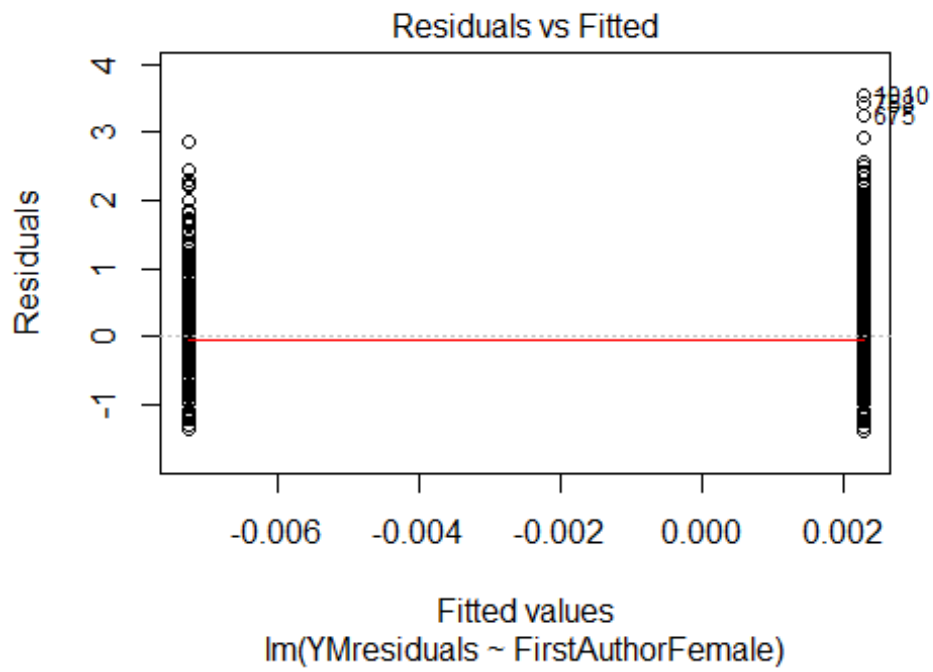

```
## [1] "List of  2 outliers with residuals above  2.5"
##        ScopusId  NLCS Year OneField Fields residuals
## 507  0031097861 3.492 1997     2205      4     2.546
## 2616 0035492228 4.506 2001     1909      2     3.370
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.1759 -0.5600  0.0182  0.5278  3.3489
##
## Coefficients:
##                   Estimate Std. Error t value Pr(>|t|)
## (Intercept)         0.8781     0.0514   17.09  < 2e-16 ***
## LastAuthorFemale1   0.0188     0.0353    0.53  0.59419
## Year1997            0.0740     0.0743    1.00  0.31935
## Year1998            0.1890     0.0739    2.56  0.01061 *
## Year1999            0.2621     0.0731    3.58  0.00034 ***
## Year2000            0.2761     0.0711    3.88  0.00010 ***
## Year2001            0.2790     0.0910    3.07  0.00218 **
## Year2002            0.1922     0.0743    2.59  0.00972 **
## Year2003            0.1574     0.0759    2.07  0.03817 *
## Year2004            0.1521     0.0690    2.21  0.02749 *
## Year2005            0.1183     0.0684    1.73  0.08344 .
```

```
## Year2006            0.1658     0.0638    2.60  0.00945 **
## Year2007            0.0867     0.0628    1.38  0.16738
## Year2008            0.0723     0.0644    1.12  0.26165
## Year2009            0.1924     0.0636    3.02  0.00251 **
## Year2010            0.1700     0.0636    2.67  0.00755 **
## Year2011            0.1159     0.0617    1.88  0.06035 .
## Year2012            0.1554     0.0618    2.51  0.01201 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.786
## Multiple R-squared:  0.00822,    Adjusted R-squared:  0.00481
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  392 weights are ~= 1. The remaining 4557 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0297  0.8710  0.9480  0.9240  0.9840  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          2.02e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Sample size for the above analysis:  4949"
## [1] ""
## [1] ""
## [1] "####################################"
## [1] "Analysis of AJSC 2216"
## [1] "####################################"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##   95   81   74   77   69   70   53   65   68  107  111  148  131  111  149
## 2011 2012
##  144  150
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##   63   54   43   44   52   28   47   43   45   89   81  116   97   89  106
```

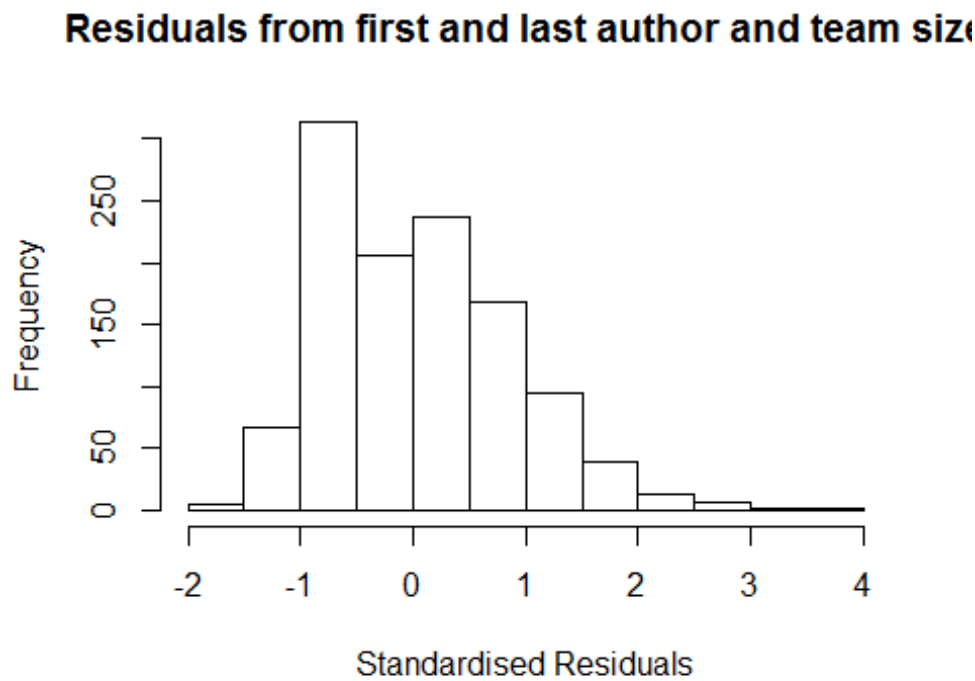
```
## 2011 2012
##  117  124
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##   59   50   40   41   50   25   46   39   36   81   72  107   92   84   98
## 2011 2012
##  113  120
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
##  Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 33, df = 16, p-value = 0.007
```




```
##
##  Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 0.16, df = 1, p-value = 0.7
```



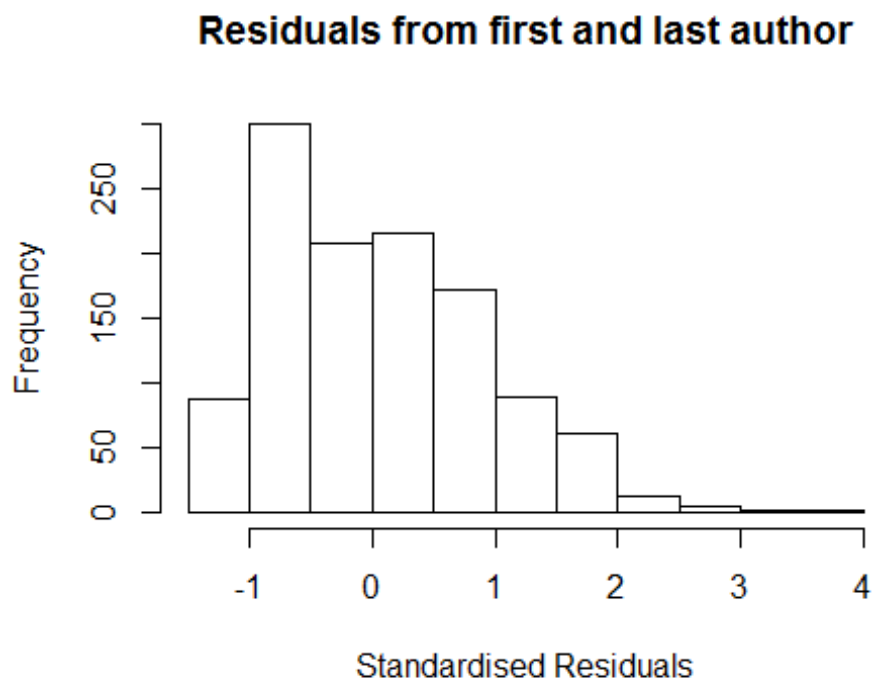

```
## [1] "Female first author team size 2018 geometric mean: 1.22460111047784"
## [1] "Male first author team size 2018 geometric mean: 1.46757035945684"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 2400, p-value = 0.03
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 1.24444598992456"
## [1] "Male last author team size 2018 geometric mean: 1.45457271403248"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 2500, p-value = 0.04
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 2.608  1           1.615
## LastAuthorFemale  2.667  1           1.633
## UniqueAuthors     1.289  4           1.032
## Year              1.419 16           1.011
```




```
## [1] "List of  9 outliers with residuals above  2.5"
##         ScopusId  NLCS Year OneField Fields residuals
## 400  85011501181 3.568 2000     1213      2     2.850
## 545   0036504271 3.965 2002     2205      4     2.523
## 637  66449129248 3.566 2004     1213      2     2.956
## 686  15844375537 3.375 2004     1213      2     2.671
## 726  60950110696 3.592 2005     1202      3     2.549
## 758  26444586054 4.731 2005     2216      1     3.330
## 1010 84973631259 4.730 2007     1206      3     3.825
## 1108 65849379504 3.117 2008     1213      2     2.538
## 1793 84861812124 3.270 2012     1111      4     2.807
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
UniqueAuthors +
##     Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##     k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.7834 -0.5826 -0.0314  0.5901  3.8250
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)           0.7849     0.1114    7.05  3.2e-12 ***
## FirstAuthorFemale1    0.0498     0.0949    0.52   0.6001
```

```
## LastAuthorFemale1    0.0447     0.0955    0.47   0.6397
## UniqueAuthors2       0.4525     0.0690    6.55  8.5e-11 ***
## UniqueAuthors3       0.8353     0.0919    9.09  < 2e-16 ***
## UniqueAuthors4       0.6809     0.1431    4.76  2.2e-06 ***
## UniqueAuthors5       0.6139     0.1956    3.14   0.0017 **
## Year1997            -0.3351     0.1454   -2.30   0.0214 *
## Year1998            -0.1847     0.1761   -1.05   0.2945
## Year1999             0.0113     0.1766    0.06   0.9490
## Year2000            -0.0672     0.1658   -0.40   0.6856
## Year2001            -0.0106     0.1885   -0.06   0.9552
## Year2002            -0.0514     0.1902   -0.27   0.7870
## Year2003             0.2656     0.1998    1.33   0.1841
## Year2004            -0.1751     0.1674   -1.05   0.2956
## Year2005             0.1633     0.1613    1.01   0.3117
## Year2006             0.1158     0.1569    0.74   0.4606
## Year2007             0.1202     0.1361    0.88   0.3775
## Year2008            -0.2054     0.1459   -1.41   0.1594
## Year2009            -0.2414     0.1417   -1.70   0.0888 .
## Year2010            -0.2429     0.1373   -1.77   0.0771 .
## Year2011            -0.1999     0.1311   -1.52   0.1276
## Year2012            -0.3222     0.1295   -2.49   0.0130 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.819
## Multiple R-squared:  0.155,  Adjusted R-squared:  0.138
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
##  observation 587 is an outlier with |weight| <= 3.8e-05 ( < 8.7e-05);
##  78 weights are ~= 1. The remaining 1074 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0608  0.8820  0.9520  0.9090  0.9770  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          8.67e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                  psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
```

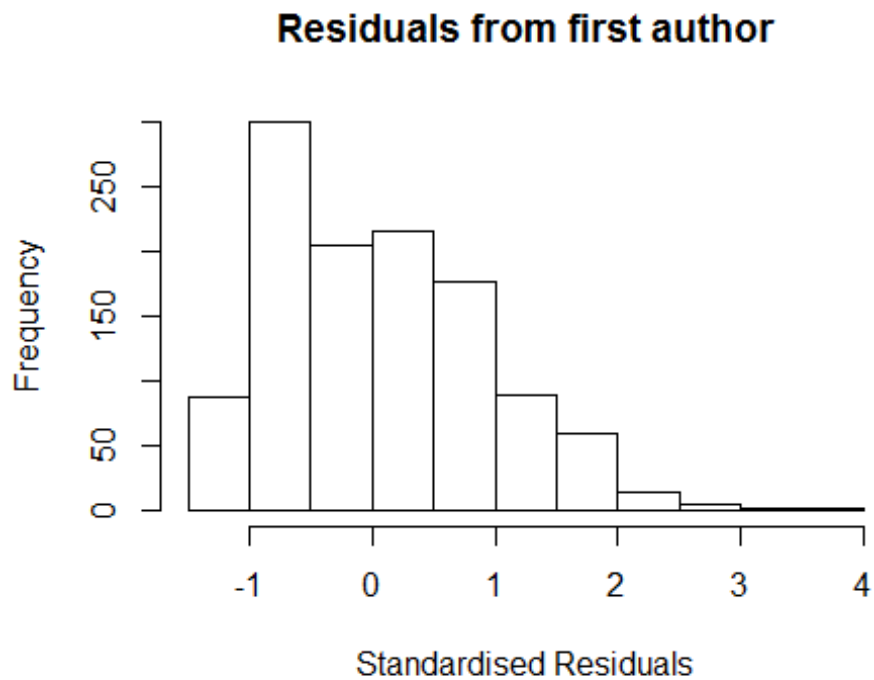
```
factors"
##                     GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 2.508  1           1.584
## LastAuthorFemale  2.551  1           1.597
## Year              1.133 16           1.004
```




```
## [1] "List of  8 outliers with residuals above  2.5"
##         ScopusId  NLCS Year OneField Fields residuals
## 400  85011501181 3.568 2000     1213      2     2.503
## 527   0036763450 3.614 2002     2205      4     2.693
## 545   0036504271 3.965 2002     2205      4     3.013
## 637  66449129248 3.566 2004     1213      2     2.633
## 758  26444586054 4.731 2005     2216      1     3.562
## 1010 84973631259 4.730 2007     1206      3     3.602
## 1514 77952198073 3.247 2010     1202      3     2.538
## 1793 84861812124 3.270 2012     1111      4     2.578
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.3761 -0.7264 -0.0235  0.6360  3.6024
##
## Coefficients:
```

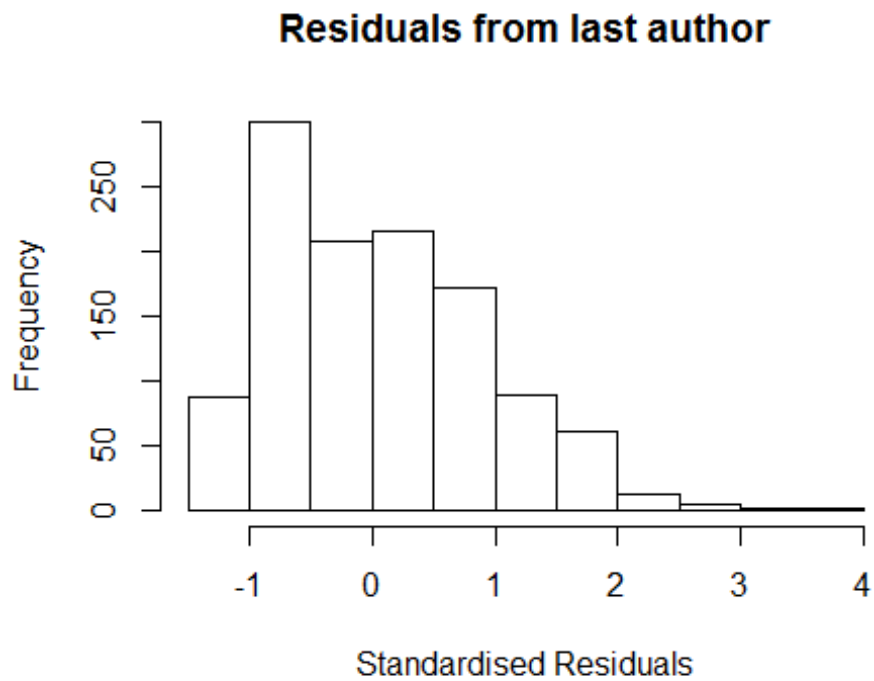
```
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)          0.92698    0.11189    8.28  3.3e-16 ***
## FirstAuthorFemale1 -0.01707    0.09758   -0.17    0.861
## LastAuthorFemale1   0.04800    0.09847    0.49    0.626
## Year1997           -0.26112    0.15163   -1.72    0.085 .
## Year1998           -0.07918    0.17603   -0.45    0.653
## Year1999           -0.02025    0.19268   -0.11    0.916
## Year2000            0.13814    0.17169    0.80    0.421
## Year2001            0.07584    0.22441    0.34    0.735
## Year2002           -0.00641    0.20130   -0.03    0.975
## Year2003            0.41824    0.22354    1.87    0.062 .
## Year2004            0.00628    0.19000    0.03    0.974
## Year2005            0.24252    0.16494    1.47    0.142
## Year2006            0.24149    0.15290    1.58    0.115
## Year2007            0.20059    0.14207    1.41    0.158
## Year2008           -0.19166    0.14804   -1.29    0.196
## Year2009           -0.14220    0.15039   -0.95    0.345
## Year2010           -0.20056    0.14082   -1.42    0.155
## Year2011           -0.10601    0.13387   -0.79    0.429
## Year2012           -0.23546    0.13470   -1.75    0.081 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.864
## Multiple R-squared:  0.0451, Adjusted R-squared:  0.0299
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
##  86 weights are ~= 1. The remaining 1067 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0433  0.8880  0.9370  0.9070  0.9820  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          8.67e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##                     GVIF Df GVIF^(1/(2*Df))
```

```
## FirstAuthorFemale 1.078  1           1.038
## Year              1.078 16           1.002
```




```
## [1] "List of  8 outliers with residuals above  2.5"
##         ScopusId  NLCS Year OneField Fields residuals
## 400  85011501181 3.568 2000     1213      2     2.503
## 527   0036763450 3.614 2002     2205      4     2.693
## 545   0036504271 3.965 2002     2205      4     3.013
## 637  66449129248 3.566 2004     1213      2     2.633
## 758  26444586054 4.731 2005     2216      1     3.562
## 1010 84973631259 4.730 2007     1206      3     3.602
## 1514 77952198073 3.247 2010     1202      3     2.538
## 1793 84861812124 3.270 2012     1111      4     2.578
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.3664 -0.7281 -0.0241  0.6379  3.5990
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         0.93068    0.11184    8.32  2.5e-16 ***
## FirstAuthorFemale1  0.01821    0.06386    0.29    0.776
```

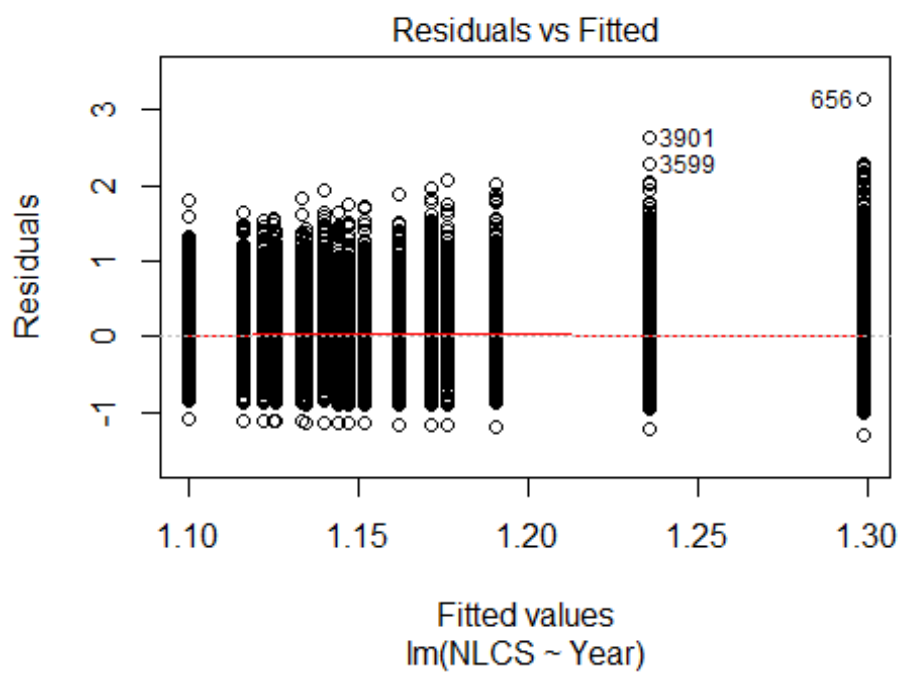
```
## Year1997           -0.26188    0.15147   -1.73    0.084 .
## Year1998           -0.08267    0.17610   -0.47    0.639
## Year1999           -0.02419    0.19288   -0.13    0.900
## Year2000            0.13580    0.17165    0.79    0.429
## Year2001            0.07310    0.22392    0.33    0.744
## Year2002           -0.00532    0.20122   -0.03    0.979
## Year2003            0.41753    0.22327    1.87    0.062 .
## Year2004            0.00603    0.18998    0.03    0.975
## Year2005            0.24320    0.16486    1.48    0.140
## Year2006            0.24244    0.15248    1.59    0.112
## Year2007            0.20034    0.14222    1.41    0.159
## Year2008           -0.19220    0.14776   -1.30    0.194
## Year2009           -0.14221    0.15058   -0.94    0.345
## Year2010           -0.20256    0.14076   -1.44    0.150
## Year2011           -0.10713    0.13383   -0.80    0.424
## Year2012           -0.23457    0.13459   -1.74    0.082 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.863
## Multiple R-squared:  0.0448, Adjusted R-squared:  0.0305
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
##  88 weights are ~= 1. The remaining 1065 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0431  0.8880  0.9360  0.9070  0.9820  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          8.67e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##                   GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.096  1           1.047
## Year             1.096 16           1.003
```



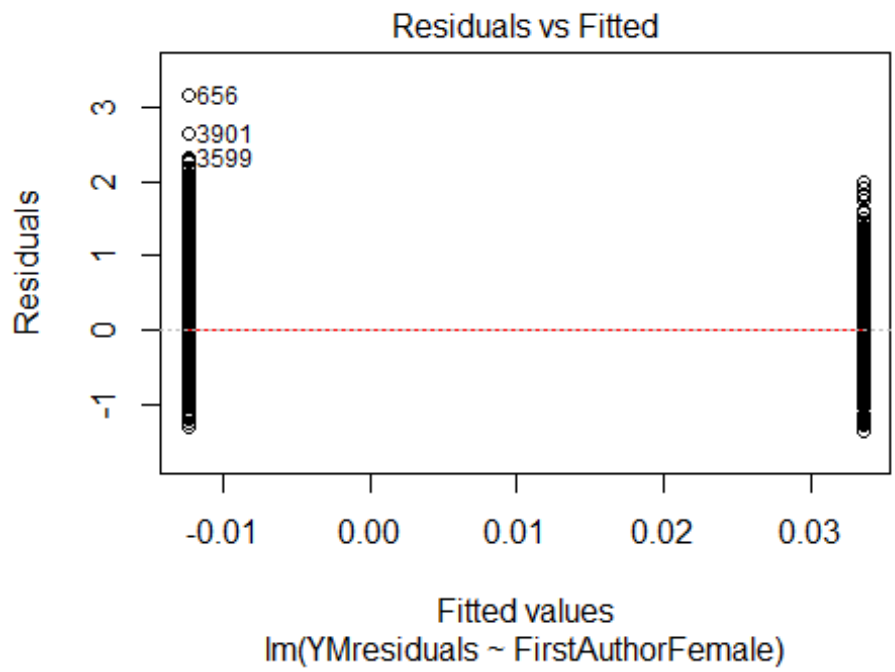

```
## [1] "List of  8 outliers with residuals above  2.5"
##         ScopusId  NLCS Year OneField Fields residuals
## 400  85011501181 3.568 2000     1213      2     2.503
## 527   0036763450 3.614 2002     2205      4     2.693
## 545   0036504271 3.965 2002     2205      4     3.013
## 637  66449129248 3.566 2004     1213      2     2.633
## 758  26444586054 4.731 2005     2216      1     3.562
## 1010 84973631259 4.730 2007     1206      3     3.602
## 1514 77952198073 3.247 2010     1202      3     2.538
## 1793 84861812124 3.270 2012     1111      4     2.578
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.3799 -0.7255 -0.0236  0.6374  3.6035
##
## Coefficients:
##                   Estimate Std. Error t value Pr(>|t|)
## (Intercept)        0.92663    0.11178    8.29  3.2e-16 ***
## LastAuthorFemale1  0.03557    0.06468    0.55    0.583
## Year1997          -0.26201    0.15177   -1.73    0.085 .
## Year1998          -0.07964    0.17598   -0.45    0.651
## Year1999          -0.02070    0.19272   -0.11    0.914
```

```
## Year2000           0.13746    0.17156    0.80    0.423
## Year2001           0.07487    0.22426    0.33    0.739
## Year2002          -0.00775    0.20141   -0.04    0.969
## Year2003           0.41772    0.22348    1.87    0.062 .
## Year2004           0.00611    0.18998    0.03    0.974
## Year2005           0.24133    0.16479    1.46    0.143
## Year2006           0.24197    0.15266    1.59    0.113
## Year2007           0.19989    0.14181    1.41    0.159
## Year2008          -0.19251    0.14804   -1.30    0.194
## Year2009          -0.14256    0.15044   -0.95    0.344
## Year2010          -0.20178    0.14079   -1.43    0.152
## Year2011          -0.10702    0.13376   -0.80    0.424
## Year2012          -0.23673    0.13482   -1.76    0.079 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.864
## Multiple R-squared:  0.045,  Adjusted R-squared:  0.0307
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
##  89 weights are ~= 1. The remaining 1064 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0431  0.8890  0.9370  0.9070  0.9810  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          8.67e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Sample size for the above analysis:  1153"
## [1] ""
## [1] ""
## [1] "####################################"
## [1] "Analysis of AJSC 2300"
## [1] "####################################"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
```

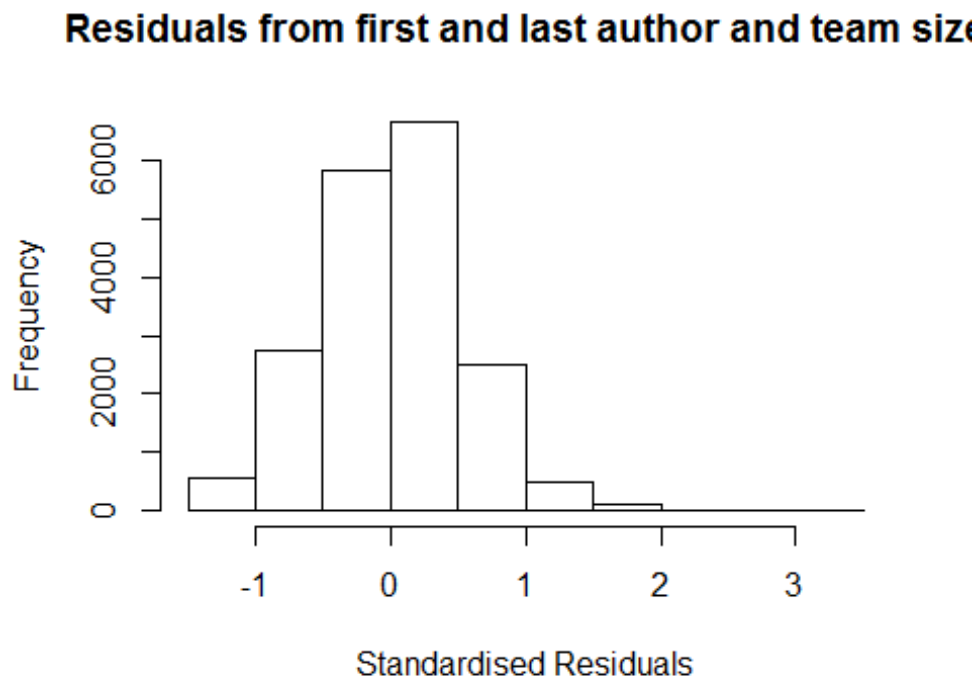
```
## 1961 2054 1596 1708 1777 1895 1783 1754 1870 1993 2329 2382 2279 1623 1493
## 2011 2012
## 1562 1720
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 1226 1274 1074 1180 1031  889 1249 1248 1309 1397 1636 1720 1573 1167 1057
## 2011 2012
## 1065 1181
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 1091 1155  981 1054  921  793 1118 1120 1162 1226 1423 1513 1390 1031  937
## 2011 2012
##  941 1042
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
##  Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 510, df = 16, p-value <2e-16
```




```
##
##  Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 95, df = 1, p-value <2e-16
```



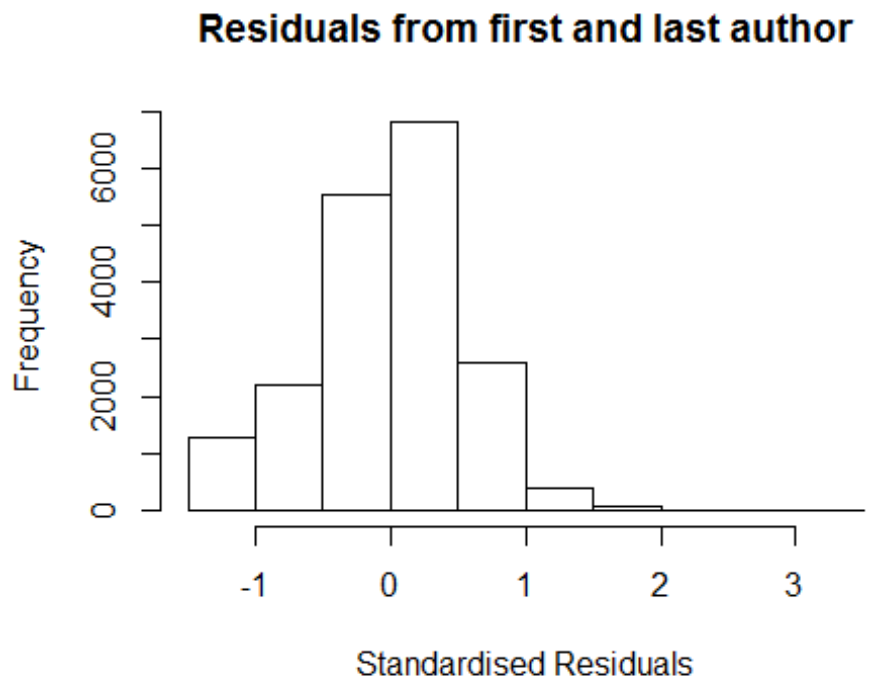

```
## [1] "Female first author team size 2018 geometric mean: 3.3681290380735"
## [1] "Male first author team size 2018 geometric mean: 2.90915837026471"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 160000, p-value = 4e-04
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 3.04297066140811"
## [1] "Male last author team size 2018 geometric mean: 3.10455314298334"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 130000, p-value = 0.7
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.036  1           1.018
## LastAuthorFemale  1.027  1           1.014
## UniqueAuthors     1.036  4           1.004
## Year              1.043 16           1.001
```




```
## [1] "List of  4 outliers with residuals above  2.5"
##         ScopusId  NLCS Year OneField Fields residuals
## 656   0001690750 4.435 1996     1909      2     3.400
## 1417  0030443129 3.584 1996     1900      2     2.549
## 1457 33746136469 3.554 1996     1900      2     2.519
## 3901  0031233750 3.858 1997     2205      3     2.563
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
UniqueAuthors +
##     Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##     k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.4661 -0.3390  0.0196  0.3462  3.4003
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)          1.03469    0.02311   44.77  < 2e-16 ***
## FirstAuthorFemale1   0.02539    0.00857    2.96   0.0031 **
## LastAuthorFemale1    0.00711    0.00968    0.73   0.4624
## UniqueAuthors2       0.30285    0.01295   23.38  < 2e-16 ***
## UniqueAuthors3       0.33109    0.01343   24.66  < 2e-16 ***
## UniqueAuthors4       0.37236    0.01447   25.73  < 2e-16 ***
## UniqueAuthors5       0.42426    0.01359   31.22  < 2e-16 ***
```

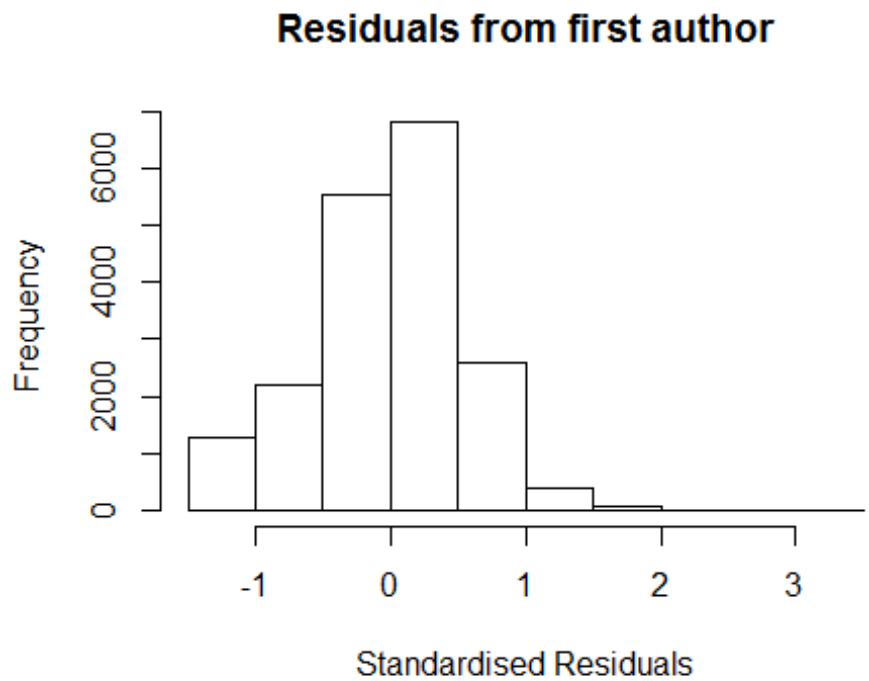
```
## Year1997           -0.04931    0.02913   -1.69   0.0905 .
## Year1998           -0.10942    0.02938   -3.72   0.0002 ***
## Year1999           -0.11704    0.02716   -4.31  1.6e-05 ***
## Year2000           -0.14090    0.02783   -5.06  4.2e-07 ***
## Year2001           -0.14874    0.02773   -5.36  8.3e-08 ***
## Year2002           -0.17166    0.02604   -6.59  4.5e-11 ***
## Year2003           -0.15673    0.02581   -6.07  1.3e-09 ***
## Year2004           -0.18531    0.02526   -7.34  2.3e-13 ***
## Year2005           -0.18444    0.02485   -7.42  1.2e-13 ***
## Year2006           -0.22319    0.02486   -8.98  < 2e-16 ***
## Year2007           -0.22925    0.02459   -9.32  < 2e-16 ***
## Year2008           -0.21411    0.02467   -8.68  < 2e-16 ***
## Year2009           -0.18200    0.02709   -6.72  1.9e-11 ***
## Year2010           -0.12751    0.02751   -4.64  3.6e-06 ***
## Year2011           -0.18188    0.02877   -6.32  2.6e-10 ***
## Year2012           -0.14084    0.02790   -5.05  4.5e-07 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.51
## Multiple R-squared:  0.082,  Adjusted R-squared:  0.0809
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
##  6 observations c(125,306,702,734,737,2026)
##   are outliers with |weight| = 0 ( < 5.3e-06);
##  1613 weights are ~= 1. The remaining 17279 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0018  0.8610  0.9500  0.8990  0.9860  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          5.29e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.025  1           1.012
```

```
## LastAuthorFemale  1.021  1           1.010
## Year              1.018 16           1.001
```



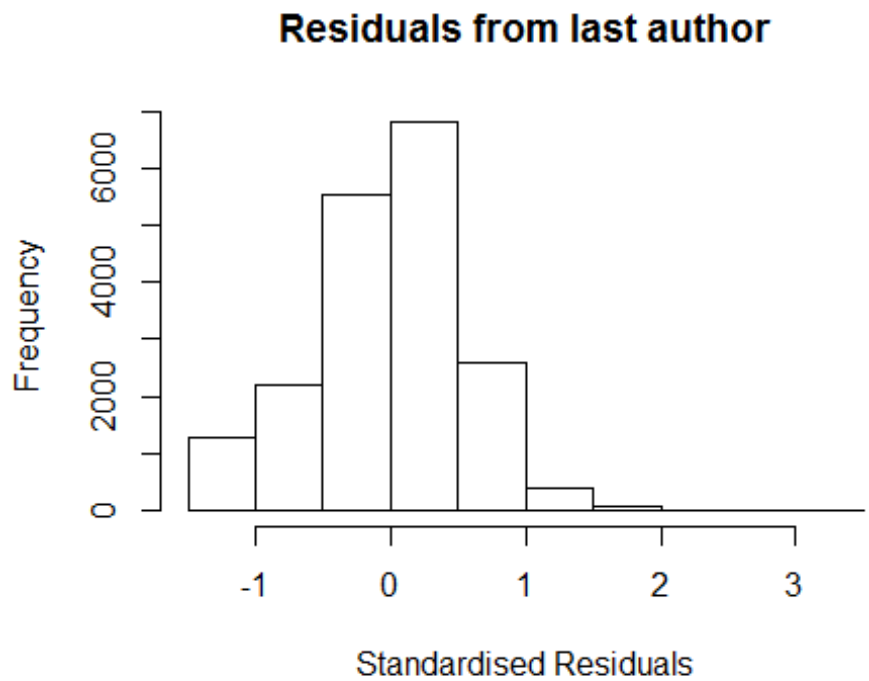

```
## [1] "List of  2 outliers with residuals above  2.5"
##        ScopusId  NLCS Year OneField Fields residuals
## 656  0001690750 4.435 1996     1909      2     3.162
## 3901 0031233750 3.858 1997     2205      3     2.635
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.3272 -0.3468  0.0305  0.3583  3.1621
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.272875   0.021726   58.59  < 2e-16 ***
## FirstAuthorFemale1  0.053759   0.008769    6.13  8.9e-10 ***
## LastAuthorFemale1   0.000598   0.009978    0.06  0.95220
## Year1997           -0.050696   0.030224   -1.68  0.09349 .
## Year1998           -0.108154   0.030564   -3.54  0.00040 ***
## Year1999           -0.106621   0.027969   -3.81  0.00014 ***
## Year2000           -0.138180   0.028741   -4.81  1.5e-06 ***
## Year2001           -0.122616   0.028584   -4.29  1.8e-05 ***
```

```
## Year2002           -0.149331   0.027145   -5.50  3.8e-08 ***
## Year2003           -0.132823   0.026617   -4.99  6.1e-07 ***
## Year2004           -0.149734   0.026042   -5.75  9.1e-09 ***
## Year2005           -0.140850   0.025696   -5.48  4.3e-08 ***
## Year2006           -0.182189   0.025778   -7.07  1.6e-12 ***
## Year2007           -0.176186   0.025287   -6.97  3.3e-12 ***
## Year2008           -0.162153   0.025390   -6.39  1.7e-10 ***
## Year2009           -0.151570   0.028112   -5.39  7.1e-08 ***
## Year2010           -0.101972   0.028485   -3.58  0.00034 ***
## Year2011           -0.142472   0.030002   -4.75  2.1e-06 ***
## Year2012           -0.101269   0.029032   -3.49  0.00049 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.52
## Multiple R-squared:  0.00807,    Adjusted R-squared:  0.00712
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
##  2 observations c(306,2026) are outliers with |weight| = 0 ( < 5.3e-06);
##  1493 weights are ~= 1. The remaining 17403 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0098  0.8610  0.9500  0.8960  0.9860  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          5.29e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.012  1           1.006
## Year              1.012 16           1.000
```




```
## [1] "List of  2 outliers with residuals above  2.5"
##        ScopusId  NLCS Year OneField Fields residuals
## 656  0001690750 4.435 1996     1909      2     3.162
## 3901 0031233750 3.858 1997     2205      3     2.635
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.3268 -0.3468  0.0305  0.3582  3.1621
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.27294    0.02170   58.67  < 2e-16 ***
## FirstAuthorFemale1  0.05389    0.00879    6.13  8.9e-10 ***
## Year1997           -0.05069    0.03022   -1.68  0.09351 .
## Year1998           -0.10814    0.03056   -3.54  0.00040 ***
## Year1999           -0.10660    0.02797   -3.81  0.00014 ***
## Year2000           -0.13816    0.02874   -4.81  1.5e-06 ***
## Year2001           -0.12260    0.02858   -4.29  1.8e-05 ***
## Year2002           -0.14931    0.02714   -5.50  3.8e-08 ***
## Year2003           -0.13278    0.02660   -4.99  6.0e-07 ***
## Year2004           -0.14970    0.02603   -5.75  9.0e-09 ***
## Year2005           -0.14083    0.02569   -5.48  4.3e-08 ***
```

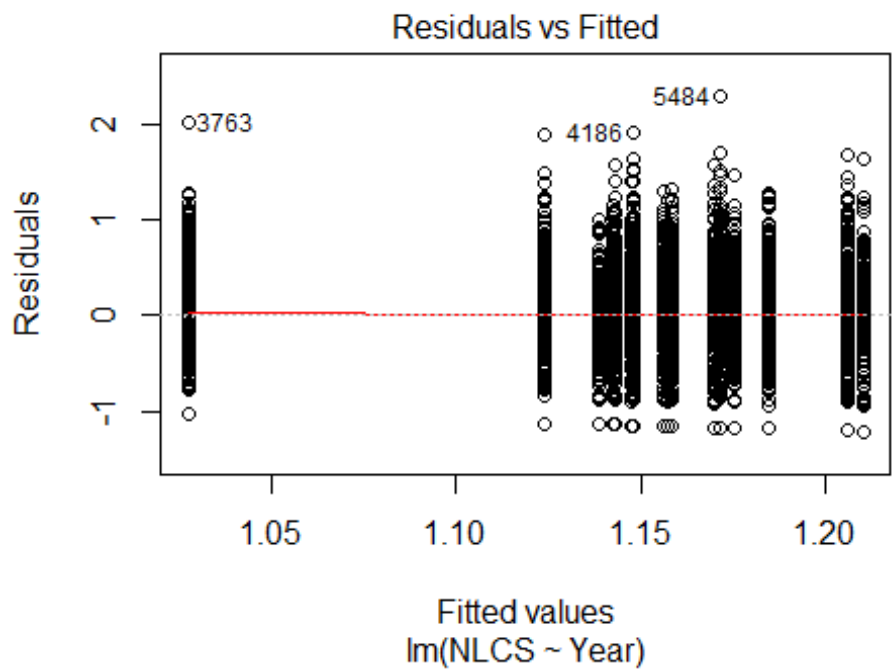
```
## Year2006           -0.18216    0.02577   -7.07  1.6e-12 ***
## Year2007           -0.17615    0.02527   -6.97  3.2e-12 ***
## Year2008           -0.16212    0.02538   -6.39  1.7e-10 ***
## Year2009           -0.15154    0.02810   -5.39  7.0e-08 ***
## Year2010           -0.10196    0.02849   -3.58  0.00035 ***
## Year2011           -0.14243    0.02998   -4.75  2.0e-06 ***
## Year2012           -0.10123    0.02902   -3.49  0.00049 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.52
## Multiple R-squared:  0.00807,    Adjusted R-squared:  0.00717
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
##  2 observations c(306,2026) are outliers with |weight| = 0 ( < 5.3e-06);
##  1493 weights are ~= 1. The remaining 17403 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0098  0.8610  0.9500  0.8960  0.9860  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          5.29e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##                   GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.008  1           1.004
## Year             1.008 16           1.000
```



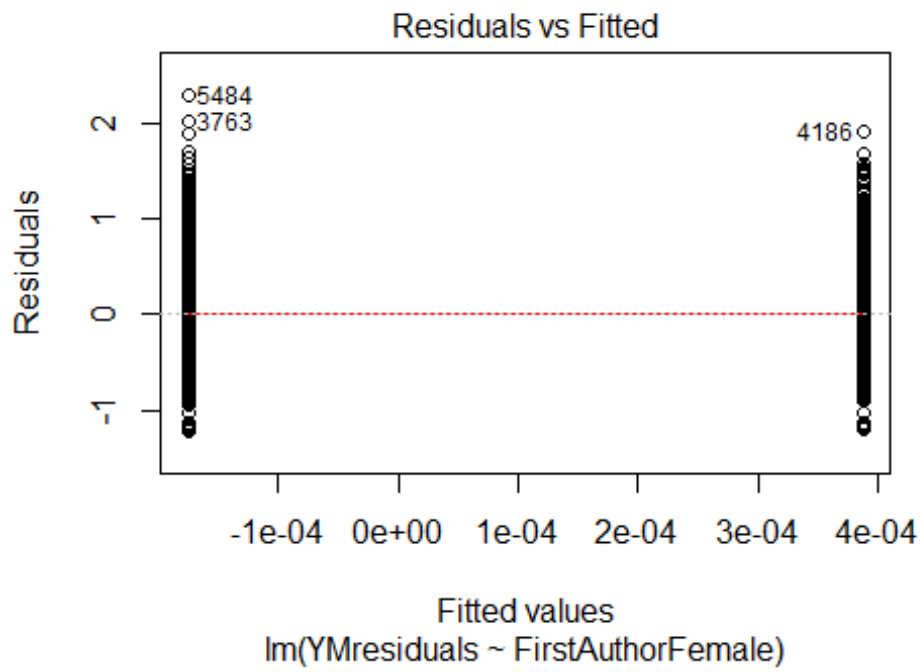

```
## [1] "List of  2 outliers with residuals above  2.5"
##        ScopusId  NLCS Year OneField Fields residuals
## 656  0001690750 4.435 1996     1909      2     3.162
## 3901 0031233750 3.858 1997     2205      3     2.635
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.2963 -0.3459  0.0293  0.3578  3.1536
##
## Coefficients:
##                   Estimate Std. Error t value Pr(>|t|)
## (Intercept)        1.28145    0.02170   59.06  < 2e-16 ***
## LastAuthorFemale1  0.01482    0.00994    1.49  0.13588
## Year1997          -0.04971    0.03027   -1.64  0.10055
## Year1998          -0.10772    0.03063   -3.52  0.00044 ***
## Year1999          -0.10574    0.02802   -3.77  0.00016 ***
## Year2000          -0.13650    0.02879   -4.74  2.1e-06 ***
## Year2001          -0.12122    0.02865   -4.23  2.3e-05 ***
## Year2002          -0.14624    0.02716   -5.38  7.4e-08 ***
## Year2003          -0.13041    0.02667   -4.89  1.0e-06 ***
## Year2004          -0.14614    0.02606   -5.61  2.1e-08 ***
## Year2005          -0.13704    0.02572   -5.33  1.0e-07 ***
```

```
## Year2006          -0.17793    0.02581   -6.89  5.6e-12 ***
## Year2007          -0.17054    0.02532   -6.74  1.7e-11 ***
## Year2008          -0.15622    0.02542   -6.15  8.1e-10 ***
## Year2009          -0.14548    0.02813   -5.17  2.3e-07 ***
## Year2010          -0.09603    0.02852   -3.37  0.00076 ***
## Year2011          -0.13661    0.03001   -4.55  5.4e-06 ***
## Year2012          -0.09531    0.02906   -3.28  0.00104 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.52
## Multiple R-squared:  0.0063, Adjusted R-squared:  0.00541
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
##  2 observations c(306,2026) are outliers with |weight| = 0 ( < 5.3e-06);
##  1533 weights are ~= 1. The remaining 17363 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0112  0.8600  0.9500  0.8960  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          5.29e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                  psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Sample size for the above analysis:  18898"
## [1] ""
## [1] ""
## [1] "##################################"
## [1] "Analysis of AJSC 2301"
## [1] "##################################"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  266  257  229  255  251  270  365  328  319  273  310  332  378  307  374
## 2011 2012
##  372  359
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
```

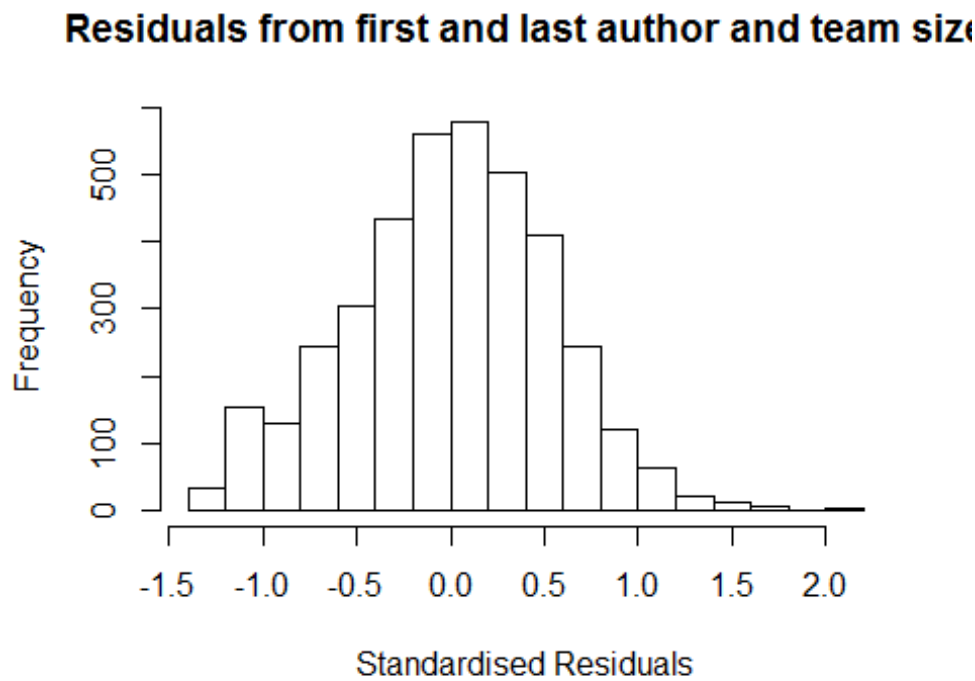
```
##  178  188  175  185  163  105  312  282  266  235  259  270  300  257  301
## 2011 2012
##  305  279
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  171  181  168  177  158   99  302  263  243  223  253  255  279  238  277
## 2011 2012
##  279  258
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
##  Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 43, df = 16, p-value = 2e-04
```




```
##
##  Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 0.11, df = 1, p-value = 0.7
```

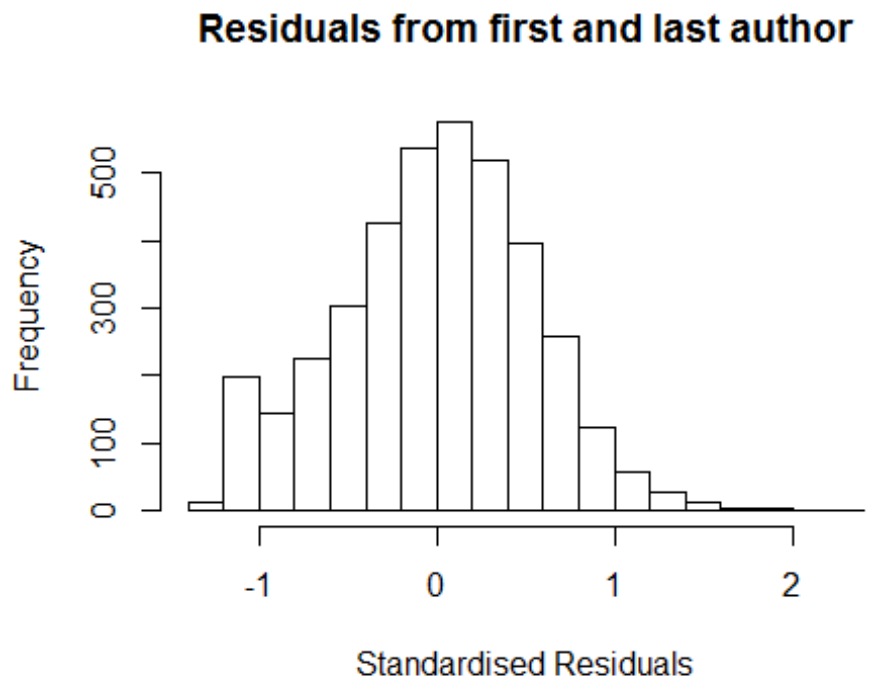



```
## [1] "Female first author team size 2018 geometric mean: 2.4453681098498"
## [1] "Male first author team size 2018 geometric mean: 2.21219226780385"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 22000, p-value = 0.1
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 2.23118747938308"
## [1] "Male last author team size 2018 geometric mean: 2.34033495153309"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 17000, p-value = 0.6
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.664  1           1.290
## LastAuthorFemale  1.661  1           1.289
## UniqueAuthors     1.113  4           1.013
## Year              1.133 16           1.004
```



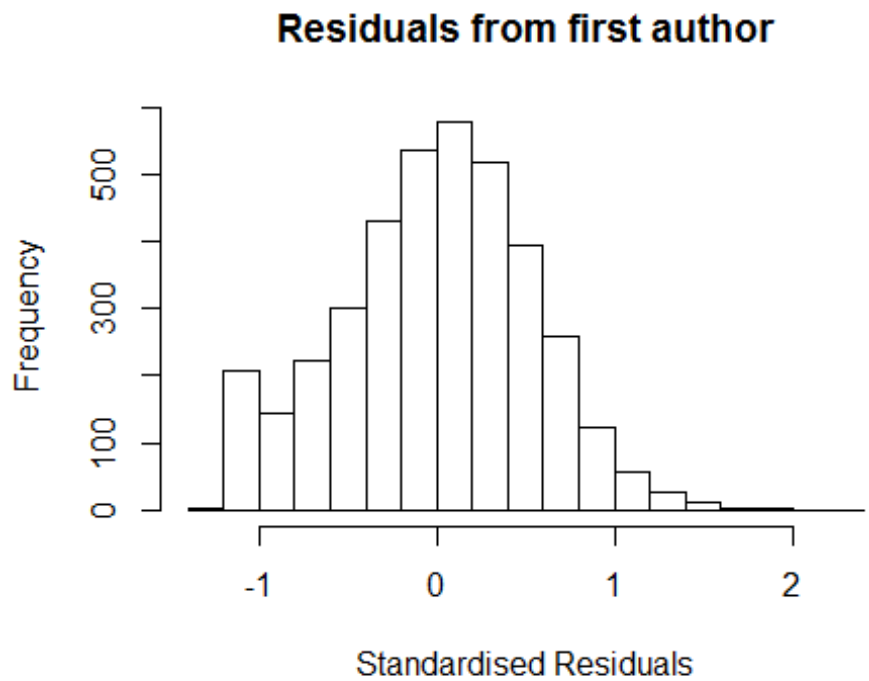

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
UniqueAuthors +
##     Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##     k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.3590 -0.3514  0.0133  0.3543  2.1684
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.07297    0.04169   25.74  < 2e-16 ***
## FirstAuthorFemale1 -0.03414    0.02508   -1.36    0.174
## LastAuthorFemale1   0.05574    0.02555    2.18    0.029 *
## UniqueAuthors2      0.14145    0.02130    6.64  3.6e-11 ***
## UniqueAuthors3      0.23029    0.02743    8.40  < 2e-16 ***
## UniqueAuthors4      0.18973    0.04297    4.42  1.0e-05 ***
## UniqueAuthors5      0.23449    0.05698    4.11  4.0e-05 ***
## Year1997            0.00933    0.06021    0.16    0.877
## Year1998            0.06490    0.05976    1.09    0.278
## Year1999            0.00805    0.05709    0.14    0.888
```

```
## Year2000            0.00922    0.05551    0.17    0.868
## Year2001            0.04968    0.07246    0.69    0.493
## Year2002            0.05535    0.05178    1.07    0.285
## Year2003            0.00512    0.05201    0.10    0.922
## Year2004            0.01600    0.05299    0.30    0.763
## Year2005            0.01120    0.05668    0.20    0.843
## Year2006            0.01477    0.05154    0.29    0.774
## Year2007           -0.12250    0.05869   -2.09    0.037 *
## Year2008            0.00765    0.05353    0.14    0.886
## Year2009           -0.01880    0.05435   -0.35    0.729
## Year2010            0.04185    0.05421    0.77    0.440
## Year2011           -0.05217    0.05409   -0.96    0.335
## Year2012           -0.02189    0.05742   -0.38    0.703
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.53
## Multiple R-squared:  0.0329, Adjusted R-squared:  0.0273
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  315 weights are ~= 1. The remaining 3509 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0566  0.8650  0.9500  0.9040  0.9860  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          2.62e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.720  1           1.312
## LastAuthorFemale  1.712  1           1.309
## Year              1.032 16           1.001
```



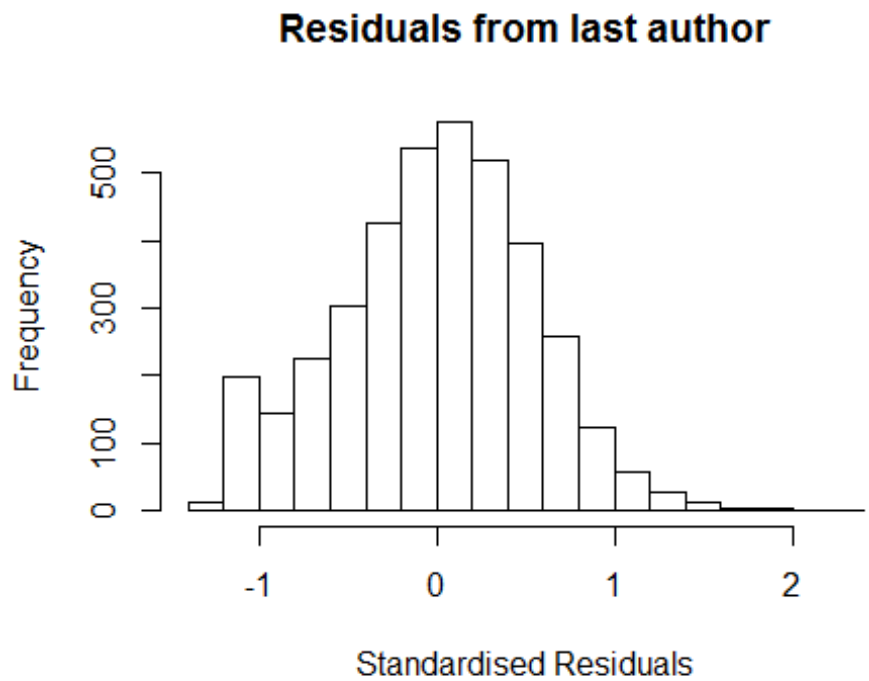

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.2399 -0.3575  0.0257  0.3662  2.3150
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)          1.13298    0.04329   26.17   <2e-16 ***
## FirstAuthorFemale1 -0.02806    0.02583   -1.09    0.277
## LastAuthorFemale1   0.04392    0.02632    1.67    0.095 .
## Year1997            0.01288    0.06253    0.21    0.837
## Year1998            0.07108    0.06155    1.15    0.248
## Year1999            0.00671    0.05852    0.11    0.909
## Year2000            0.01304    0.05741    0.23    0.820
## Year2001            0.06296    0.07425    0.85    0.397
## Year2002            0.06064    0.05345    1.13    0.257
## Year2003            0.02631    0.05341    0.49    0.622
## Year2004            0.03283    0.05411    0.61    0.544
## Year2005            0.03267    0.05829    0.56    0.575
```

```
## Year2006            0.04839    0.05297    0.91    0.361
## Year2007           -0.10297    0.06017   -1.71    0.087 .
## Year2008            0.02385    0.05527    0.43    0.666
## Year2009            0.00944    0.05522    0.17    0.864
## Year2010            0.05955    0.05563    1.07    0.285
## Year2011           -0.02562    0.05584   -0.46    0.646
## Year2012            0.00598    0.05864    0.10    0.919
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.537
## Multiple R-squared:  0.00623,    Adjusted R-squared:  0.00153
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  316 weights are ~= 1. The remaining 3508 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0235  0.8640  0.9490  0.9040  0.9860  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          2.62e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.021  1           1.010
## Year              1.021 16           1.001
```




```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.2083 -0.3554  0.0238  0.3661  2.3127
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.136933   0.043150   26.35   <2e-16 ***
## FirstAuthorFemale1  0.000553   0.019884    0.03    0.978
## Year1997            0.012417   0.062509    0.20    0.843
## Year1998            0.071369   0.061608    1.16    0.247
## Year1999            0.007373   0.058496    0.13    0.900
## Year2000            0.014264   0.057247    0.25    0.803
## Year2001            0.062034   0.074540    0.83    0.405
## Year2002            0.061447   0.053430    1.15    0.250
## Year2003            0.025075   0.053254    0.47    0.638
## Year2004            0.034107   0.053959    0.63    0.527
## Year2005            0.032945   0.058175    0.57    0.571
## Year2006            0.047990   0.052888    0.91    0.364
```

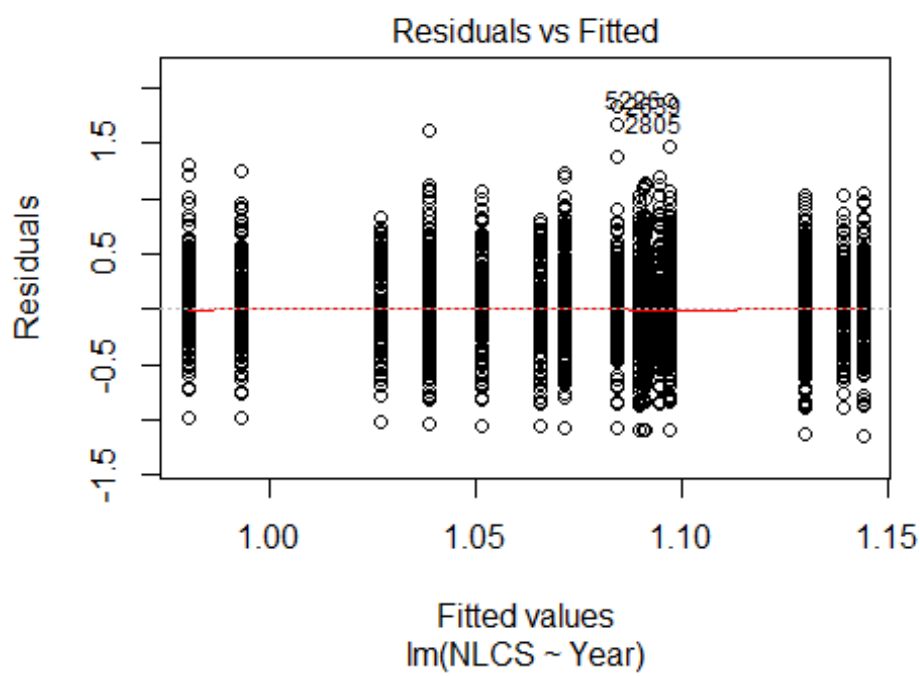
```
## Year2007           -0.102330   0.059986   -1.71    0.088 .
## Year2008            0.025744   0.055176    0.47    0.641
## Year2009            0.008085   0.055100    0.15    0.883
## Year2010            0.059877   0.055516    1.08    0.281
## Year2011           -0.025110   0.055763   -0.45    0.653
## Year2012            0.004364   0.058614    0.07    0.941
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.538
## Multiple R-squared:  0.00551,    Adjusted R-squared:  0.00107
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  321 weights are ~= 1. The remaining 3503 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0247  0.8630  0.9490  0.9040  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          2.62e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##                   GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.017  1           1.008
## Year             1.017 16           1.001
```



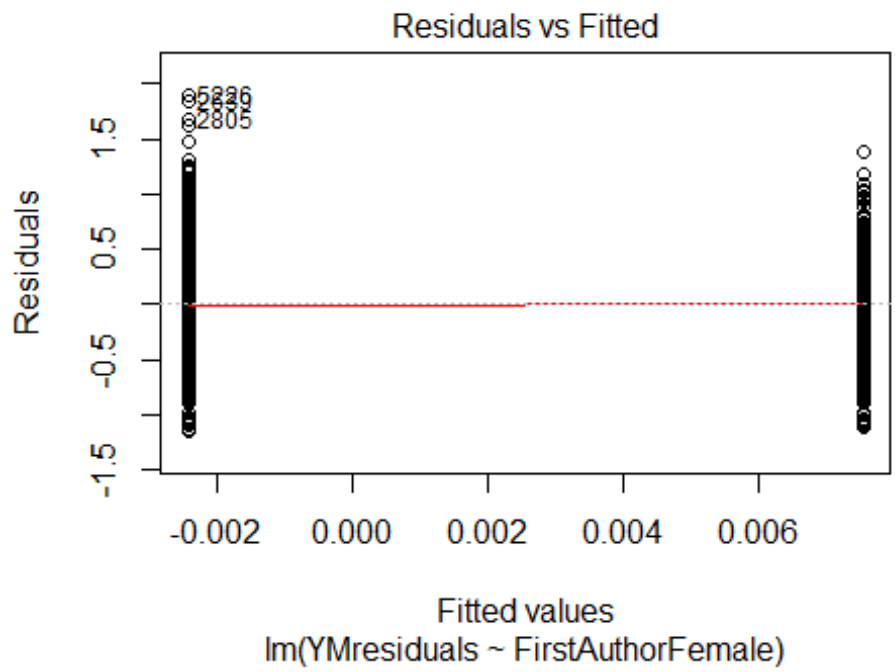

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.2178 -0.3566  0.0273  0.3668  2.3197
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.13012    0.04324   26.14   <2e-16 ***
## LastAuthorFemale1   0.02520    0.02032    1.24    0.215
## Year1997            0.01388    0.06259    0.22    0.825
## Year1998            0.07182    0.06168    1.16    0.244
## Year1999            0.00714    0.05854    0.12    0.903
## Year2000            0.01276    0.05739    0.22    0.824
## Year2001            0.06243    0.07449    0.84    0.402
## Year2002            0.06138    0.05355    1.15    0.252
## Year2003            0.02556    0.05343    0.48    0.632
## Year2004            0.03287    0.05414    0.61    0.544
## Year2005            0.03302    0.05831    0.57    0.571
## Year2006            0.04765    0.05300    0.90    0.369
```

```
## Year2007          -0.10373    0.06020   -1.72    0.085 .
## Year2008           0.02389    0.05537    0.43    0.666
## Year2009           0.00915    0.05527    0.17    0.868
## Year2010           0.05869    0.05571    1.05    0.292
## Year2011          -0.02753    0.05592   -0.49    0.622
## Year2012           0.00413    0.05869    0.07    0.944
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.537
## Multiple R-squared:  0.00592,    Adjusted R-squared:  0.00148
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  325 weights are ~= 1. The remaining 3499 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0225  0.8640  0.9490  0.9030  0.9860  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          2.62e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Sample size for the above analysis:  3824"
## [1] ""
## [1] ""
## [1] "####################################"
## [1] "Analysis of AJSC 2302"
## [1] "####################################"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  241  225  229  194  199  257  216  200  242  285  240  313  279  315  358
## 2011 2012
##  419  427
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  158  163  144  126  110  179  153  129  160  180  156  205  179  211  260
## 2011 2012
```

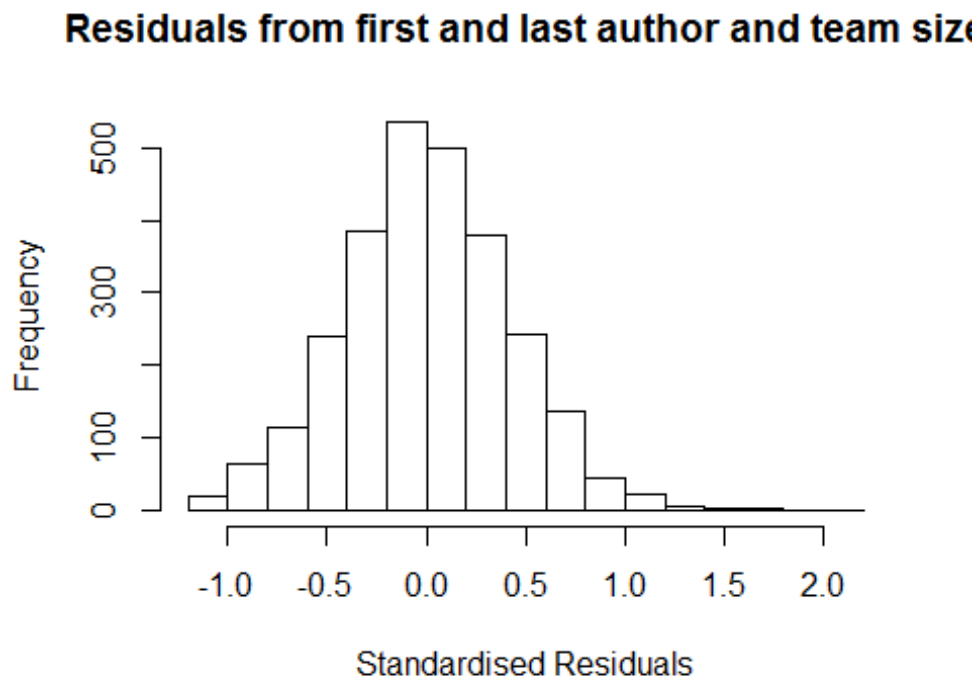
```
##  299  307
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  141  149  123  110   98  162  139  116  144  152  131  175  159  179  221
## 2011 2012
##  243  256
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
##  Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 18, df = 16, p-value = 0.3
```




```
##
##  Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 0.59, df = 1, p-value = 0.4
```

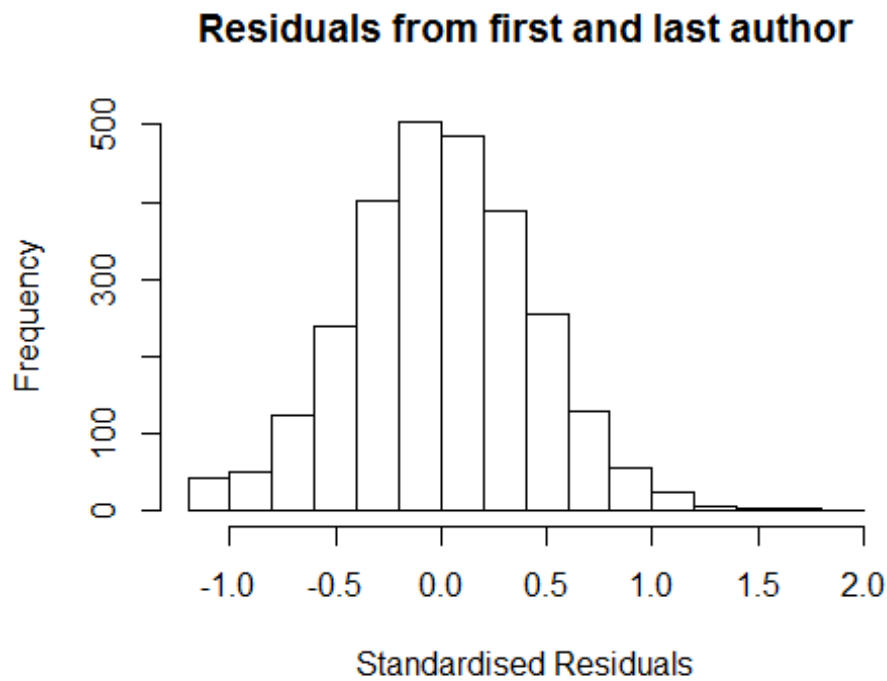



```
## [1] "Female first author team size 2018 geometric mean: 3.73417339341682"
## [1] "Male first author team size 2018 geometric mean: 3.30685699915333"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 6800, p-value = 0.1
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 3.68718735918837"
## [1] "Male last author team size 2018 geometric mean: 3.35881121420504"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 5600, p-value = 0.2
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##                     GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.086  1           1.042
## LastAuthorFemale  1.078  1           1.038
## UniqueAuthors     1.160  4           1.019
## Year              1.224 16           1.006
```



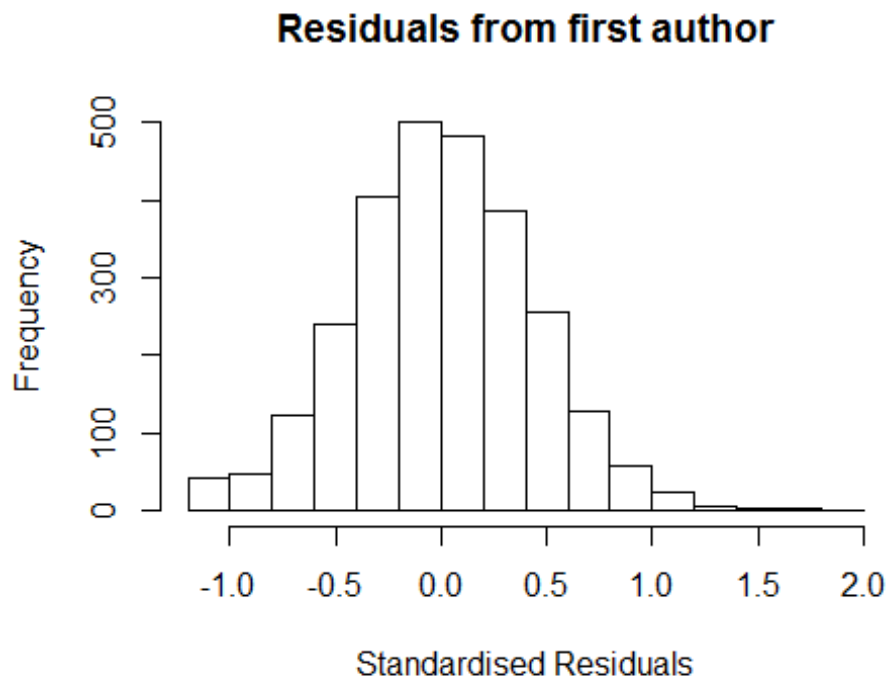

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
UniqueAuthors +
##     Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##     k.max = 1000))
##  \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.17616 -0.26481 -0.00265  0.27848  2.11925
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)         0.992904   0.042444   23.39  < 2e-16 ***
## FirstAuthorFemale1 -0.003116   0.019913   -0.16   0.8757
## LastAuthorFemale1   0.000795   0.022173    0.04   0.9714
## UniqueAuthors2      0.178182   0.025956    6.86  8.3e-12 ***
## UniqueAuthors3      0.216300   0.026165    8.27  < 2e-16 ***
## UniqueAuthors4      0.237084   0.029802    7.96  2.6e-15 ***
## UniqueAuthors5      0.306828   0.028373   10.81  < 2e-16 ***
## Year1997           -0.071255   0.053835   -1.32   0.1858
## Year1998           -0.178407   0.056627   -3.15   0.0016 **
## Year1999           -0.117450   0.057238   -2.05   0.0403 *
```

```
## Year2000           -0.094459   0.056594   -1.67   0.0952 .
## Year2001           -0.168945   0.052494   -3.22   0.0013 **
## Year2002           -0.127280   0.053039   -2.40   0.0165 *
## Year2003           -0.042573   0.052180   -0.82   0.4146
## Year2004           -0.083860   0.049866   -1.68   0.0927 .
## Year2005           -0.065600   0.050774   -1.29   0.1965
## Year2006           -0.134901   0.051844   -2.60   0.0093 **
## Year2007           -0.114571   0.048257   -2.37   0.0177 *
## Year2008           -0.159260   0.054740   -2.91   0.0037 **
## Year2009           -0.132919   0.048783   -2.72   0.0065 **
## Year2010           -0.089213   0.048452   -1.84   0.0657 .
## Year2011           -0.123574   0.047585   -2.60   0.0095 **
## Year2012           -0.124155   0.047981   -2.59   0.0097 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.409
## Multiple R-squared:  0.0585, Adjusted R-squared:  0.0507
## Convergence in 10 IRWLS iterations
##
## Robustness weights:
##  observation 2640 is an outlier with |weight| = 0 ( < 3.7e-05);
##  233 weights are ~= 1. The remaining 2464 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0545  0.8700  0.9500  0.9040  0.9860  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          3.71e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.069  1           1.034
## LastAuthorFemale  1.060  1           1.030
## Year              1.067 16           1.002
```



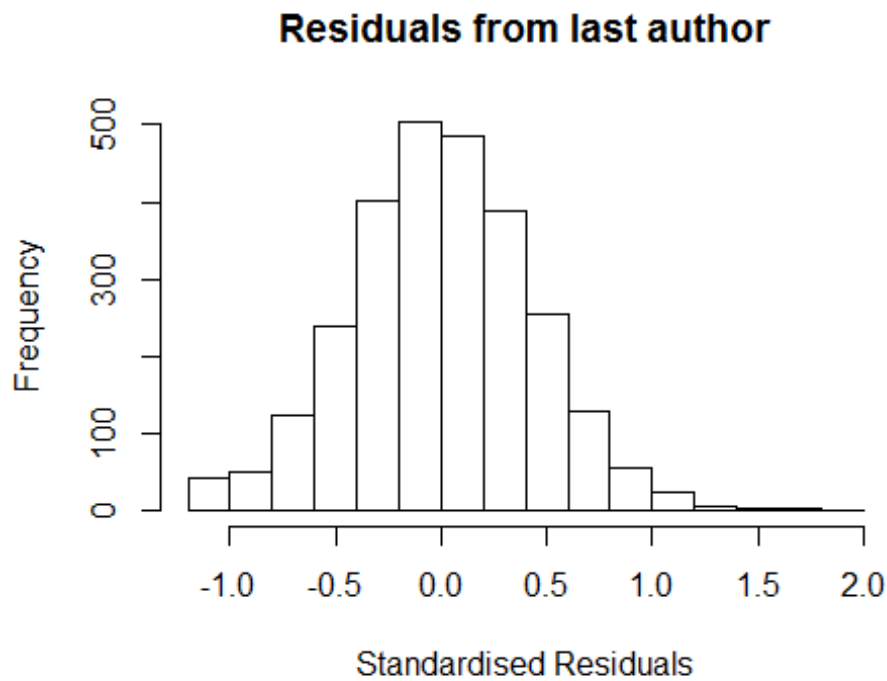

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.14142 -0.28412 -0.00236  0.28563  1.91784
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.141417   0.038181   29.89  < 2e-16 ***
## FirstAuthorFemale1  0.016230   0.020191    0.80  0.42159
## LastAuthorFemale1   0.008798   0.022564    0.39  0.69662
## Year1997           -0.057678   0.054929   -1.05  0.29379
## Year1998           -0.184257   0.056104   -3.28  0.00104 **
## Year1999           -0.098886   0.057714   -1.71  0.08676 .
## Year2000           -0.080337   0.058510   -1.37  0.16985
## Year2001           -0.178896   0.053990   -3.31  0.00093 ***
## Year2002           -0.099731   0.054286   -1.84  0.06630 .
## Year2003           -0.000524   0.052424   -0.01  0.99203
## Year2004           -0.070777   0.051159   -1.38  0.16664
## Year2005           -0.038880   0.051471   -0.76  0.45008
```

```
## Year2006           -0.087412   0.052004   -1.68  0.09290 .
## Year2007           -0.078914   0.048074   -1.64  0.10081
## Year2008           -0.128201   0.054063   -2.37  0.01779 *
## Year2009           -0.084473   0.049174   -1.72  0.08594 .
## Year2010           -0.036655   0.048522   -0.76  0.45005
## Year2011           -0.082957   0.048116   -1.72  0.08480 .
## Year2012           -0.071256   0.048091   -1.48  0.13854
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.421
## Multiple R-squared:  0.012,  Adjusted R-squared:  0.00541
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  252 weights are ~= 1. The remaining 2446 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0031  0.8670  0.9500  0.9040  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          3.71e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.039  1           1.019
## Year              1.039 16           1.001
```




```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.14227 -0.28551 -0.00359  0.28604  1.91663
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.14e+00   3.82e-02   29.87  < 2e-16 ***
## FirstAuthorFemale1  1.76e-02   1.99e-02    0.89  0.37578
## Year1997           -5.76e-02   5.50e-02   -1.05  0.29471
## Year1998           -1.84e-01   5.61e-02   -3.28  0.00106 **
## Year1999           -9.86e-02   5.78e-02   -1.71  0.08783 .
## Year2000           -8.05e-02   5.86e-02   -1.37  0.16953
## Year2001           -1.78e-01   5.38e-02   -3.31  0.00095 ***
## Year2002           -9.90e-02   5.43e-02   -1.82  0.06838 .
## Year2003            3.59e-05   5.24e-02    0.00  0.99945
## Year2004           -7.08e-02   5.12e-02   -1.38  0.16697
## Year2005           -3.85e-02   5.15e-02   -0.75  0.45488
## Year2006           -8.74e-02   5.20e-02   -1.68  0.09316 .
```

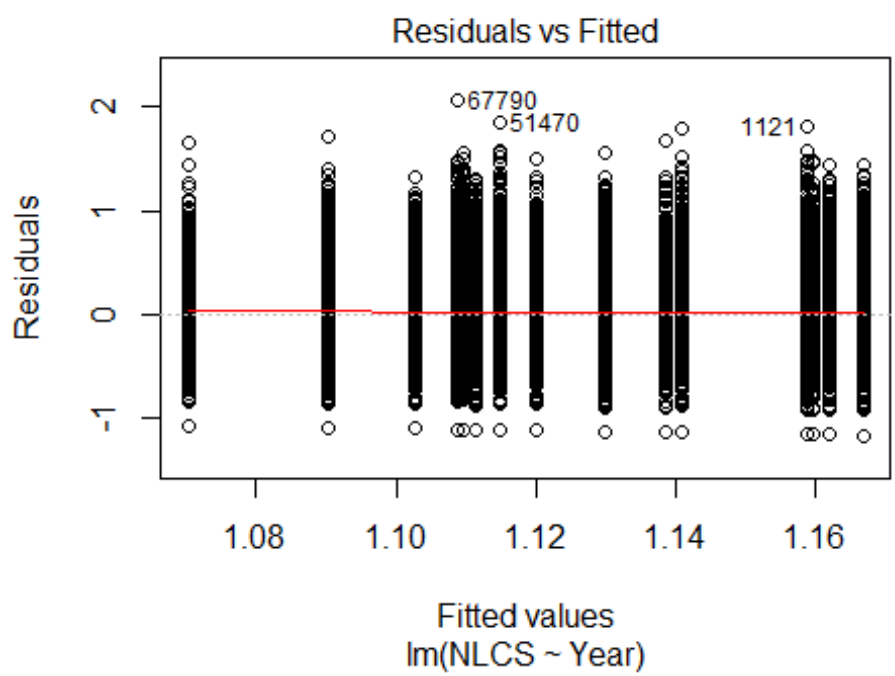
```
## Year2007           -7.83e-02   4.80e-02   -1.63  0.10270
## Year2008           -1.28e-01   5.40e-02   -2.36  0.01831 *
## Year2009           -8.38e-02   4.91e-02   -1.70  0.08832 .
## Year2010           -3.58e-02   4.84e-02   -0.74  0.46027
## Year2011           -8.24e-02   4.81e-02   -1.71  0.08687 .
## Year2012           -7.09e-02   4.81e-02   -1.47  0.14049
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.421
## Multiple R-squared:  0.012,  Adjusted R-squared:  0.00572
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  249 weights are ~= 1. The remaining 2449 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0031  0.8670  0.9500  0.9040  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          3.71e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##                   GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.031  1           1.015
## Year             1.031 16           1.001
```



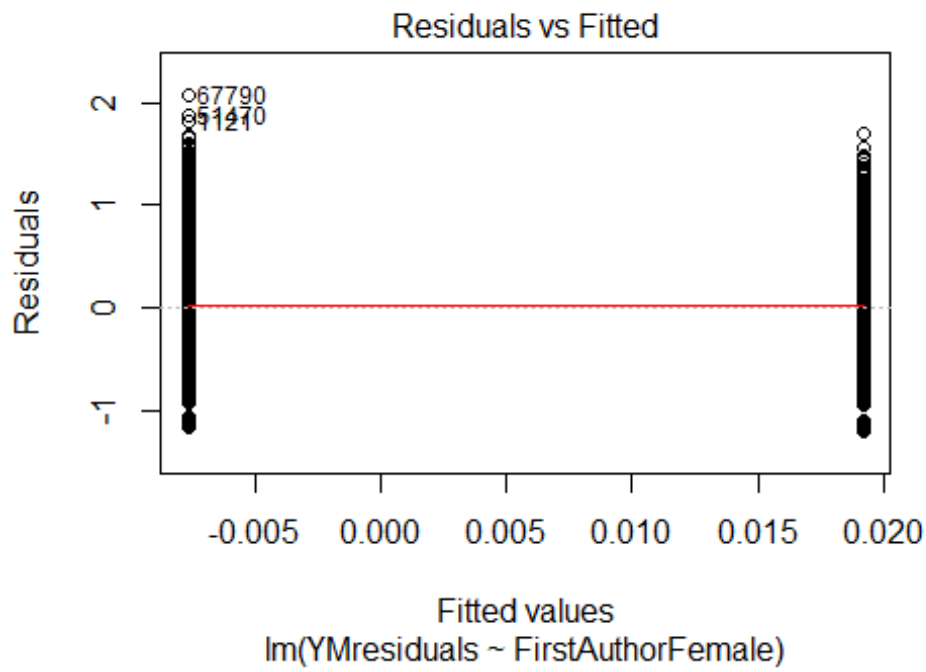

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.14236 -0.28605 -0.00141  0.28326  1.91354
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)        1.142359   0.038217   29.89   <2e-16 ***
## LastAuthorFemale1  0.011896   0.022198    0.54   0.5921
## Year1997          -0.056142   0.054768   -1.03   0.3054
## Year1998          -0.183017   0.056043   -3.27   0.0011 **
## Year1999          -0.096622   0.057480   -1.68   0.0929 .
## Year2000          -0.077893   0.058490   -1.33   0.1831
## Year2001          -0.177011   0.054067   -3.27   0.0011 **
## Year2002          -0.096988   0.054054   -1.79   0.0729 .
## Year2003           0.000682   0.052417    0.01   0.9896
## Year2004          -0.068307   0.051081   -1.34   0.1813
## Year2005          -0.036216   0.051226   -0.71   0.4796
## Year2006          -0.083502   0.051763   -1.61   0.1068
```

```
## Year2007          -0.075928   0.047930   -1.58   0.1133
## Year2008          -0.126012   0.054012   -2.33   0.0197 *
## Year2009          -0.082366   0.049008   -1.68   0.0929 .
## Year2010          -0.033452   0.048325   -0.69   0.4889
## Year2011          -0.079186   0.047813   -1.66   0.0978 .
## Year2012          -0.067895   0.047826   -1.42   0.1558
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.421
## Multiple R-squared:  0.0118, Adjusted R-squared:  0.00553
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  260 weights are ~= 1. The remaining 2438 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0035  0.8660  0.9490  0.9030  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          3.71e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Sample size for the above analysis:  2698"
## [1] ""
## [1] ""
## [1] "####################################"
## [1] "Analysis of AJSC 2303"
## [1] "####################################"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 3006 3322 3343 3378 3172 3496 3533 3622 3493 3771 4190 4568 4465 4159 4247
## 2011 2012
## 4239 4431
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 1923 2086 2160 2194 1698 1615 2502 2470 2446 2705 3046 3380 3232 3033 3091
## 2011 2012
```

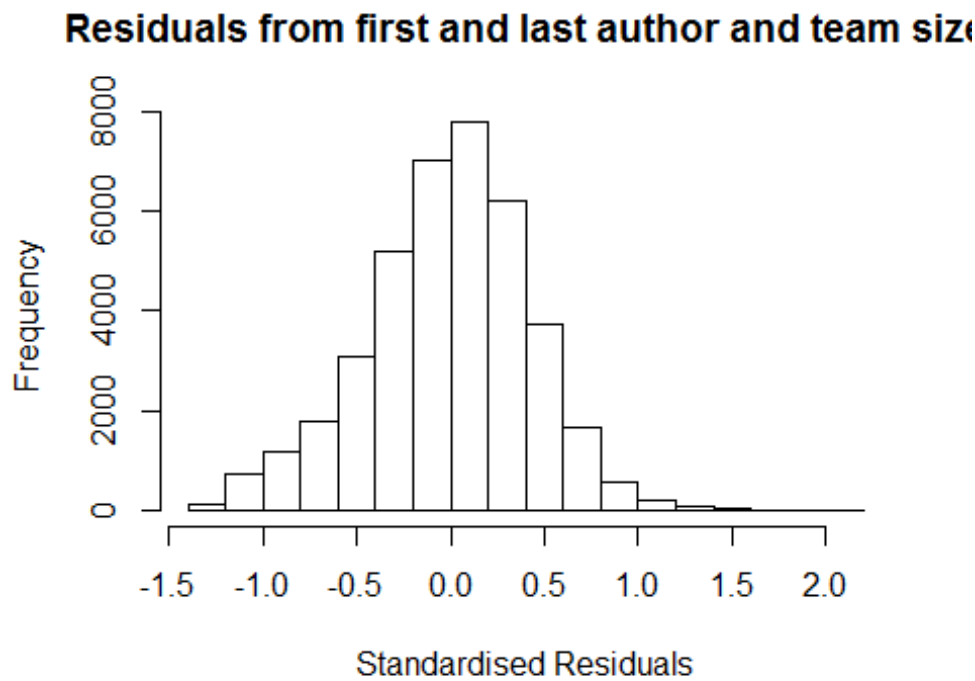
```
## 3118 3254
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 1754 1921 1956 1990 1534 1409 2208 2206 2186 2424 2704 2998 2865 2710 2742
## 2011 2012
## 2756 2924
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
##  Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 110, df = 16, p-value = 2e-16
```




```
##
##  Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 24, df = 1, p-value = 1e-06
```

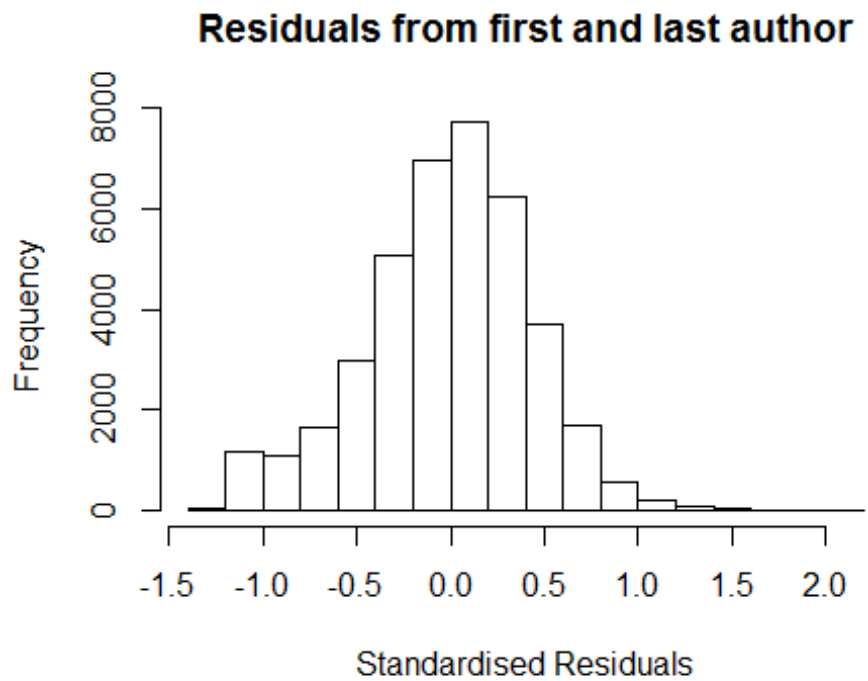



```
## [1] "Female first author team size 2018 geometric mean: 3.5647673217587"
## [1] "Male first author team size 2018 geometric mean: 3.38334514000622"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 1200000, p-value = 0.06
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 3.52638113768845"
## [1] "Male last author team size 2018 geometric mean: 3.42833029072316"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 980000, p-value = 0.2
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##                     GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.045  1           1.022
## LastAuthorFemale  1.028  1           1.014
## UniqueAuthors     1.055  4           1.007
## Year              1.076 16           1.002
```



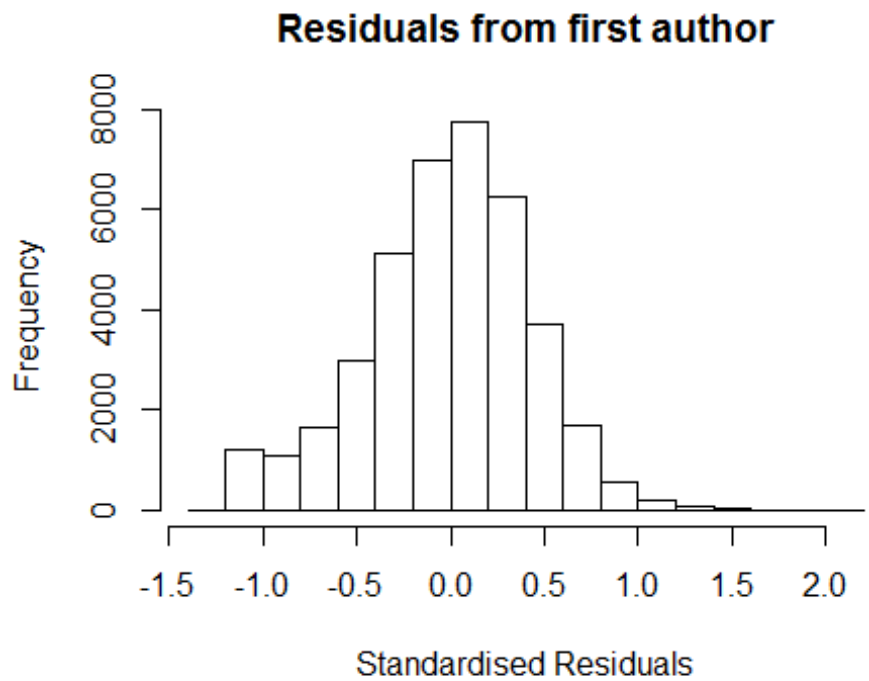

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
UniqueAuthors +
##     Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##     k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.3245 -0.2746  0.0157  0.2756  2.1044
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.03230    0.01285   80.31  < 2e-16 ***
## FirstAuthorFemale1  0.02418    0.00485    4.98  6.3e-07 ***
## LastAuthorFemale1   0.00913    0.00548    1.67  0.09568 .
## UniqueAuthors2      0.11801    0.00794   14.86  < 2e-16 ***
## UniqueAuthors3      0.14662    0.00812   18.05  < 2e-16 ***
## UniqueAuthors4      0.19098    0.00880   21.72  < 2e-16 ***
## UniqueAuthors5      0.26681    0.00858   31.09  < 2e-16 ***
## Year1997            0.02535    0.01515    1.67  0.09423 .
## Year1998            0.00769    0.01523    0.50  0.61374
## Year1999           -0.02547    0.01533   -1.66  0.09669 .
```

```
## Year2000           -0.00032    0.01533   -0.02  0.98333
## Year2001           -0.01965    0.01629   -1.21  0.22766
## Year2002           -0.06945    0.01491   -4.66  3.2e-06 ***
## Year2003           -0.05270    0.01445   -3.65  0.00027 ***
## Year2004           -0.09854    0.01483   -6.64  3.1e-11 ***
## Year2005           -0.07833    0.01431   -5.47  4.5e-08 ***
## Year2006           -0.06569    0.01391   -4.72  2.4e-06 ***
## Year2007           -0.06803    0.01362   -5.00  5.9e-07 ***
## Year2008           -0.07044    0.01396   -5.05  4.5e-07 ***
## Year2009           -0.05044    0.01409   -3.58  0.00034 ***
## Year2010           -0.06775    0.01412   -4.80  1.6e-06 ***
## Year2011           -0.08475    0.01430   -5.93  3.1e-09 ***
## Year2012           -0.08155    0.01427   -5.72  1.1e-08 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.407
## Multiple R-squared:  0.0373, Adjusted R-squared:  0.0368
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
##  2 observations c(26656,35675) are outliers with |weight| = 0 ( < 2.5e-
06);
##  3271 weights are ~= 1. The remaining 36014 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0083  0.8650  0.9500  0.8970  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          2.55e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.041  1           1.021
## LastAuthorFemale  1.029  1           1.014
## Year              1.027 16           1.001
```



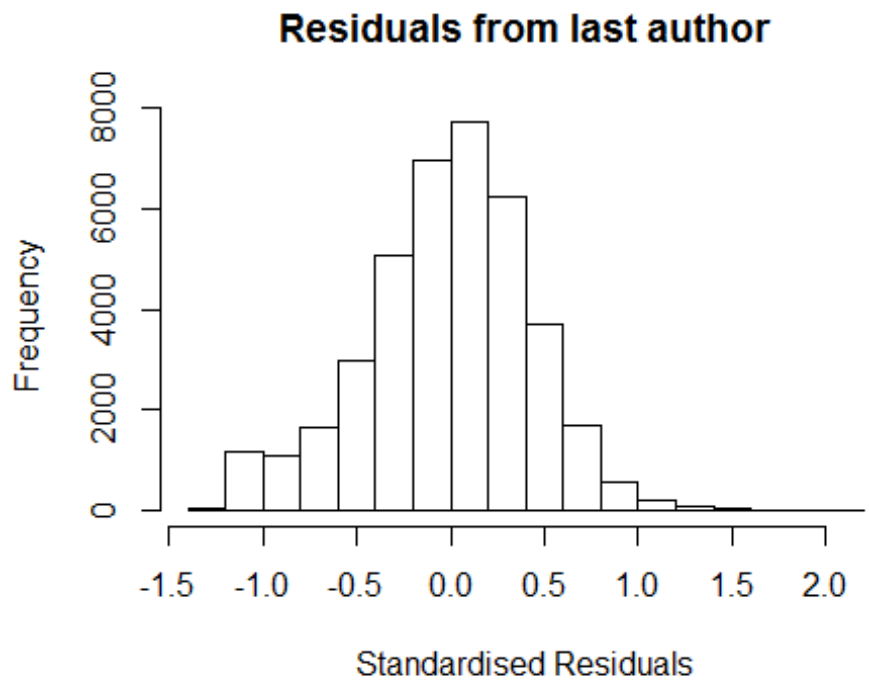

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.2098 -0.2779  0.0156  0.2772  2.0640
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.150966   0.011443  100.58  < 2e-16 ***
## FirstAuthorFemale1  0.030963   0.004902    6.32  2.7e-10 ***
## LastAuthorFemale1   0.006190   0.005557    1.11  0.26529
## Year1997            0.021676   0.015272    1.42  0.15581
## Year1998            0.013161   0.015309    0.86  0.38996
## Year1999           -0.020239   0.015361   -1.32  0.18767
## Year2000            0.008178   0.015359    0.53  0.59444
## Year2001           -0.000976   0.016411   -0.06  0.95260
## Year2002           -0.052196   0.015021   -3.47  0.00051 ***
## Year2003           -0.034455   0.014494   -2.38  0.01745 *
## Year2004           -0.075030   0.014977   -5.01  5.5e-07 ***
## Year2005           -0.053534   0.014364   -3.73  0.00019 ***
```

```
## Year2006           -0.037953   0.013982   -2.71  0.00664 **
## Year2007           -0.040110   0.013682   -2.93  0.00337 **
## Year2008           -0.035930   0.013981   -2.57  0.01017 *
## Year2009           -0.012129   0.014135   -0.86  0.39084
## Year2010           -0.028751   0.014137   -2.03  0.04199 *
## Year2011           -0.044960   0.014426   -3.12  0.00183 **
## Year2012           -0.036364   0.014339   -2.54  0.01122 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.411
## Multiple R-squared:  0.00414,    Adjusted R-squared:  0.00368
## Convergence in 10 IRWLS iterations
##
## Robustness weights:
##  observation 35675 is an outlier with |weight| = 0 ( < 2.5e-06);
##  3256 weights are ~= 1. The remaining 36030 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0045  0.8650  0.9500  0.8960  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          2.55e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                  psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.021  1           1.011
## Year              1.021 16           1.001
```




```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.2053 -0.2779  0.0151  0.2771  2.0629
##
## Coefficients:
##                      Estimate Std. Error t value Pr(>|t|)
## (Intercept)          1.151736   0.011426  100.80  < 2e-16 ***
## FirstAuthorFemale1   0.031972   0.004860    6.58  4.8e-11 ***
## Year1997             0.021634   0.015272    1.42  0.15661
## Year1998             0.013193   0.015310    0.86  0.38885
## Year1999            -0.020220   0.015363   -1.32  0.18813
## Year2000             0.008337   0.015359    0.54  0.58724
## Year2001            -0.000841   0.016412   -0.05  0.95914
## Year2002            -0.052023   0.015021   -3.46  0.00053 ***
## Year2003            -0.034312   0.014494   -2.37  0.01792 *
## Year2004            -0.074894   0.014977   -5.00  5.7e-07 ***
## Year2005            -0.053297   0.014360   -3.71  0.00021 ***
## Year2006            -0.037696   0.013979   -2.70  0.00701 **
```

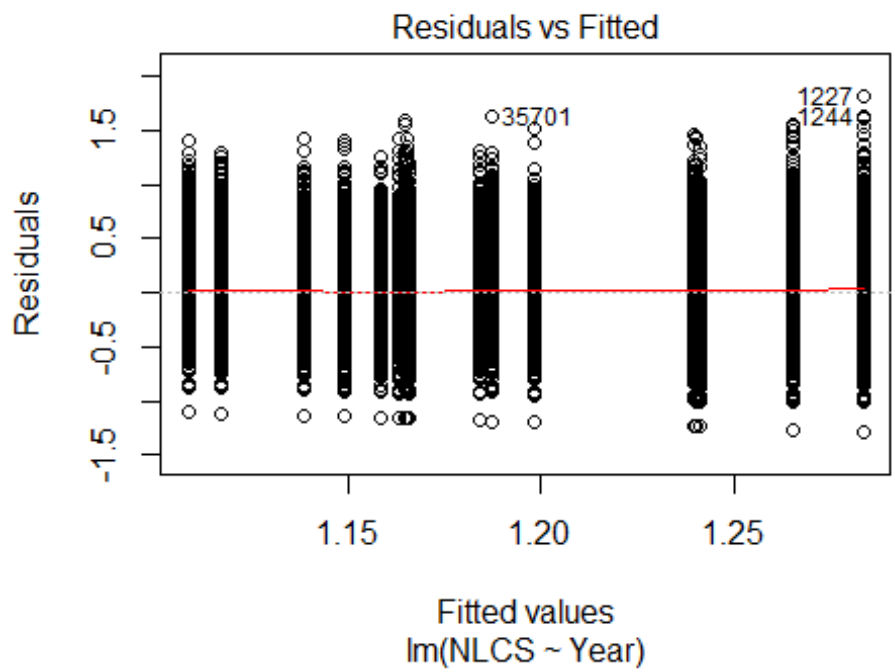
```
## Year2007           -0.039842   0.013678   -2.91  0.00358 **
## Year2008           -0.035595   0.013977   -2.55  0.01088 *
## Year2009           -0.011811   0.014129   -0.84  0.40317
## Year2010           -0.028447   0.014132   -2.01  0.04412 *
## Year2011           -0.044675   0.014423   -3.10  0.00195 **
## Year2012           -0.035877   0.014328   -2.50  0.01229 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.411
## Multiple R-squared:  0.0041, Adjusted R-squared:  0.00367
## Convergence in 10 IRWLS iterations
##
## Robustness weights:
##  observation 35675 is an outlier with |weight| = 0 ( < 2.5e-06);
##  3246 weights are ~= 1. The remaining 36040 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0047  0.8650  0.9500  0.8960  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          2.55e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##                   GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.009  1           1.004
## Year             1.009 16           1.000
```



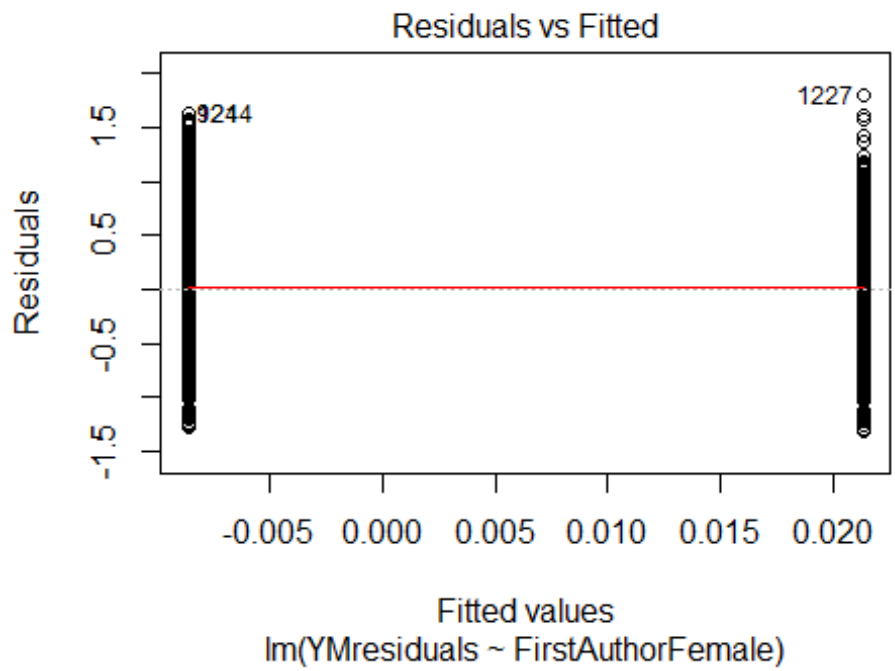

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.1906 -0.2785  0.0167  0.2776  2.0544
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.156153   0.011415  101.28  < 2e-16 ***
## LastAuthorFemale1   0.012480   0.005505    2.27  0.02340 *
## Year1997            0.021993   0.015277    1.44  0.14997
## Year1998            0.013707   0.015310    0.90  0.37064
## Year1999           -0.019485   0.015362   -1.27  0.20466
## Year2000            0.009290   0.015357    0.60  0.54521
## Year2001           -0.000465   0.016412   -0.03  0.97739
## Year2002           -0.050253   0.015025   -3.34  0.00082 ***
## Year2003           -0.032624   0.014492   -2.25  0.02438 *
## Year2004           -0.072575   0.014976   -4.85  1.3e-06 ***
## Year2005           -0.051207   0.014372   -3.56  0.00037 ***
## Year2006           -0.034947   0.013972   -2.50  0.01238 *
```

```
## Year2007          -0.036900   0.013674   -2.70  0.00697 **
## Year2008          -0.032150   0.013961   -2.30  0.02130 *
## Year2009          -0.008092   0.014113   -0.57  0.56640
## Year2010          -0.025019   0.014113   -1.77  0.07628 .
## Year2011          -0.040517   0.014407   -2.81  0.00492 **
## Year2012          -0.032160   0.014319   -2.25  0.02471 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.411
## Multiple R-squared:  0.00309,    Adjusted R-squared:  0.00266
## Convergence in 10 IRWLS iterations
##
## Robustness weights:
##  observation 35675 is an outlier with |weight| = 0 ( < 2.5e-06);
##  3249 weights are ~= 1. The remaining 36037 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0059  0.8650  0.9500  0.8960  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          2.55e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Sample size for the above analysis:  39287"
## [1] ""
## [1] ""
## [1] "####################################"
## [1] "Analysis of AJSC 2304"
## [1] "####################################"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 1655 1646 1695 1813 1687 1748 1849 1825 1777 2060 2077 2133 2128 1998 1876
## 2011 2012
## 2100 2177
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 1069 1059 1079 1183  937  828 1243 1179 1178 1349 1412 1460 1451 1342 1280
```

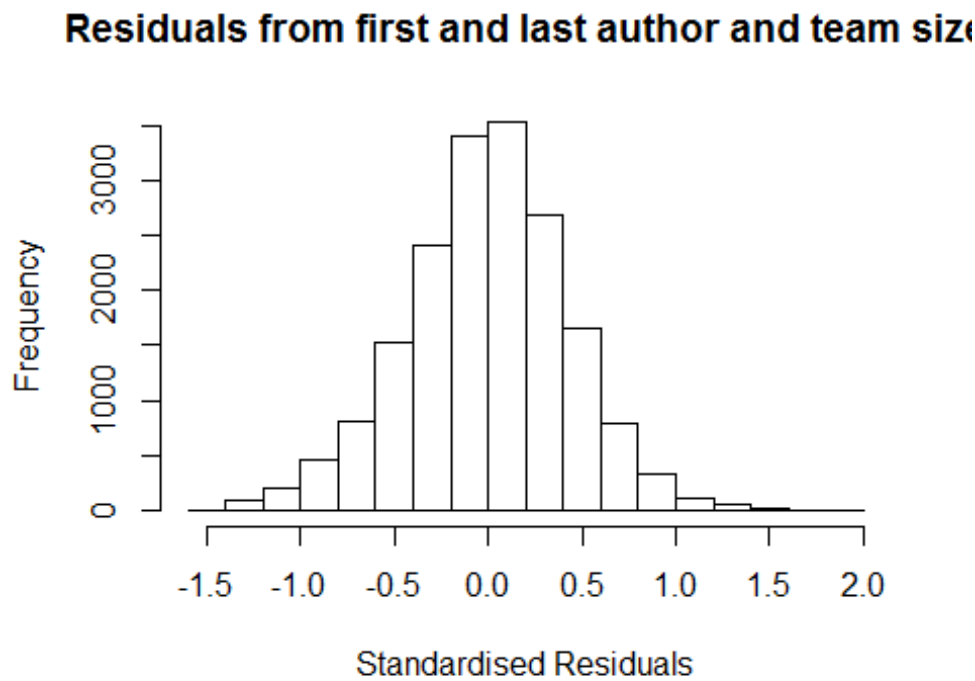
```
## 2011 2012
## 1362 1424
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  936  933  959 1023  817  711 1077 1009 1022 1177 1209 1266 1260 1154 1101
## 2011 2012
## 1192 1219
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
##  Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 270, df = 16, p-value <2e-16
```




```
##
##  Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 67, df = 1, p-value <2e-16
```

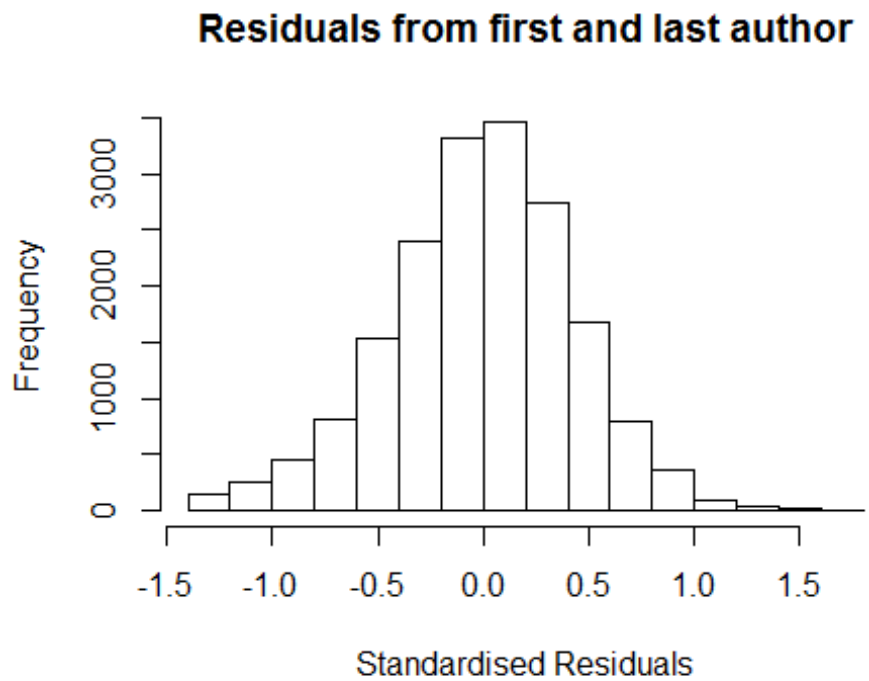



```
## [1] "Female first author team size 2018 geometric mean: 4.2416618458563"
## [1] "Male first author team size 2018 geometric mean: 4.02710157461803"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 2e+05, p-value = 0.3
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 4.05042579474398"
## [1] "Male last author team size 2018 geometric mean: 4.12561706185569"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 170000, p-value = 0.6
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##                     GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.031  1           1.015
## LastAuthorFemale  1.020  1           1.010
## UniqueAuthors     1.051  4           1.006
## Year              1.063 16           1.002
```



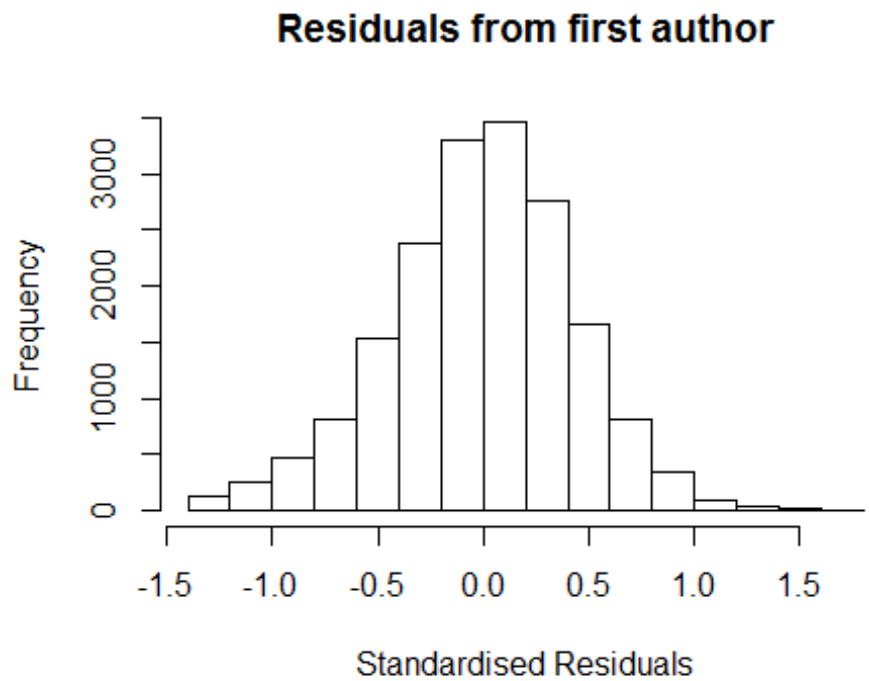

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
UniqueAuthors +
##     Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##     k.max = 1000))
##  \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.42215 -0.27099  0.00736  0.27285  1.98429
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.08504    0.02180   49.78  < 2e-16 ***
## FirstAuthorFemale1  0.01619    0.00694    2.33  0.01978 *
## LastAuthorFemale1   0.01349    0.00820    1.65  0.09980 .
## UniqueAuthors2      0.22461    0.01489   15.08  < 2e-16 ***
## UniqueAuthors3      0.24885    0.01490   16.70  < 2e-16 ***
## UniqueAuthors4      0.28541    0.01527   18.69  < 2e-16 ***
## UniqueAuthors5      0.33711    0.01479   22.79  < 2e-16 ***
## Year1997           -0.02868    0.02376   -1.21  0.22739
## Year1998           -0.06786    0.02343   -2.90  0.00378 **
## Year1999           -0.10364    0.02269   -4.57  4.9e-06 ***
```

```
## Year2000           -0.08466    0.02270   -3.73  0.00019 ***
## Year2001           -0.09515    0.02328   -4.09  4.4e-05 ***
## Year2002           -0.13612    0.02090   -6.51  7.6e-11 ***
## Year2003           -0.16854    0.02125   -7.93  2.3e-15 ***
## Year2004           -0.17581    0.02105   -8.35  < 2e-16 ***
## Year2005           -0.17165    0.02079   -8.26  < 2e-16 ***
## Year2006           -0.21549    0.02125  -10.14  < 2e-16 ***
## Year2007           -0.24591    0.02079  -11.83  < 2e-16 ***
## Year2008           -0.21771    0.02099  -10.37  < 2e-16 ***
## Year2009           -0.19434    0.02149   -9.04  < 2e-16 ***
## Year2010           -0.19100    0.02147   -8.90  < 2e-16 ***
## Year2011           -0.19778    0.02149   -9.20  < 2e-16 ***
## Year2012           -0.16337    0.02160   -7.56  4.1e-14 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.407
## Multiple R-squared:  0.0615, Adjusted R-squared:  0.0604
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
##  observation 656 is an outlier with |weight| = 0 ( < 5.5e-06);
##  1565 weights are ~= 1. The remaining 16499 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0061  0.8650  0.9500  0.8980  0.9860  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          5.54e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.022  1           1.011
## LastAuthorFemale  1.017  1           1.008
## Year              1.020 16           1.001
```



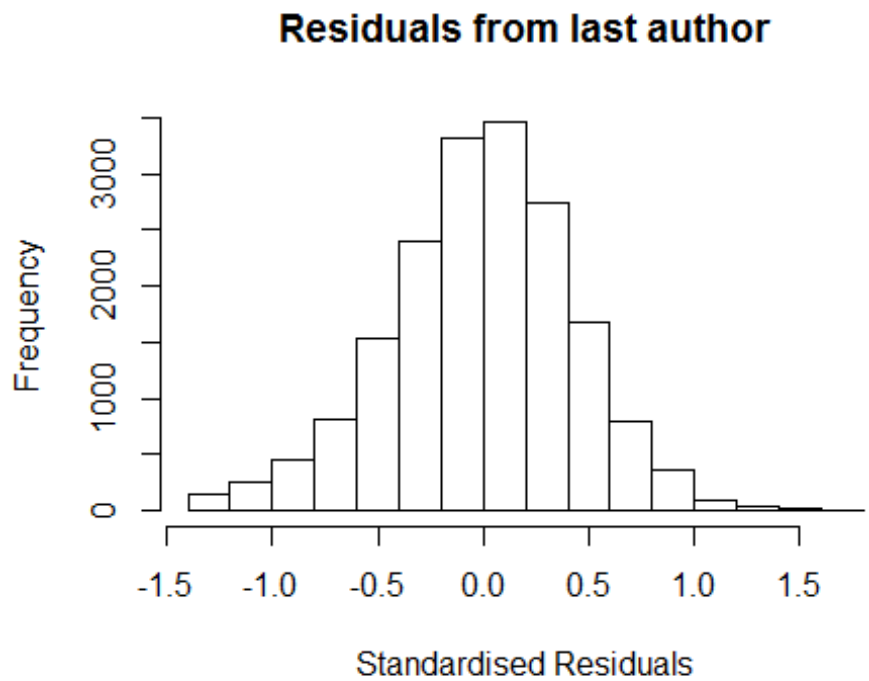

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.35508 -0.27595  0.00862  0.27919  1.74392
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)          1.31064    0.01762   74.40  < 2e-16 ***
## FirstAuthorFemale1   0.03037    0.00701    4.33  1.5e-05 ***
## LastAuthorFemale1    0.01407    0.00830    1.69   0.0901 .
## Year1997            -0.02565    0.02424   -1.06   0.2901
## Year1998            -0.06783    0.02373   -2.86   0.0043 **
## Year1999            -0.09862    0.02306   -4.28  1.9e-05 ***
## Year2000            -0.07410    0.02302   -3.22   0.0013 **
## Year2001            -0.07614    0.02371   -3.21   0.0013 **
## Year2002            -0.12543    0.02122   -5.91  3.5e-09 ***
## Year2003            -0.15667    0.02151   -7.28  3.4e-13 ***
## Year2004            -0.15785    0.02142   -7.37  1.8e-13 ***
## Year2005            -0.15051    0.02118   -7.11  1.2e-12 ***
```

```
## Year2006           -0.19318    0.02158   -8.95  < 2e-16 ***
## Year2007           -0.21740    0.02109  -10.31  < 2e-16 ***
## Year2008           -0.19000    0.02117   -8.98  < 2e-16 ***
## Year2009           -0.16626    0.02182   -7.62  2.7e-14 ***
## Year2010           -0.16323    0.02174   -7.51  6.2e-14 ***
## Year2011           -0.16369    0.02189   -7.48  7.9e-14 ***
## Year2012           -0.12150    0.02185   -5.56  2.7e-08 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.413
## Multiple R-squared:  0.0183, Adjusted R-squared:  0.0173
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
##  1509 weights are ~= 1. The remaining 16556 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0355  0.8660  0.9500  0.8970  0.9860  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          5.54e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.013  1           1.007
## Year              1.013 16           1.000
```




```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.34416 -0.27644  0.00867  0.27841  1.75484
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.31227    0.01760   74.57  < 2e-16 ***
## FirstAuthorFemale1  0.03189    0.00698    4.57  4.9e-06 ***
## Year1997           -0.02552    0.02424   -1.05   0.2925
## Year1998           -0.06754    0.02373   -2.85   0.0044 **
## Year1999           -0.09832    0.02307   -4.26  2.0e-05 ***
## Year2000           -0.07359    0.02301   -3.20   0.0014 **
## Year2001           -0.07570    0.02371   -3.19   0.0014 **
## Year2002           -0.12468    0.02122   -5.88  4.3e-09 ***
## Year2003           -0.15582    0.02150   -7.25  4.4e-13 ***
## Year2004           -0.15686    0.02141   -7.33  2.5e-13 ***
## Year2005           -0.14999    0.02118   -7.08  1.5e-12 ***
## Year2006           -0.19274    0.02159   -8.93  < 2e-16 ***
```

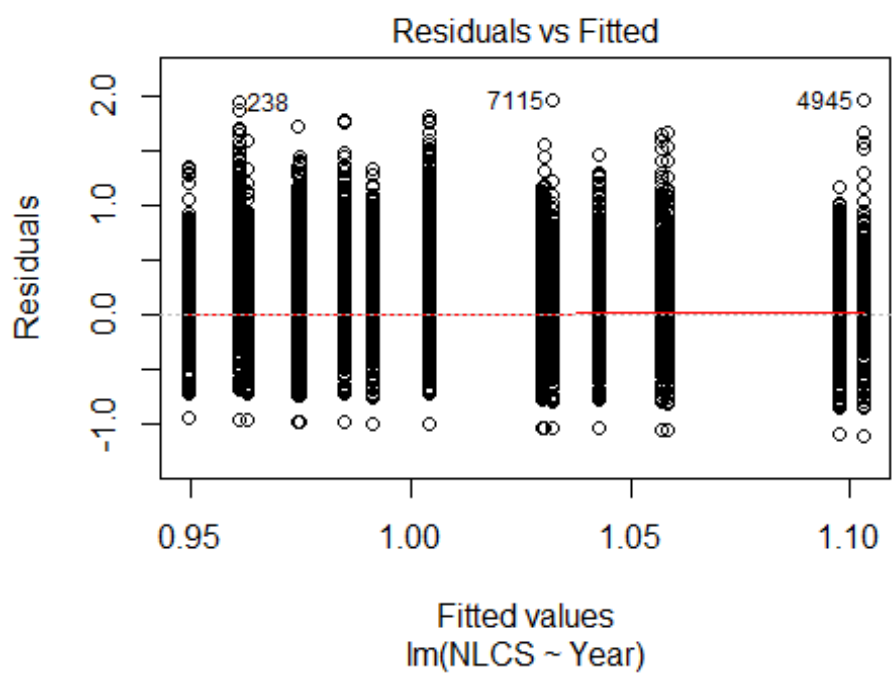
```
## Year2007           -0.21675    0.02109  -10.28  < 2e-16 ***
## Year2008           -0.18894    0.02115   -8.93  < 2e-16 ***
## Year2009           -0.16553    0.02181   -7.59  3.4e-14 ***
## Year2010           -0.16277    0.02173   -7.49  7.3e-14 ***
## Year2011           -0.16262    0.02188   -7.43  1.1e-13 ***
## Year2012           -0.12056    0.02184   -5.52  3.4e-08 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.413
## Multiple R-squared:  0.0181, Adjusted R-squared:  0.0172
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
##  1519 weights are ~= 1. The remaining 16546 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.032   0.866   0.950   0.897   0.986   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          5.54e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##                   GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.008  1           1.004
## Year             1.008 16           1.000
```



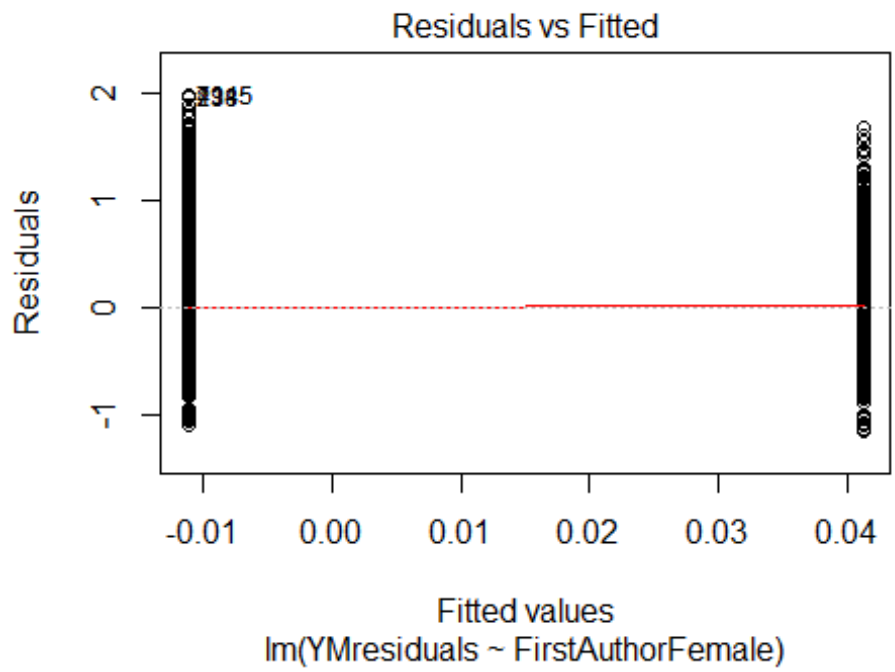

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.33456 -0.27623  0.00887  0.27854  1.76444
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.31610    0.01751   75.17  < 2e-16 ***
## LastAuthorFemale1   0.01846    0.00825    2.24   0.0254 *
## Year1997           -0.02428    0.02422   -1.00   0.3161
## Year1998           -0.06690    0.02369   -2.82   0.0047 **
## Year1999           -0.09720    0.02303   -4.22  2.4e-05 ***
## Year2000           -0.07198    0.02297   -3.13   0.0017 **
## Year2001           -0.07474    0.02369   -3.15   0.0016 **
## Year2002           -0.12335    0.02117   -5.83  5.8e-09 ***
## Year2003           -0.15434    0.02147   -7.19  6.8e-13 ***
## Year2004           -0.15565    0.02137   -7.28  3.4e-13 ***
## Year2005           -0.14834    0.02113   -7.02  2.3e-12 ***
## Year2006           -0.18974    0.02153   -8.81  < 2e-16 ***
```

```
## Year2007          -0.21391    0.02106  -10.16  < 2e-16 ***
## Year2008          -0.18607    0.02111   -8.81  < 2e-16 ***
## Year2009          -0.16211    0.02176   -7.45  9.9e-14 ***
## Year2010          -0.15833    0.02169   -7.30  3.0e-13 ***
## Year2011          -0.15922    0.02182   -7.30  3.0e-13 ***
## Year2012          -0.11684    0.02180   -5.36  8.4e-08 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.413
## Multiple R-squared:  0.0173, Adjusted R-squared:  0.0164
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
##  1530 weights are ~= 1. The remaining 16535 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0287  0.8660  0.9510  0.8970  0.9860  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          5.54e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Sample size for the above analysis:  18065"
## [1] ""
## [1] ""
## [1] "####################################"
## [1] "Analysis of AJSC 2305"
## [1] "####################################"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 1001  989  865  835  846  752  710  677  734  947 1053 1008  996 1061 1051
## 2011 2012
## 1155 1129
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  650  627  491  481  445  329  446  406  464  584  666  641  630  670  661
## 2011 2012
```

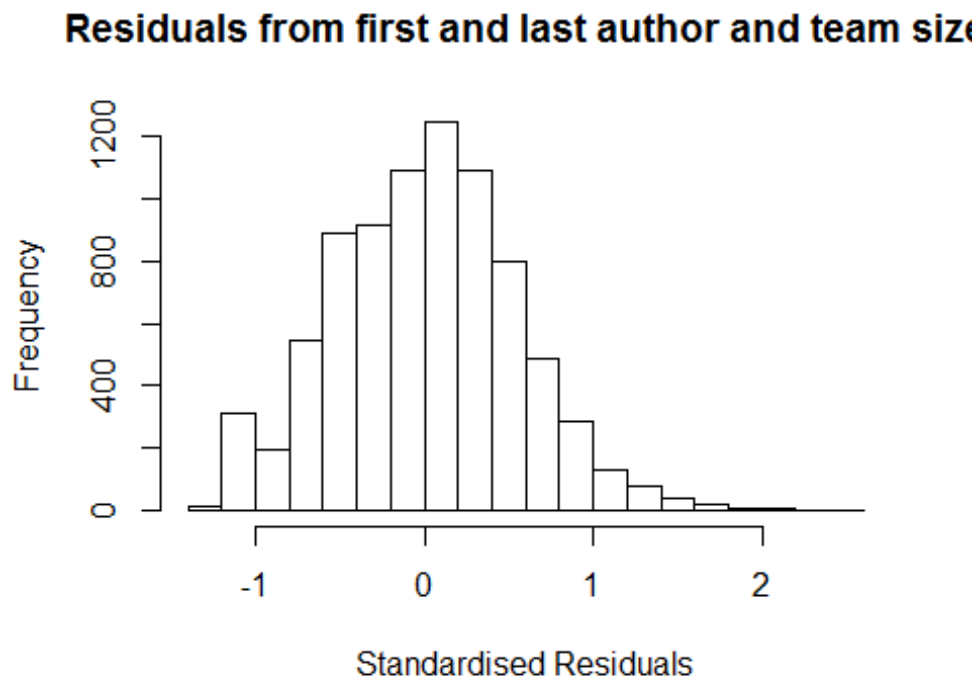
```
##  683  740
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  553  549  429  427  389  290  385  349  387  492  547  546  531  583  537
## 2011 2012
##  556  598
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
##  Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 380, df = 16, p-value <2e-16
```




```
##
##  Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 17, df = 1, p-value = 3e-05
```

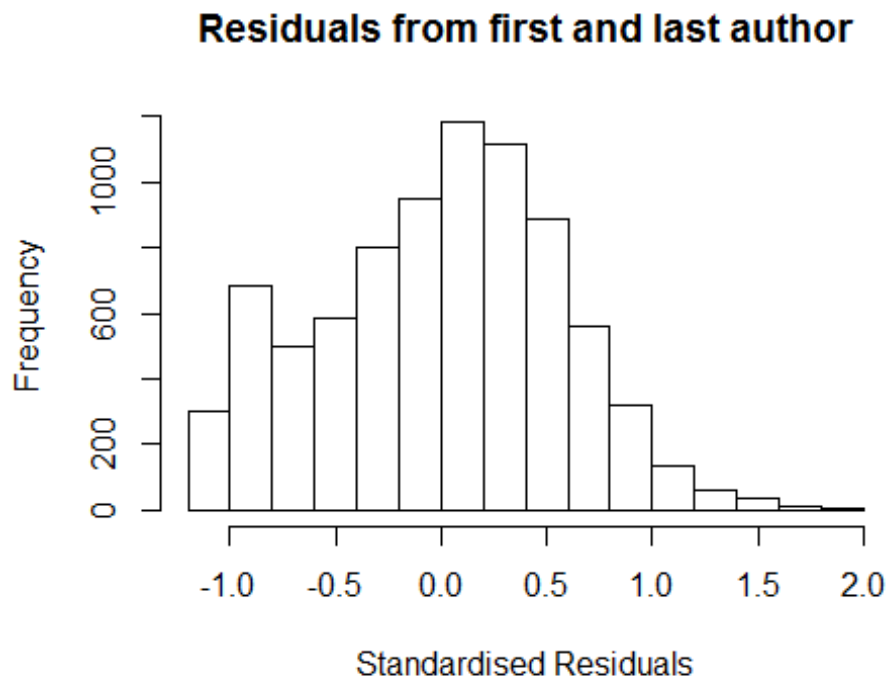



```
## [1] "Female first author team size 2018 geometric mean: 4.15813835261863"
## [1] "Male first author team size 2018 geometric mean: 3.26298187037333"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 69000, p-value = 5e-07
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 4.10858098328586"
## [1] "Male last author team size 2018 geometric mean: 3.35632710548395"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 58000, p-value = 1e-04
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##                     GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.053  1           1.026
## LastAuthorFemale  1.029  1           1.014
## UniqueAuthors     1.148  4           1.017
## Year              1.165 16           1.005
```



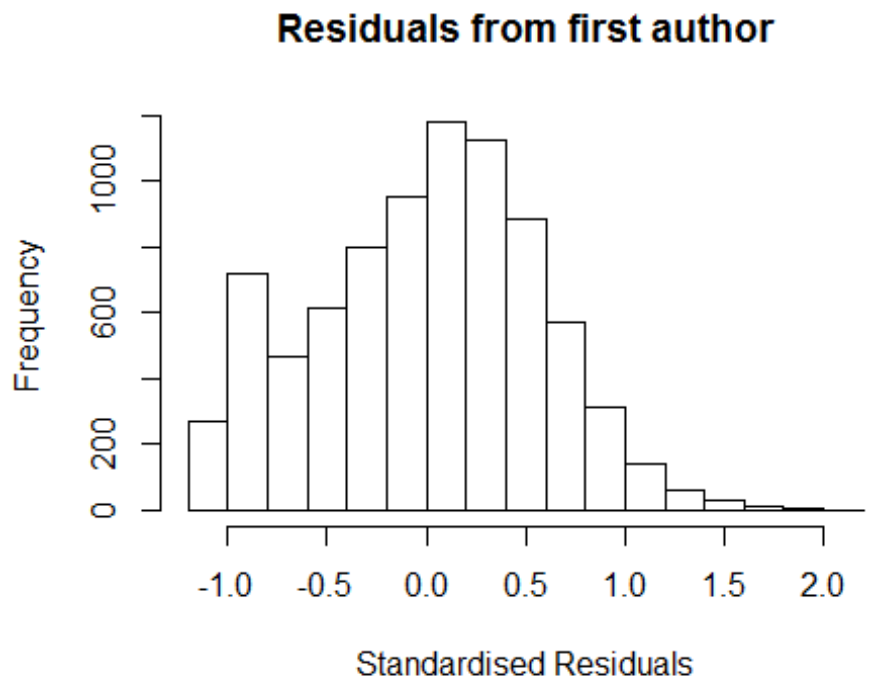

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
UniqueAuthors +
##     Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##     k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.2680 -0.3774  0.0177  0.3638  2.4039
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         0.56559    0.03315   17.06   <2e-16 ***
## FirstAuthorFemale1  0.02368    0.01465    1.62   0.1060
## LastAuthorFemale1  -0.01703    0.01654   -1.03   0.3032
## UniqueAuthors2      0.50549    0.02151   23.50   <2e-16 ***
## UniqueAuthors3      0.52888    0.02184   24.22   <2e-16 ***
## UniqueAuthors4      0.58585    0.02343   25.00   <2e-16 ***
## UniqueAuthors5      0.61982    0.02231   27.78   <2e-16 ***
## Year1997           -0.02212    0.04462   -0.50   0.6201
## Year1998            0.01140    0.04512    0.25   0.8006
## Year1999            0.03940    0.04303    0.92   0.3599
```

```
## Year2000            0.09286    0.04476    2.07   0.0380 *
## Year2001            0.10155    0.04452    2.28   0.0226 *
## Year2002            0.04449    0.04173    1.07   0.2864
## Year2003           -0.00115    0.04154   -0.03   0.9780
## Year2004            0.06428    0.04026    1.60   0.1104
## Year2005           -0.07957    0.03810   -2.09   0.0368 *
## Year2006           -0.07214    0.03910   -1.84   0.0651 .
## Year2007           -0.05450    0.03840   -1.42   0.1558
## Year2008           -0.10119    0.03827   -2.64   0.0082 **
## Year2009           -0.10387    0.03754   -2.77   0.0057 **
## Year2010           -0.10835    0.03888   -2.79   0.0053 **
## Year2011           -0.04653    0.03820   -1.22   0.2232
## Year2012           -0.02393    0.03814   -0.63   0.5303
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.511
## Multiple R-squared:  0.151,  Adjusted R-squared:  0.149
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
##  observation 2378 is an outlier with |weight| = 0 ( < 1.2e-05);
##  652 weights are ~= 1. The remaining 7495 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0252  0.8640  0.9430  0.8980  0.9840  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.23e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.027  1           1.014
## LastAuthorFemale  1.012  1           1.006
## Year              1.034 16           1.001
```



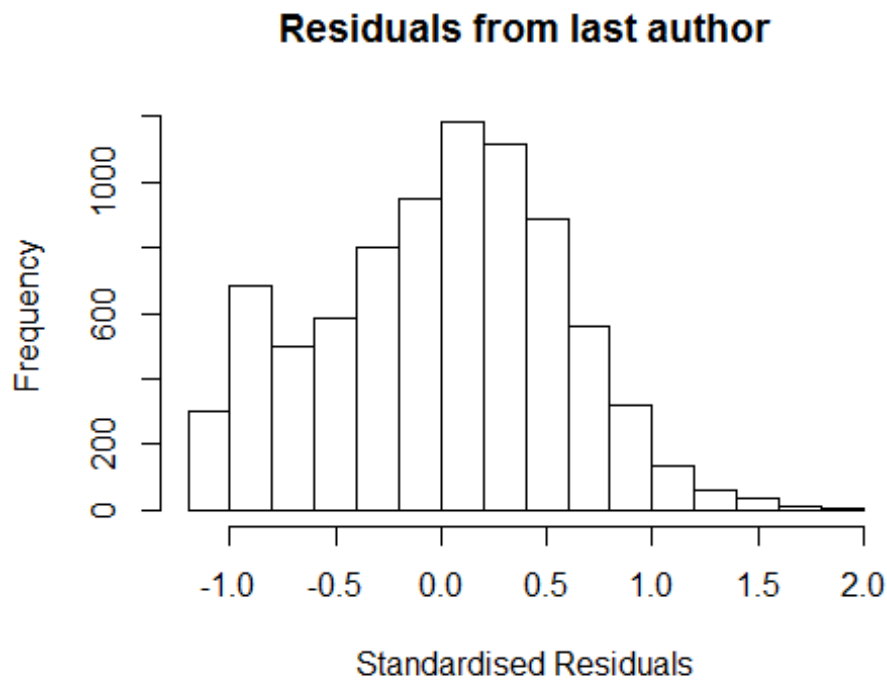

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.1626 -0.4096  0.0423  0.3980  1.9984
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)          0.90558    0.03525   25.69  < 2e-16 ***
## FirstAuthorFemale1   0.05976    0.01540    3.88  0.00011 ***
## LastAuthorFemale1   -0.01482    0.01785   -0.83  0.40646
## Year1997            -0.00449    0.05291   -0.08  0.93234
## Year1998             0.01917    0.05191    0.37  0.71192
## Year1999             0.08648    0.04776    1.81  0.07022 .
## Year2000             0.13997    0.05107    2.74  0.00614 **
## Year2001             0.19723    0.05007    3.94  8.3e-05 ***
## Year2002             0.13938    0.04769    2.92  0.00348 **
## Year2003             0.10377    0.04499    2.31  0.02110 *
## Year2004             0.17384    0.04480    3.88  0.00011 ***
## Year2005             0.03685    0.04247    0.87  0.38550
```

```
## Year2006            0.04685    0.04293    1.09  0.27514
## Year2007            0.06907    0.04219    1.64  0.10168
## Year2008            0.03233    0.04161    0.78  0.43724
## Year2009            0.01542    0.04098    0.38  0.70678
## Year2010            0.03920    0.04290    0.91  0.36084
## Year2011            0.09527    0.04189    2.27  0.02296 *
## Year2012            0.13104    0.04162    3.15  0.00165 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.58
## Multiple R-squared:  0.011,  Adjusted R-squared:  0.0088
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
##  671 weights are ~= 1. The remaining 7477 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.211   0.864   0.948   0.909   0.984   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.23e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.027  1           1.013
## Year              1.027 16           1.001
```




```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.1589 -0.4096  0.0423  0.3974  2.0003
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)          0.90373    0.03535   25.56  < 2e-16 ***
## FirstAuthorFemale1   0.05741    0.01559    3.68  0.00023 ***
## Year1997            -0.00508    0.05313   -0.10  0.92388
## Year1998             0.01977    0.05217    0.38  0.70477
## Year1999             0.08675    0.04790    1.81  0.07017 .
## Year2000             0.14043    0.05120    2.74  0.00611 **
## Year2001             0.19776    0.05019    3.94  8.2e-05 ***
## Year2002             0.13971    0.04780    2.92  0.00348 **
## Year2003             0.10354    0.04507    2.30  0.02161 *
## Year2004             0.17358    0.04488    3.87  0.00011 ***
## Year2005             0.03688    0.04257    0.87  0.38628
## Year2006             0.04733    0.04303    1.10  0.27141
```

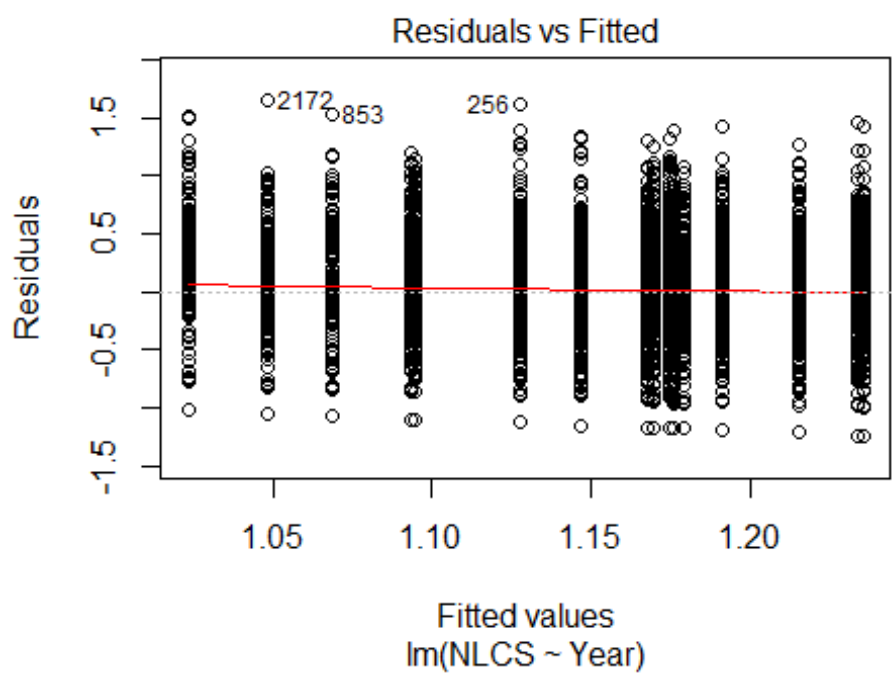
```
## Year2007            0.06921    0.04230    1.64  0.10190
## Year2008            0.03220    0.04170    0.77  0.44005
## Year2009            0.01516    0.04106    0.37  0.71198
## Year2010            0.03886    0.04300    0.90  0.36627
## Year2011            0.09560    0.04200    2.28  0.02287 *
## Year2012            0.13066    0.04170    3.13  0.00173 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.576
## Multiple R-squared:  0.011,  Adjusted R-squared:  0.00893
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
##  662 weights are ~= 1. The remaining 7486 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.203   0.862   0.947   0.908   0.984   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.23e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##                  GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.01  1           1.005
## Year             1.01 16           1.000
```



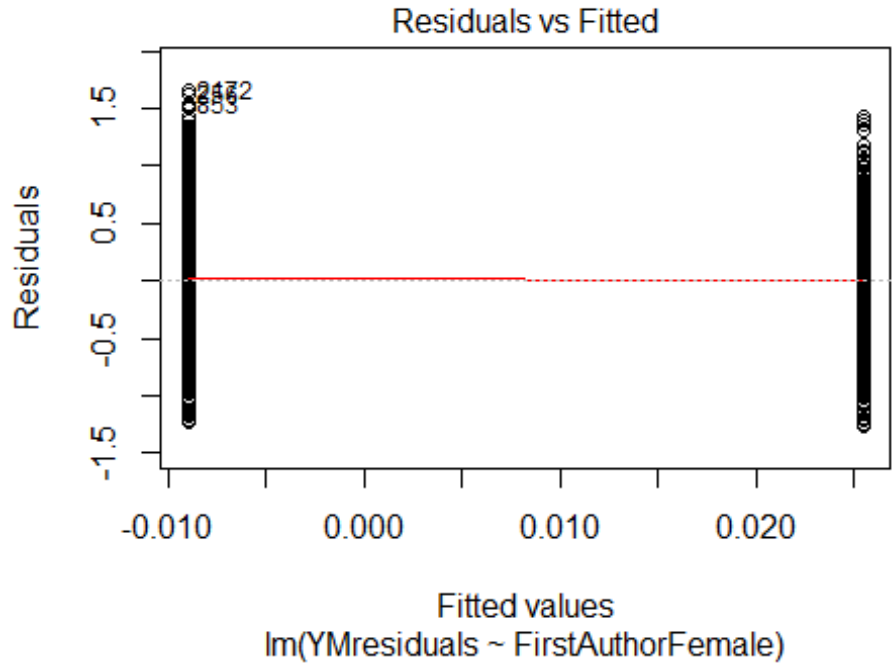

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.1102 -0.4143  0.0434  0.3992  1.9921
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         0.91191    0.03528   25.84  < 2e-16 ***
## LastAuthorFemale1 -0.00223    0.01785   -0.12   0.9007
## Year1997          -0.00359    0.05303   -0.07   0.9460
## Year1998           0.02040    0.05203    0.39   0.6950
## Year1999           0.08958    0.04776    1.88   0.0607 .
## Year2000           0.14445    0.05114    2.82   0.0047 **
## Year2001           0.19831    0.05019    3.95  7.8e-05 ***
## Year2002           0.14281    0.04764    3.00   0.0027 **
## Year2003           0.10754    0.04494    2.39   0.0167 *
## Year2004           0.17689    0.04478    3.95  7.9e-05 ***
## Year2005           0.04170    0.04244    0.98   0.3259
## Year2006           0.05232    0.04296    1.22   0.2233
```

```
## Year2007           0.07694    0.04216    1.82   0.0680 .
## Year2008           0.03985    0.04157    0.96   0.3378
## Year2009           0.02143    0.04092    0.52   0.6006
## Year2010           0.04556    0.04288    1.06   0.2880
## Year2011           0.10360    0.04183    2.48   0.0133 *
## Year2012           0.13926    0.04152    3.35   0.0008 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.576
## Multiple R-squared:  0.00945,    Adjusted R-squared:  0.00738
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
##  648 weights are ~= 1. The remaining 7500 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.208   0.862   0.947   0.908   0.984   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.23e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Sample size for the above analysis:  8148"
## [1] ""
## [1] ""
## [1] "####################################"
## [1] "Analysis of AJSC 2306"
## [1] "####################################"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  282  309  229  318  408  378  332  376  344  364  331  386  418  478  498
## 2011 2012
##  547  569
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  176  166  168  228  244  214  249  280  264  271  266  297  320  364  382
## 2011 2012
```

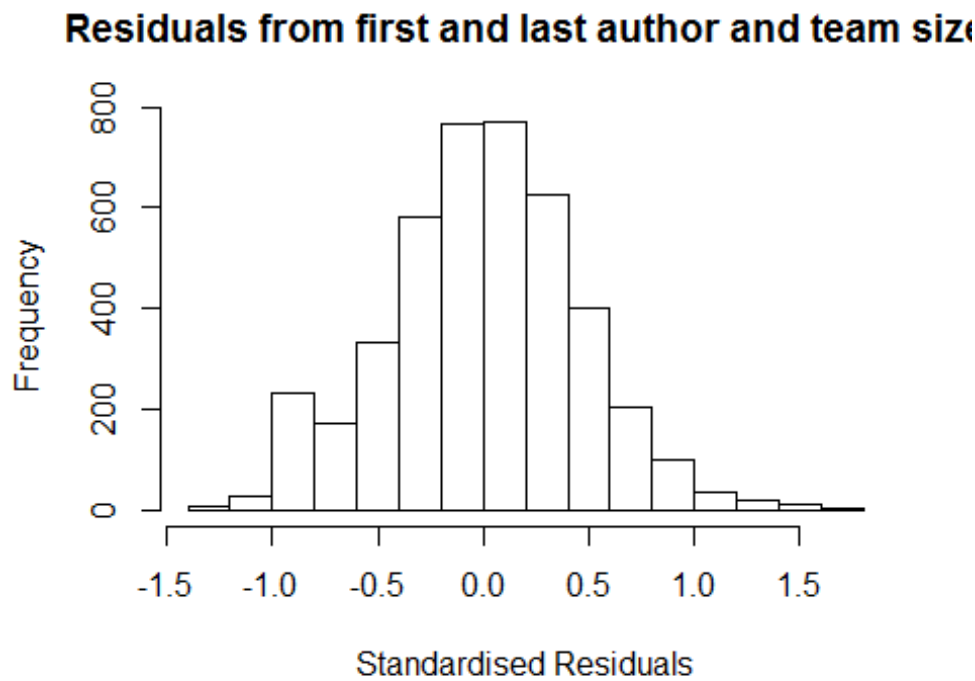
```
##  422  451
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  158  157  156  210  223  189  228  257  231  234  248  259  285  327  345
## 2011 2012
##  379  406
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
##  Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 120, df = 16, p-value <2e-16
```




```
##
##  Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 7.5, df = 1, p-value = 0.006
```

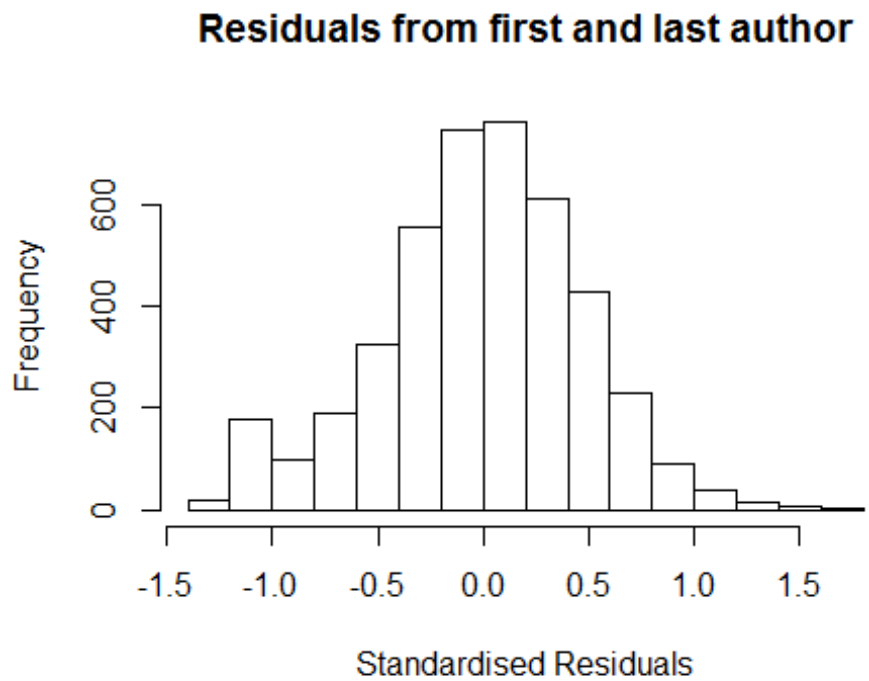



```
## [1] "Female first author team size 2018 geometric mean: 3.1346468232773"
## [1] "Male first author team size 2018 geometric mean: 3.09895528712188"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 38000, p-value = 0.8
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 3.10874067775015"
## [1] "Male last author team size 2018 geometric mean: 3.11492803111365"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 34000, p-value = 0.9
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.049  1           1.024
## LastAuthorFemale  1.041  1           1.020
## UniqueAuthors     1.110  4           1.013
## Year              1.144 16           1.004
```



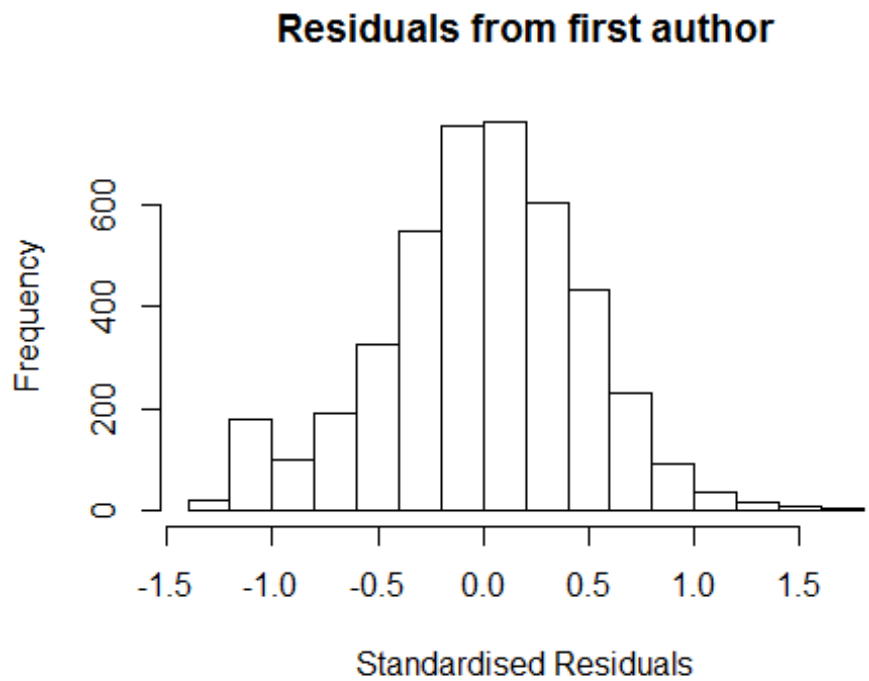

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
UniqueAuthors +
##     Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##     k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.3263 -0.2831  0.0067  0.2976  1.6538
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         0.91737    0.04513   20.33   <2e-16 ***
## FirstAuthorFemale1  0.00975    0.01601    0.61     0.54
## LastAuthorFemale1   0.01378    0.01752    0.79     0.43
## UniqueAuthors2      0.22707    0.02348    9.67   <2e-16 ***
## UniqueAuthors3      0.28407    0.02315   12.27   <2e-16 ***
## UniqueAuthors4      0.35135    0.02563   13.71   <2e-16 ***
## UniqueAuthors5      0.42083    0.02498   16.85   <2e-16 ***
## Year1997           -0.05185    0.06877   -0.75     0.45
## Year1998           -0.02224    0.06300   -0.35     0.72
## Year1999           -0.00530    0.05347   -0.10     0.92
```

```
## Year2000           -0.02572    0.05359   -0.48     0.63
## Year2001           -0.06553    0.05803   -1.13     0.26
## Year2002            0.02628    0.05362    0.49     0.62
## Year2003           -0.00876    0.04962   -0.18     0.86
## Year2004            0.00170    0.05058    0.03     0.97
## Year2005            0.06172    0.04984    1.24     0.22
## Year2006            0.00213    0.04905    0.04     0.97
## Year2007            0.04070    0.05080    0.80     0.42
## Year2008            0.03901    0.04881    0.80     0.42
## Year2009           -0.03052    0.04808   -0.63     0.53
## Year2010            0.02525    0.04840    0.52     0.60
## Year2011           -0.01536    0.04820   -0.32     0.75
## Year2012           -0.00897    0.04842   -0.19     0.85
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.431
## Multiple R-squared:  0.104,  Adjusted R-squared:  0.0993
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
##  364 weights are ~= 1. The remaining 3928 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.108   0.860   0.950   0.896   0.986   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          2.33e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.042  1           1.021
## LastAuthorFemale  1.033  1           1.016
## Year              1.034 16           1.001
```



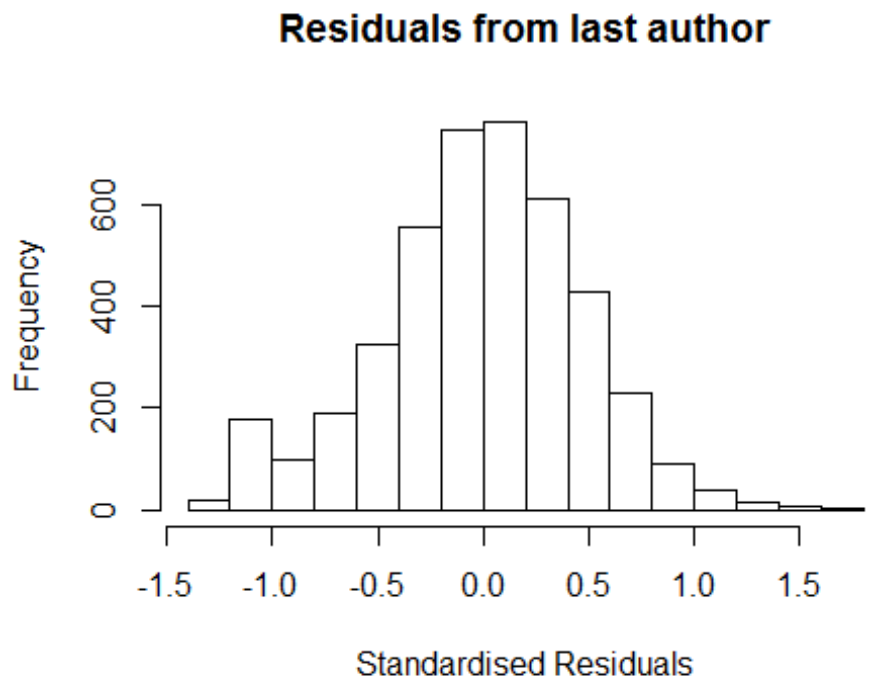

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.2499 -0.2871  0.0106  0.3056  1.6470
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)          1.10588    0.04728   23.39   <2e-16 ***
## FirstAuthorFemale1   0.02994    0.01661    1.80    0.072 .
## LastAuthorFemale1    0.01035    0.01837    0.56    0.573
## Year1997            -0.06865    0.07811   -0.88    0.380
## Year1998            -0.01161    0.07063   -0.16    0.869
## Year1999            -0.00255    0.05817   -0.04    0.965
## Year2000            -0.00896    0.05803   -0.15    0.877
## Year2001            -0.05692    0.06479   -0.88    0.380
## Year2002             0.07313    0.05912    1.24    0.216
## Year2003             0.01845    0.05502    0.34    0.737
## Year2004             0.05454    0.05557    0.98    0.326
## Year2005             0.12324    0.05431    2.27    0.023 *
```

```
## Year2006            0.07283    0.05366    1.36    0.175
## Year2007            0.09752    0.05500    1.77    0.076 .
## Year2008            0.11403    0.05379    2.12    0.034 *
## Year2009            0.03940    0.05310    0.74    0.458
## Year2010            0.07284    0.05356    1.36    0.174
## Year2011            0.02185    0.05295    0.41    0.680
## Year2012            0.06334    0.05325    1.19    0.234
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.446
## Multiple R-squared:  0.0126, Adjusted R-squared:  0.00846
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
##  349 weights are ~= 1. The remaining 3943 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.143   0.861   0.950   0.895   0.986   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          2.33e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.021  1           1.011
## Year              1.021 16           1.001
```




```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.25399 -0.28678  0.00912  0.30478  1.64578
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.10681    0.04730   23.40   <2e-16 ***
## FirstAuthorFemale1  0.03219    0.01652    1.95    0.051 .
## Year1997           -0.06863    0.07810   -0.88    0.380
## Year1998           -0.01186    0.07066   -0.17    0.867
## Year1999           -0.00234    0.05821   -0.04    0.968
## Year2000           -0.00822    0.05808   -0.14    0.887
## Year2001           -0.05659    0.06481   -0.87    0.383
## Year2002            0.07377    0.05913    1.25    0.212
## Year2003            0.01950    0.05504    0.35    0.723
## Year2004            0.05494    0.05560    0.99    0.323
## Year2005            0.12366    0.05436    2.27    0.023 *
## Year2006            0.07292    0.05372    1.36    0.175
```

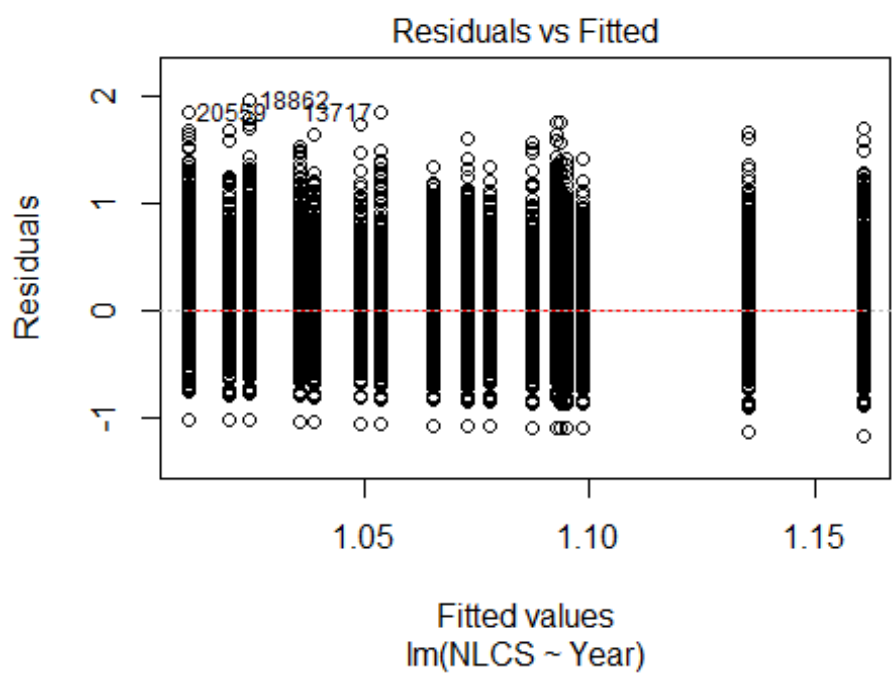
```
## Year2007            0.09852    0.05505    1.79    0.074 .
## Year2008            0.11499    0.05381    2.14    0.033 *
## Year2009            0.04003    0.05315    0.75    0.451
## Year2010            0.07361    0.05361    1.37    0.170
## Year2011            0.02293    0.05295    0.43    0.665
## Year2012            0.06428    0.05325    1.21    0.227
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.446
## Multiple R-squared:  0.0125, Adjusted R-squared:  0.00859
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
##  355 weights are ~= 1. The remaining 3937 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.144   0.860   0.951   0.894   0.986   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          2.33e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##                   GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.013  1           1.006
## Year             1.013 16           1.000
```



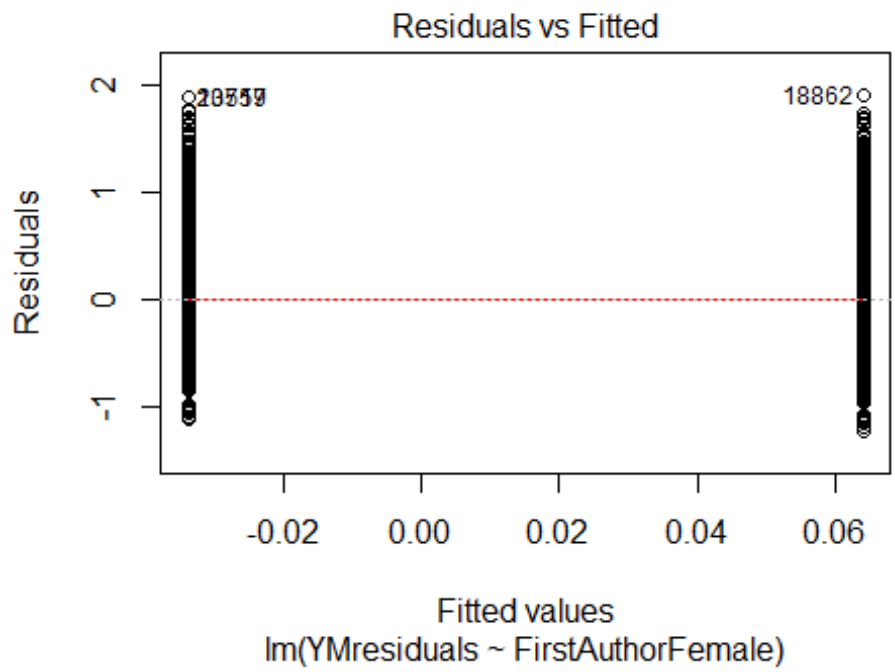

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.23475 -0.28691  0.00681  0.30898  1.64200
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)        1.109267   0.047234   23.48   <2e-16 ***
## LastAuthorFemale1  0.018083   0.018206    0.99    0.321
## Year1997          -0.068902   0.078132   -0.88    0.378
## Year1998          -0.009827   0.070472   -0.14    0.889
## Year1999          -0.000956   0.058055   -0.02    0.987
## Year2000          -0.007548   0.057904   -0.13    0.896
## Year2001          -0.055265   0.064844   -0.85    0.394
## Year2002           0.076070   0.059063    1.29    0.198
## Year2003           0.019842   0.054944    0.36    0.718
## Year2004           0.057567   0.055466    1.04    0.299
## Year2005           0.125484   0.054249    2.31    0.021 *
## Year2006           0.077194   0.053556    1.44    0.150
```

```
## Year2007           0.101061   0.054908    1.84    0.066 .
## Year2008           0.117283   0.053717    2.18    0.029 *
## Year2009           0.042629   0.052999    0.80    0.421
## Year2010           0.078203   0.053334    1.47    0.143
## Year2011           0.026424   0.052814    0.50    0.617
## Year2012           0.067559   0.053189    1.27    0.204
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.446
## Multiple R-squared:  0.0118, Adjusted R-squared:  0.00792
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
##  359 weights are ~= 1. The remaining 3933 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.146   0.861   0.951   0.894   0.986   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          2.33e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Sample size for the above analysis:  4292"
## [1] ""
## [1] ""
## [1] "####################################"
## [1] "Analysis of AJSC 2307"
## [1] "####################################"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 1236 1171 1199 1225 1381 1354 1296 1023 1075 1059 1111 1146 1123 1291 1244
## 2011 2012
## 1466 1351
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  743  722  685  810  680  648  870  684  695  680  737  789  766  873  915
## 2011 2012
```

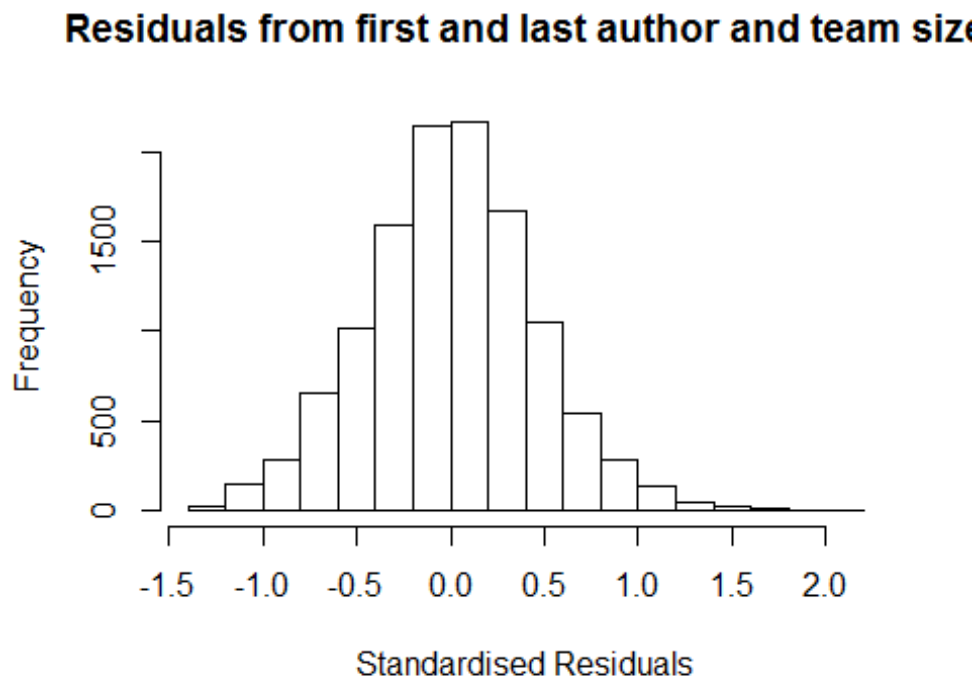
```
## 1094 1002
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  676  636  610  704  610  573  762  593  591  567  641  704  663  788  809
## 2011 2012
##  958  888
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
##  Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 130, df = 16, p-value <2e-16
```




```
##
##  Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 0.0014, df = 1, p-value = 1
```

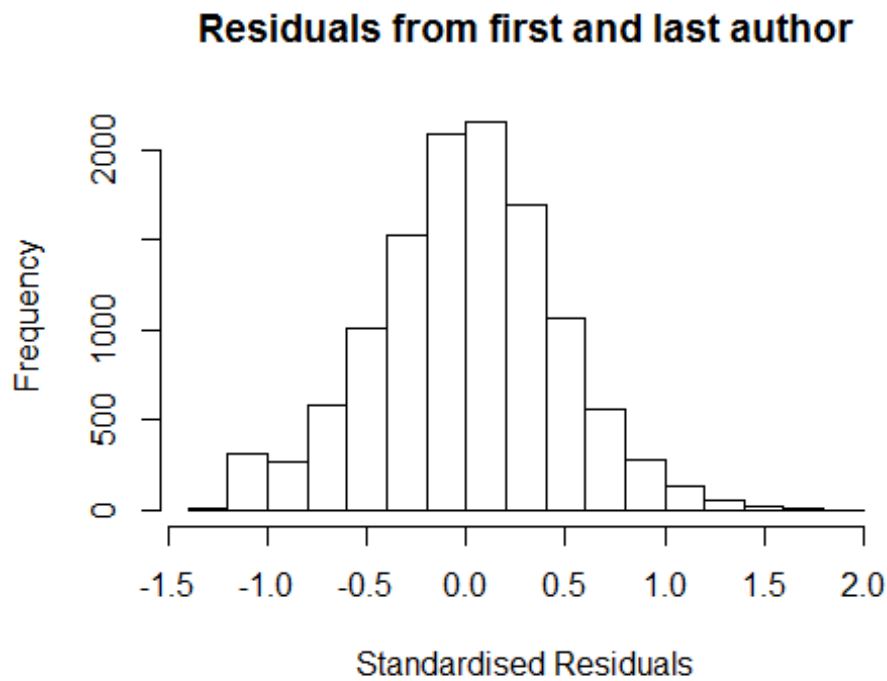



```
## [1] "Female first author team size 2018 geometric mean: 4.75951128447363"
## [1] "Male first author team size 2018 geometric mean: 3.9695936925648"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 180000, p-value = 2e-05
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 4.64362987117008"
## [1] "Male last author team size 2018 geometric mean: 4.17829541184902"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 160000, p-value = 0.009
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##                     GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.026  1           1.013
## LastAuthorFemale  1.016  1           1.008
## UniqueAuthors     1.052  4           1.006
## Year              1.062 16           1.002
```



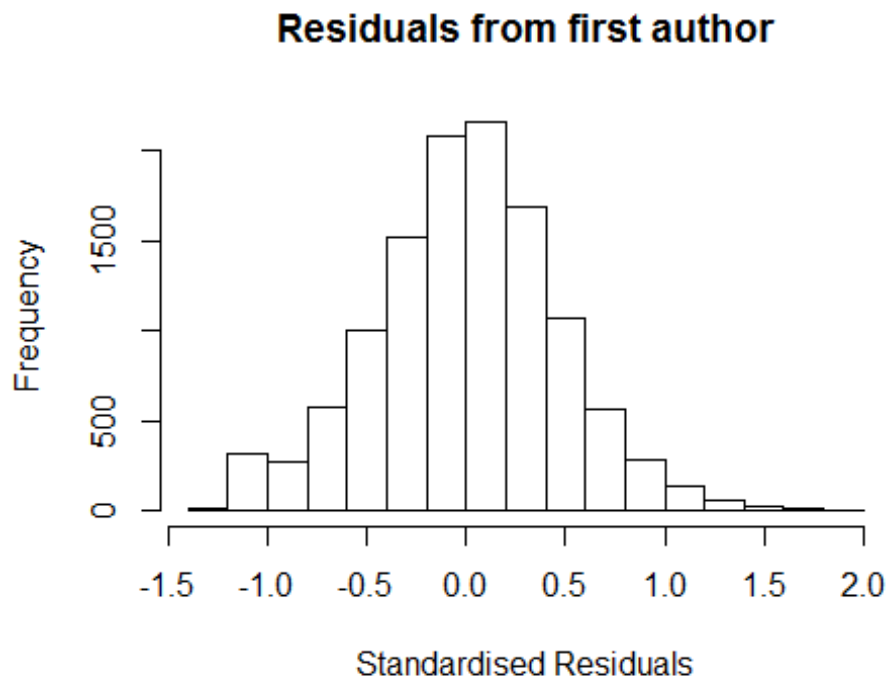

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
UniqueAuthors +
##     Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##     k.max = 1000))
##  \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.35141 -0.28821  0.00266  0.28999  2.01322
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         0.89064    0.02512   35.46  < 2e-16 ***
## FirstAuthorFemale1  0.06642    0.00884    7.51  6.3e-14 ***
## LastAuthorFemale1   0.01442    0.01006    1.43  0.15169
## UniqueAuthors2      0.21016    0.01826   11.51  < 2e-16 ***
## UniqueAuthors3      0.25122    0.01804   13.92  < 2e-16 ***
## UniqueAuthors4      0.27430    0.01841   14.90  < 2e-16 ***
## UniqueAuthors5      0.39435    0.01704   23.15  < 2e-16 ***
## Year1997           -0.02867    0.02777   -1.03  0.30182
## Year1998           -0.08023    0.02759   -2.91  0.00365 **
## Year1999           -0.06932    0.02684   -2.58  0.00981 **
```

```
## Year2000           -0.07953    0.02693   -2.95  0.00315 **
## Year2001           -0.12338    0.02827   -4.36  1.3e-05 ***
## Year2002           -0.10601    0.02600   -4.08  4.6e-05 ***
## Year2003           -0.12894    0.02598   -4.96  7.0e-07 ***
## Year2004           -0.11807    0.02619   -4.51  6.6e-06 ***
## Year2005           -0.09160    0.02685   -3.41  0.00065 ***
## Year2006           -0.15393    0.02720   -5.66  1.6e-08 ***
## Year2007           -0.16259    0.02497   -6.51  7.7e-11 ***
## Year2008           -0.18651    0.02676   -6.97  3.3e-12 ***
## Year2009           -0.18169    0.02583   -7.03  2.1e-12 ***
## Year2010           -0.19343    0.02591   -7.46  8.9e-14 ***
## Year2011           -0.19559    0.02550   -7.67  1.8e-14 ***
## Year2012           -0.12670    0.02673   -4.74  2.2e-06 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.43
## Multiple R-squared:  0.0856, Adjusted R-squared:  0.0839
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
##  observation 9121 is an outlier with |weight| <= 8.3e-07 ( < 8.5e-06);
##  993 weights are ~= 1. The remaining 10779 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0012  0.8640  0.9510  0.8990  0.9860  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          8.49e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.015  1           1.007
## LastAuthorFemale  1.012  1           1.006
## Year              1.020 16           1.001
```



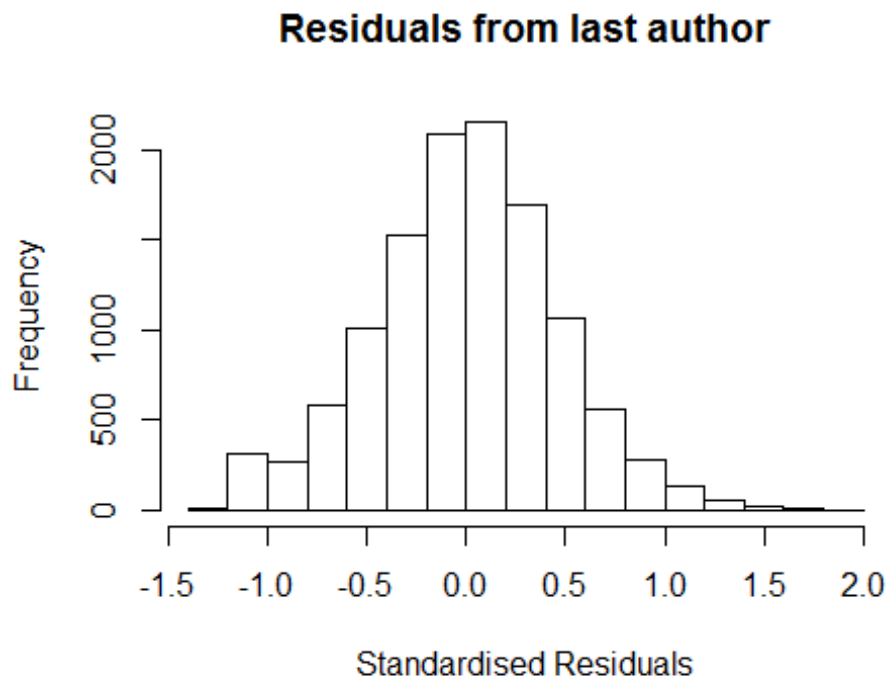

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.24468 -0.29332  0.00775  0.29575  1.91780
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)          1.13782    0.02076   54.80  < 2e-16 ***
## FirstAuthorFemale1   0.09307    0.00904   10.30  < 2e-16 ***
## LastAuthorFemale1    0.01379    0.01034    1.33  0.18230
## Year1997            -0.03257    0.02814   -1.16  0.24711
## Year1998            -0.08254    0.02774   -2.98  0.00293 **
## Year1999            -0.07886    0.02744   -2.87  0.00406 **
## Year2000            -0.08816    0.02742   -3.21  0.00131 **
## Year2001            -0.12490    0.02959   -4.22  2.4e-05 ***
## Year2002            -0.09782    0.02662   -3.68  0.00024 ***
## Year2003            -0.10971    0.02646   -4.15  3.4e-05 ***
## Year2004            -0.09200    0.02694   -3.42  0.00064 ***
## Year2005            -0.06600    0.02771   -2.38  0.01725 *
```

```
## Year2006           -0.13011    0.02769   -4.70  2.7e-06 ***
## Year2007           -0.13582    0.02548   -5.33  9.9e-08 ***
## Year2008           -0.16159    0.02746   -5.88  4.1e-09 ***
## Year2009           -0.15075    0.02634   -5.72  1.1e-08 ***
## Year2010           -0.16169    0.02648   -6.11  1.0e-09 ***
## Year2011           -0.17761    0.02619   -6.78  1.3e-11 ***
## Year2012           -0.09375    0.02741   -3.42  0.00063 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.438
## Multiple R-squared:  0.0188, Adjusted R-squared:  0.0173
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
##  977 weights are ~= 1. The remaining 10796 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0165  0.8620  0.9510  0.8970  0.9860  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          8.49e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.011  1           1.006
## Year              1.011 16           1.000
```




```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.23428 -0.29287  0.00798  0.29617  1.91353
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)          1.13958    0.02070   55.04  < 2e-16 ***
## FirstAuthorFemale1   0.09471    0.00906   10.46  < 2e-16 ***
## Year1997            -0.03181    0.02812   -1.13  0.25808
## Year1998            -0.08156    0.02771   -2.94  0.00325 **
## Year1999            -0.07861    0.02743   -2.87  0.00416 **
## Year2000            -0.08718    0.02741   -3.18  0.00147 **
## Year2001            -0.12374    0.02956   -4.19  2.9e-05 ***
## Year2002            -0.09644    0.02659   -3.63  0.00029 ***
## Year2003            -0.10893    0.02645   -4.12  3.8e-05 ***
## Year2004            -0.09111    0.02694   -3.38  0.00072 ***
## Year2005            -0.06553    0.02769   -2.37  0.01797 *
## Year2006            -0.12904    0.02767   -4.66  3.1e-06 ***
```

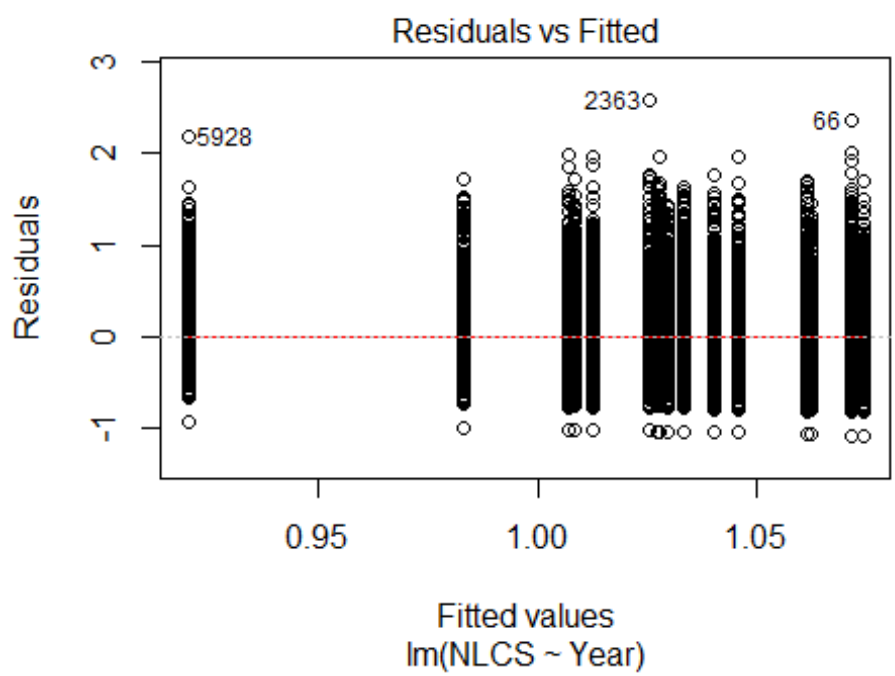
```
## Year2007           -0.13518    0.02548   -5.30  1.1e-07 ***
## Year2008           -0.16051    0.02744   -5.85  5.1e-09 ***
## Year2009           -0.14993    0.02634   -5.69  1.3e-08 ***
## Year2010           -0.16082    0.02647   -6.07  1.3e-09 ***
## Year2011           -0.17641    0.02618   -6.74  1.7e-11 ***
## Year2012           -0.09227    0.02740   -3.37  0.00076 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.438
## Multiple R-squared:  0.0186, Adjusted R-squared:  0.0172
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
##  1002 weights are ~= 1. The remaining 10771 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0175  0.8630  0.9500  0.8960  0.9860  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          8.49e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##                   GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.009  1           1.004
## Year             1.009 16           1.000
```



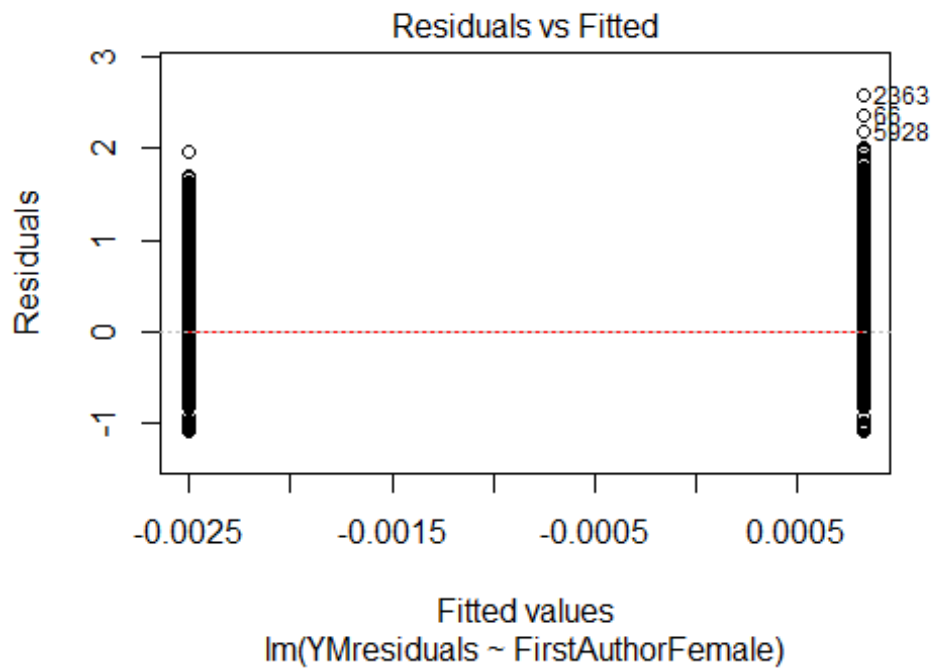

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.18799 -0.29136  0.00758  0.29724  1.97609
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.1602     0.0205   56.59  < 2e-16 ***
## LastAuthorFemale1   0.0278     0.0104    2.68  0.00734 **
## Year1997           -0.0292     0.0280   -1.04  0.29654
## Year1998           -0.0816     0.0277   -2.95  0.00322 **
## Year1999           -0.0743     0.0274   -2.72  0.00658 **
## Year2000           -0.0857     0.0273   -3.13  0.00172 **
## Year2001           -0.1231     0.0295   -4.18  3.0e-05 ***
## Year2002           -0.0926     0.0265   -3.49  0.00048 ***
## Year2003           -0.1054     0.0263   -4.01  6.2e-05 ***
## Year2004           -0.0850     0.0268   -3.17  0.00155 **
## Year2005           -0.0575     0.0277   -2.08  0.03760 *
## Year2006           -0.1211     0.0277   -4.38  1.2e-05 ***
```

```
## Year2007           -0.1282     0.0254   -5.04  4.6e-07 ***
## Year2008           -0.1515     0.0274   -5.52  3.4e-08 ***
## Year2009           -0.1397     0.0263   -5.30  1.2e-07 ***
## Year2010           -0.1492     0.0264   -5.64  1.7e-08 ***
## Year2011           -0.1693     0.0262   -6.46  1.1e-10 ***
## Year2012           -0.0775     0.0274   -2.83  0.00470 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.439
## Multiple R-squared:  0.00973,    Adjusted R-squared:  0.0083
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
##  1000 weights are ~= 1. The remaining 10773 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0062  0.8610  0.9500  0.8960  0.9860  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          8.49e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Sample size for the above analysis:  11773"
## [1] ""
## [1] ""
## [1] "####################################"
## [1] "Analysis of AJSC 2308"
## [1] "####################################"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  910  921 1006 1075 1248 1246 1098 1083 1144 1232 1271 1306 1437 1814 1793
## 2011 2012
## 1980 1998
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  673  642  714  745  831  797  838  855  923  979  992 1037 1114 1486 1422
## 2011 2012
```

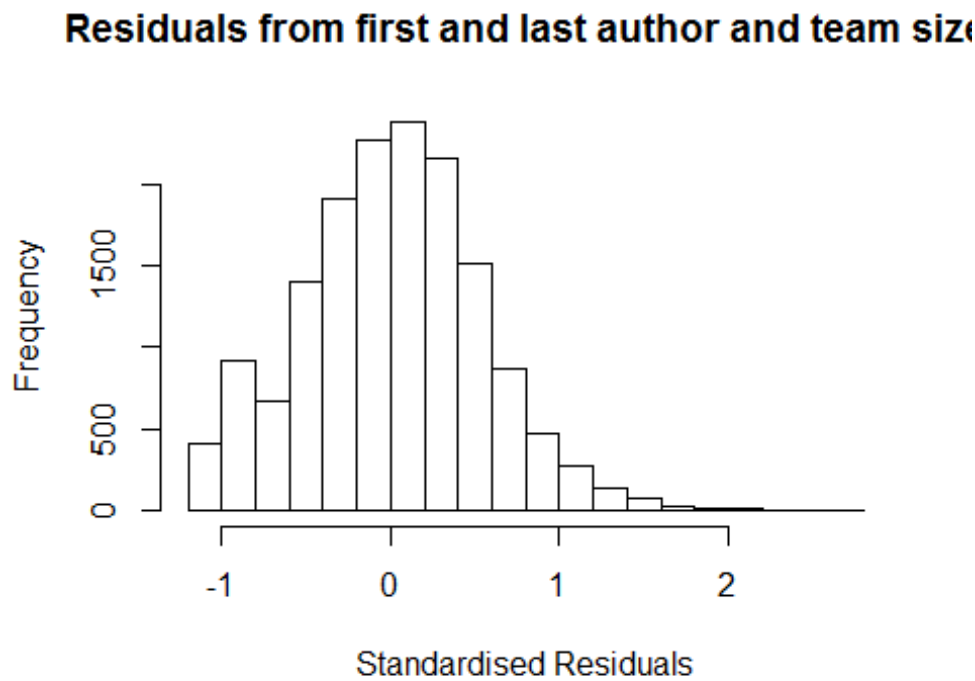
```
## 1523 1603
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  632  580  668  680  763  729  760  779  850  883  880  911  995 1320 1271
## 2011 2012
## 1365 1429
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
##  Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 170, df = 16, p-value <2e-16
```




```
##
##  Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 0.0017, df = 1, p-value = 1
```

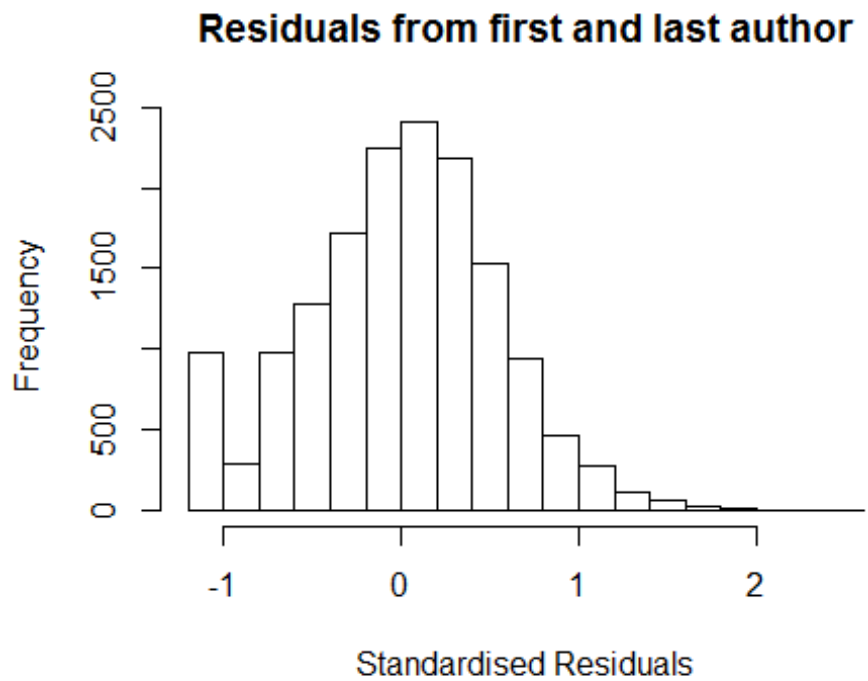



```
## [1] "Female first author team size 2018 geometric mean: 2.76367285294029"
## [1] "Male first author team size 2018 geometric mean: 2.59177705103728"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 350000, p-value = 0.07
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 2.6600127350141"
## [1] "Male last author team size 2018 geometric mean: 2.64977781259452"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 3e+05, p-value = 0.8
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.108  1           1.052
## LastAuthorFemale  1.100  1           1.049
## UniqueAuthors     1.049  4           1.006
## Year              1.073 16           1.002
```



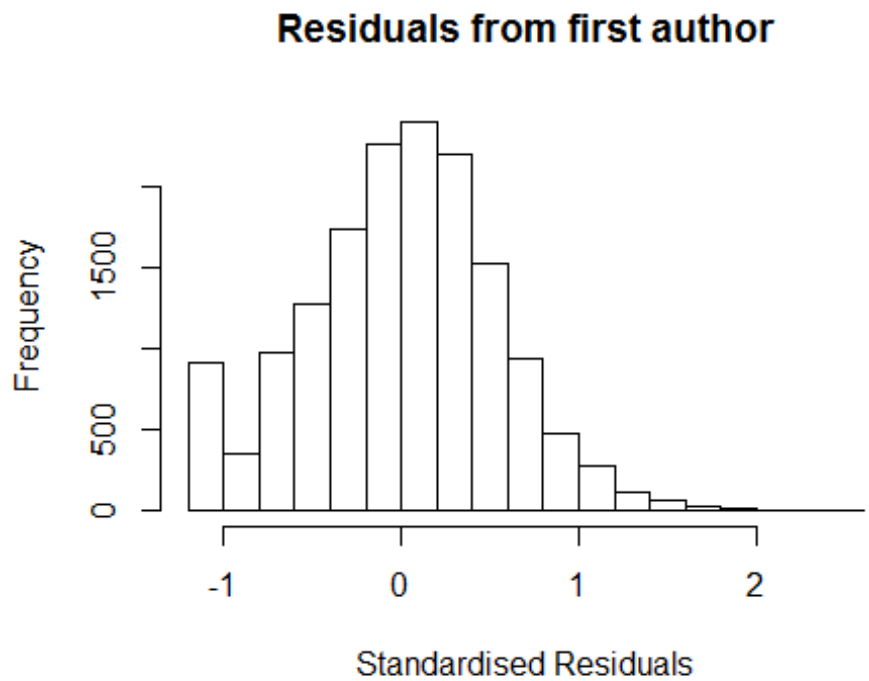

```
## [1] "List of  2 outliers with residuals above  2.5"
##        ScopusId  NLCS Year OneField Fields residuals
## 66   0030429663 3.430 1996     1903      4     2.516
## 2363 0031856333 3.602 1998     2100      2     2.736
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
UniqueAuthors +
##     Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##     k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.1895 -0.3477  0.0144  0.3486  2.7365
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         0.91372    0.02695   33.90  < 2e-16 ***
## FirstAuthorFemale1 -0.01566    0.01054   -1.49    0.137
## LastAuthorFemale1   0.00280    0.01130    0.25    0.805
## UniqueAuthors2      0.20561    0.01216   16.91  < 2e-16 ***
## UniqueAuthors3      0.22919    0.01289   17.78  < 2e-16 ***
## UniqueAuthors4      0.24470    0.01462   16.73  < 2e-16 ***
## UniqueAuthors5      0.29110    0.01411   20.63  < 2e-16 ***
## Year1997           -0.02049    0.03591   -0.57    0.568
## Year1998           -0.04821    0.03385   -1.42    0.154
```

```
## Year1999           -0.05915    0.03425   -1.73    0.084 .
## Year2000           -0.05556    0.03419   -1.63    0.104
## Year2001           -0.14516    0.03360   -4.32  1.6e-05 ***
## Year2002           -0.06869    0.03370   -2.04    0.042 *
## Year2003           -0.04583    0.03271   -1.40    0.161
## Year2004           -0.01982    0.03218   -0.62    0.538
## Year2005           -0.04903    0.03126   -1.57    0.117
## Year2006           -0.01817    0.03112   -0.58    0.559
## Year2007           -0.00302    0.03069   -0.10    0.922
## Year2008           -0.02767    0.03026   -0.91    0.361
## Year2009           -0.06044    0.02970   -2.04    0.042 *
## Year2010           -0.04739    0.03009   -1.58    0.115
## Year2011           -0.06718    0.03038   -2.21    0.027 *
## Year2012           -0.06595    0.02980   -2.21    0.027 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.518
## Multiple R-squared:  0.0467, Adjusted R-squared:  0.0453
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
##  2 observations c(48,1456) are outliers with |weight| = 0 ( < 6.5e-06);
##  1330 weights are ~= 1. The remaining 14163 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0432  0.8700  0.9510  0.9040  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          6.45e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.099  1           1.048
## LastAuthorFemale  1.088  1           1.043
## Year              1.033 16           1.001
```



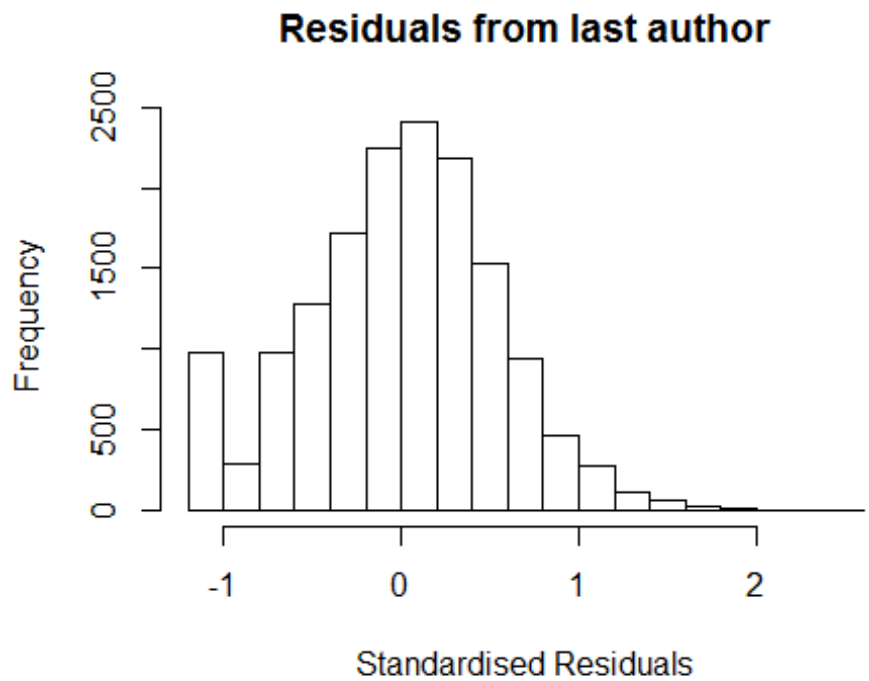

```
## [1] "List of  1 outliers with residuals above  2.5"
##        ScopusId  NLCS Year OneField Fields residuals
## 2363 0031856333 3.602 1998     2100      2     2.589
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.0780 -0.3510  0.0225  0.3555  2.5892
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.04898    0.02644   39.67  < 2e-16 ***
## FirstAuthorFemale1 -0.00957    0.01075   -0.89    0.374
## LastAuthorFemale1  -0.01220    0.01155   -1.06    0.291
## Year1997           -0.01340    0.03678   -0.36    0.716
## Year1998           -0.03616    0.03452   -1.05    0.295
## Year1999           -0.04251    0.03519   -1.21    0.227
## Year2000           -0.04884    0.03476   -1.41    0.160
## Year2001           -0.13934    0.03445   -4.04  5.3e-05 ***
## Year2002           -0.07102    0.03473   -2.04    0.041 *
## Year2003           -0.03485    0.03348   -1.04    0.298
## Year2004           -0.00496    0.03298   -0.15    0.880
## Year2005           -0.01964    0.03174   -0.62    0.536
```

```
## Year2006            0.02130    0.03142    0.68    0.498
## Year2007            0.02899    0.03107    0.93    0.351
## Year2008            0.01250    0.03053    0.41    0.682
## Year2009           -0.01821    0.03005   -0.61    0.544
## Year2010           -0.00754    0.03042   -0.25    0.804
## Year2011           -0.02258    0.03069   -0.74    0.462
## Year2012           -0.02351    0.03017   -0.78    0.436
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.524
## Multiple R-squared:  0.00437,    Adjusted R-squared:  0.00321
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
##  observation 1456 is an outlier with |weight| = 0 ( < 6.5e-06);
##  1301 weights are ~= 1. The remaining 14193 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0034  0.8610  0.9500  0.9030  0.9860  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          6.45e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.025  1           1.012
## Year              1.025 16           1.001
```




```
## [1] "List of  1 outliers with residuals above  2.5"
##        ScopusId  NLCS Year OneField Fields residuals
## 2363 0031856333 3.602 1998     2100      2     2.589
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.0763 -0.3520  0.0234  0.3551  2.5902
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.04795    0.02643   39.66  < 2e-16 ***
## FirstAuthorFemale1 -0.01383    0.01047   -1.32     0.19
## Year1997           -0.01354    0.03678   -0.37     0.71
## Year1998           -0.03615    0.03451   -1.05     0.29
## Year1999           -0.04277    0.03518   -1.22     0.22
## Year2000           -0.04898    0.03475   -1.41     0.16
## Year2001           -0.13989    0.03443   -4.06  4.9e-05 ***
## Year2002           -0.07133    0.03471   -2.05     0.04 *
## Year2003           -0.03555    0.03345   -1.06     0.29
## Year2004           -0.00554    0.03295   -0.17     0.87
## Year2005           -0.02007    0.03172   -0.63     0.53
## Year2006            0.02063    0.03139    0.66     0.51
```

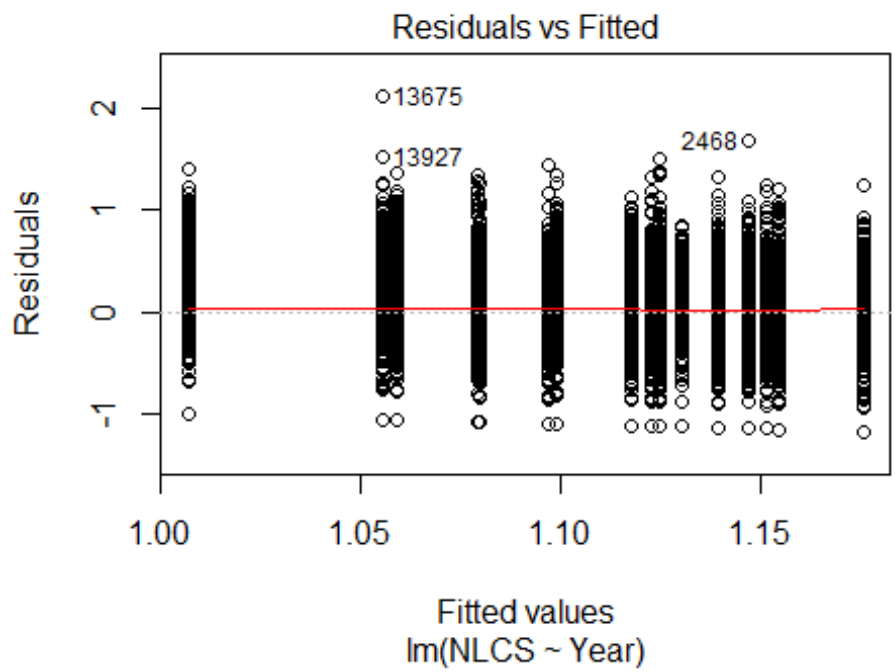
```
## Year2007            0.02838    0.03104    0.91     0.36
## Year2008            0.01195    0.03050    0.39     0.70
## Year2009           -0.01897    0.03002   -0.63     0.53
## Year2010           -0.00811    0.03039   -0.27     0.79
## Year2011           -0.02343    0.03065   -0.76     0.44
## Year2012           -0.02432    0.03013   -0.81     0.42
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.524
## Multiple R-squared:  0.00429,    Adjusted R-squared:  0.00319
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
##  observation 1456 is an outlier with |weight| = 0 ( < 6.5e-06);
##  1286 weights are ~= 1. The remaining 14208 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0034  0.8610  0.9500  0.9030  0.9860  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          6.45e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                    psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##                   GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.015  1           1.007
## Year             1.015 16           1.000
```



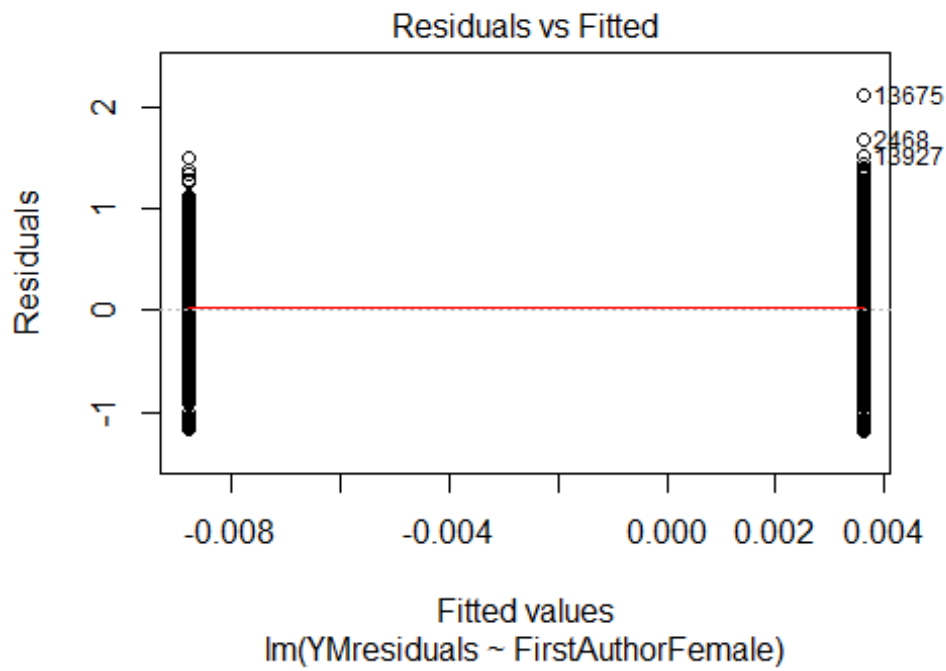

```
## [1] "List of  1 outliers with residuals above  2.5"
##        ScopusId  NLCS Year OneField Fields residuals
## 2363 0031856333 3.602 1998     2100      2     2.589
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.0762 -0.3518  0.0234  0.3554  2.5904
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.04799    0.02643   39.65  < 2e-16 ***
## LastAuthorFemale1 -0.01588    0.01122   -1.42     0.16
## Year1997          -0.01336    0.03679   -0.36     0.72
## Year1998          -0.03643    0.03454   -1.05     0.29
## Year1999          -0.04264    0.03520   -1.21     0.23
## Year2000          -0.04910    0.03477   -1.41     0.16
## Year2001          -0.13951    0.03447   -4.05  5.2e-05 ***
## Year2002          -0.07123    0.03475   -2.05     0.04 *
## Year2003          -0.03514    0.03350   -1.05     0.29
## Year2004          -0.00535    0.03299   -0.16     0.87
## Year2005          -0.02040    0.03173   -0.64     0.52
## Year2006           0.02059    0.03142    0.66     0.51
```

```
## Year2007           0.02816    0.03106    0.91     0.36
## Year2008           0.01165    0.03054    0.38     0.70
## Year2009          -0.01918    0.03002   -0.64     0.52
## Year2010          -0.00830    0.03042   -0.27     0.79
## Year2011          -0.02360    0.03067   -0.77     0.44
## Year2012          -0.02445    0.03017   -0.81     0.42
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.524
## Multiple R-squared:  0.00431,    Adjusted R-squared:  0.00322
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
##  observation 1456 is an outlier with |weight| = 0 ( < 6.5e-06);
##  1306 weights are ~= 1. The remaining 14188 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0034  0.8610  0.9500  0.9030  0.9860  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          6.45e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Sample size for the above analysis:  15495"
## [1] ""
## [1] ""
## [1] "####################################"
## [1] "Analysis of AJSC 2309"
## [1] "####################################"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  576  590  708  641  639  607  621  573  570  787  798  908  893 1094 1069
## 2011 2012
## 1208 1209
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  459  473  551  516  373  297  521  445  462  637  635  767  726  919  876
```

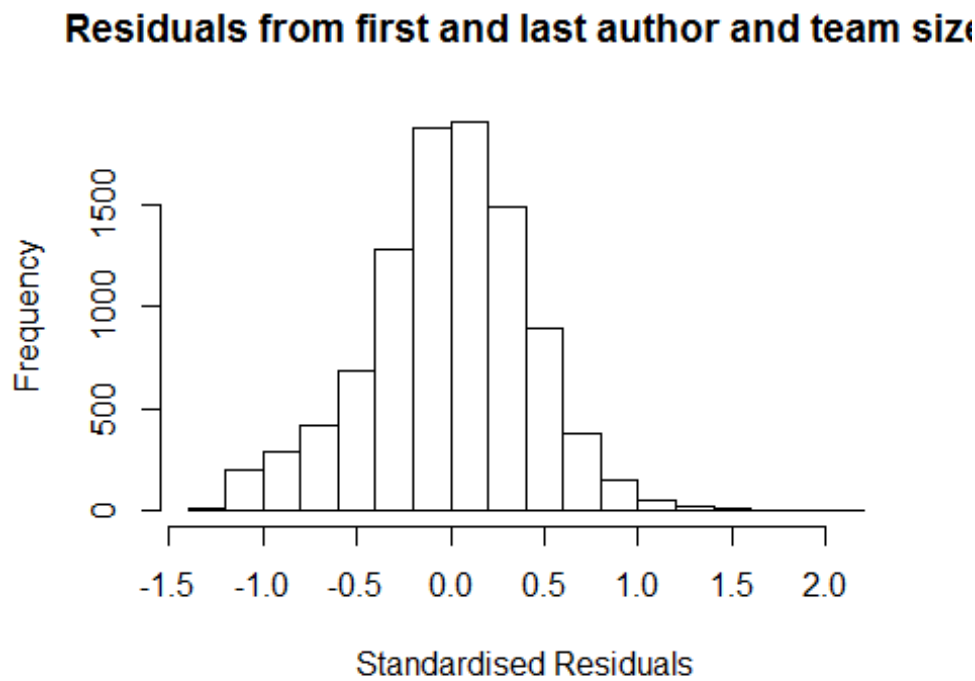
```
## 2011 2012
## 1026  994
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  427  433  506  474  339  259  457  398  415  574  558  674  646  846  793
## 2011 2012
##  929  900
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
##  Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 140, df = 16, p-value <2e-16
```




```
##
##  Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 1.6, df = 1, p-value = 0.2
```

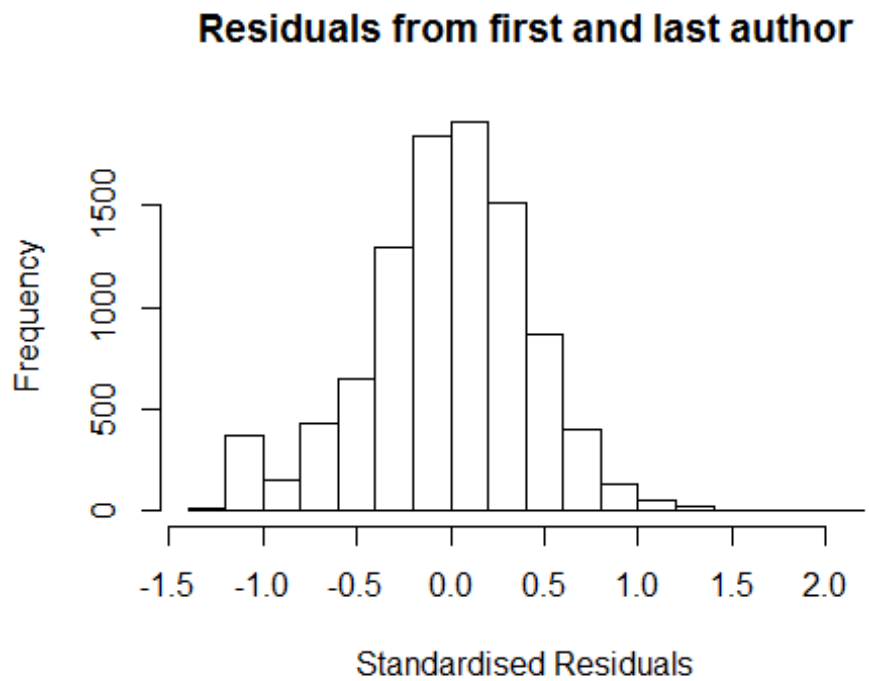



```
## [1] "Female first author team size 2018 geometric mean: 3.38197751287651"
## [1] "Male first author team size 2018 geometric mean: 3.27517424622129"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 160000, p-value = 0.7
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 3.39596491366806"
## [1] "Male last author team size 2018 geometric mean: 3.28821778133546"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 130000, p-value = 0.5
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.041  1           1.020
## LastAuthorFemale  1.032  1           1.016
## UniqueAuthors     1.101  4           1.012
## Year              1.129 16           1.004
```



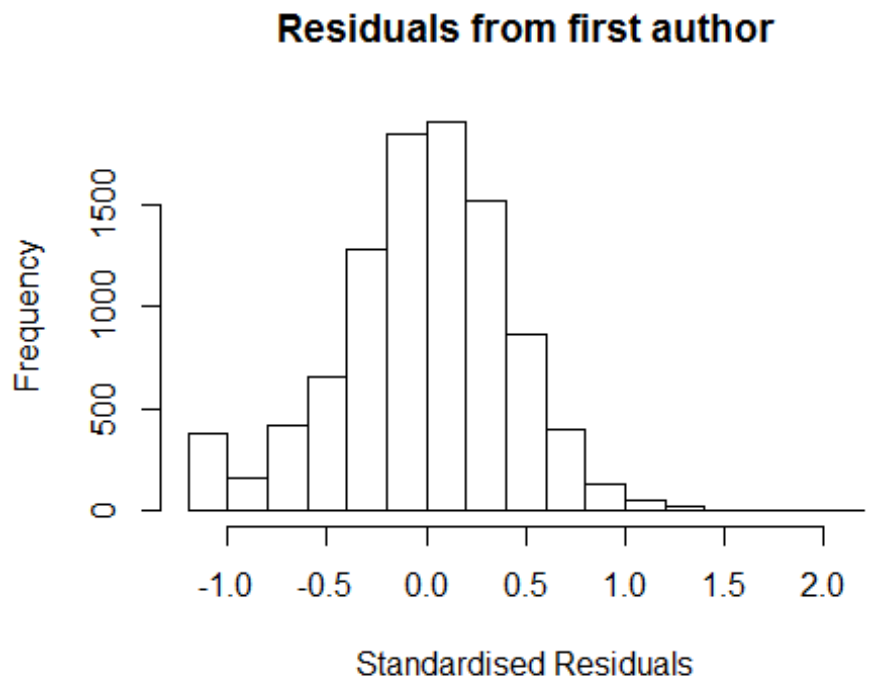

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
UniqueAuthors +
##     Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##     k.max = 1000))
##  \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.24540 -0.26453  0.00724  0.27118  2.10514
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         0.99761    0.02557   39.02  < 2e-16 ***
## FirstAuthorFemale1 -0.01194    0.00991   -1.20    0.228
## LastAuthorFemale1  -0.00508    0.01098   -0.46    0.644
## UniqueAuthors2      0.14741    0.01672    8.82  < 2e-16 ***
## UniqueAuthors3      0.17075    0.01678   10.18  < 2e-16 ***
## UniqueAuthors4      0.18060    0.01797   10.05  < 2e-16 ***
## UniqueAuthors5      0.21019    0.01814   11.59  < 2e-16 ***
## Year1997            0.05462    0.02995    1.82    0.068 .
## Year1998            0.00916    0.02821    0.32    0.745
## Year1999            0.01255    0.02913    0.43    0.667
```

```
## Year2000            0.06718    0.03296    2.04    0.042 *
## Year2001            0.01626    0.03215    0.51    0.613
## Year2002           -0.00958    0.02932   -0.33    0.744
## Year2003            0.02485    0.02880    0.86    0.388
## Year2004           -0.00595    0.02879   -0.21    0.836
## Year2005           -0.05345    0.02724   -1.96    0.050 *
## Year2006           -0.02684    0.02816   -0.95    0.341
## Year2007           -0.01727    0.02656   -0.65    0.515
## Year2008           -0.04738    0.02755   -1.72    0.085 .
## Year2009           -0.06118    0.02751   -2.22    0.026 *
## Year2010           -0.03304    0.02700   -1.22    0.221
## Year2011           -0.08016    0.02698   -2.97    0.003 **
## Year2012           -0.13414    0.02779   -4.83  1.4e-06 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.393
## Multiple R-squared:  0.0347, Adjusted R-squared:  0.0325
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
##  observation 8501 is an outlier with |weight| = 0 ( < 1e-05);
##  822 weights are ~= 1. The remaining 8805 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0313  0.8610  0.9500  0.8920  0.9860  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.04e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.035  1           1.017
## LastAuthorFemale  1.030  1           1.015
## Year              1.033 16           1.001
```



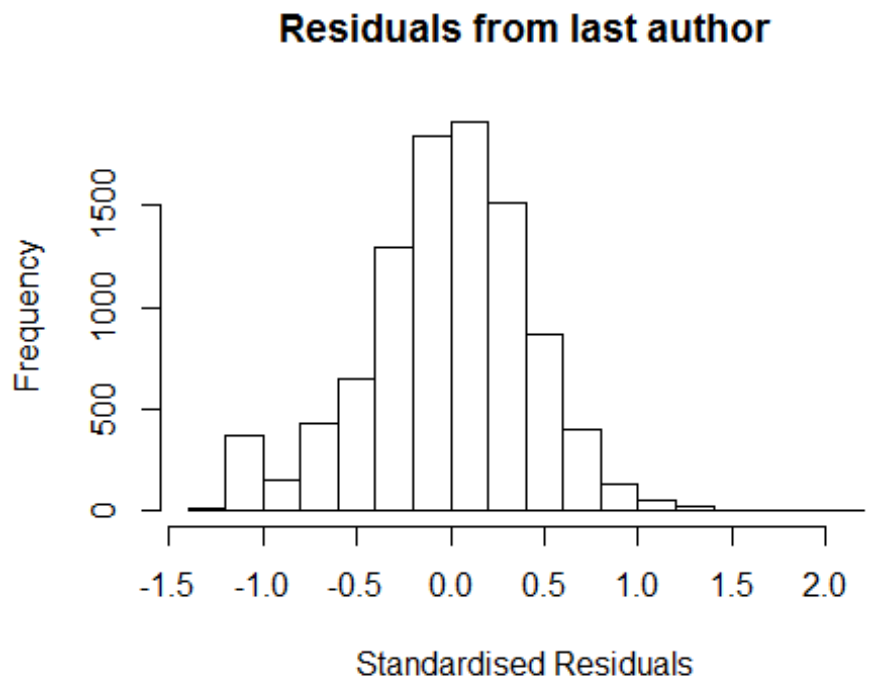

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.20062 -0.26738  0.00793  0.26625  2.09469
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)          1.12787    0.02205   51.16  < 2e-16 ***
## FirstAuthorFemale1 -0.00738    0.00997   -0.74  0.45928
## LastAuthorFemale1  -0.01018    0.01112   -0.92  0.35998
## Year1997            0.05010    0.03010    1.66  0.09610 .
## Year1998            0.01012    0.02806    0.36  0.71845
## Year1999            0.01774    0.02901    0.61  0.54078
## Year2000            0.07275    0.03303    2.20  0.02763 *
## Year2001            0.01692    0.03161    0.54  0.59260
## Year2002           -0.00181    0.02933   -0.06  0.95080
## Year2003            0.03560    0.02857    1.25  0.21278
## Year2004            0.01552    0.02874    0.54  0.58920
## Year2005           -0.03349    0.02718   -1.23  0.21785
```

```
## Year2006           -0.00309    0.02799   -0.11  0.91212
## Year2007            0.00340    0.02655    0.13  0.89805
## Year2008           -0.02570    0.02742   -0.94  0.34864
## Year2009           -0.03550    0.02762   -1.29  0.19881
## Year2010           -0.00682    0.02688   -0.25  0.79961
## Year2011           -0.05256    0.02691   -1.95  0.05082 .
## Year2012           -0.10384    0.02782   -3.73  0.00019 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.395
## Multiple R-squared:  0.0102, Adjusted R-squared:  0.00834
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
##  observation 8501 is an outlier with |weight| = 0 ( < 1e-05);
##  801 weights are ~= 1. The remaining 8826 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0292  0.8610  0.9510  0.8910  0.9860  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.04e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.021  1           1.010
## Year              1.021 16           1.001
```




```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.19922 -0.26672  0.00713  0.26619  2.09680
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.12659    0.02199   51.23  < 2e-16 ***
## FirstAuthorFemale1 -0.00916    0.00993   -0.92  0.35647
## Year1997            0.05022    0.03009    1.67  0.09515 .
## Year1998            0.01010    0.02804    0.36  0.71882
## Year1999            0.01802    0.02900    0.62  0.53431
## Year2000            0.07263    0.03301    2.20  0.02781 *
## Year2001            0.01681    0.03158    0.53  0.59461
## Year2002           -0.00199    0.02933   -0.07  0.94602
## Year2003            0.03544    0.02857    1.24  0.21481
## Year2004            0.01520    0.02872    0.53  0.59666
## Year2005           -0.03384    0.02715   -1.25  0.21268
## Year2006           -0.00350    0.02796   -0.13  0.90025
```

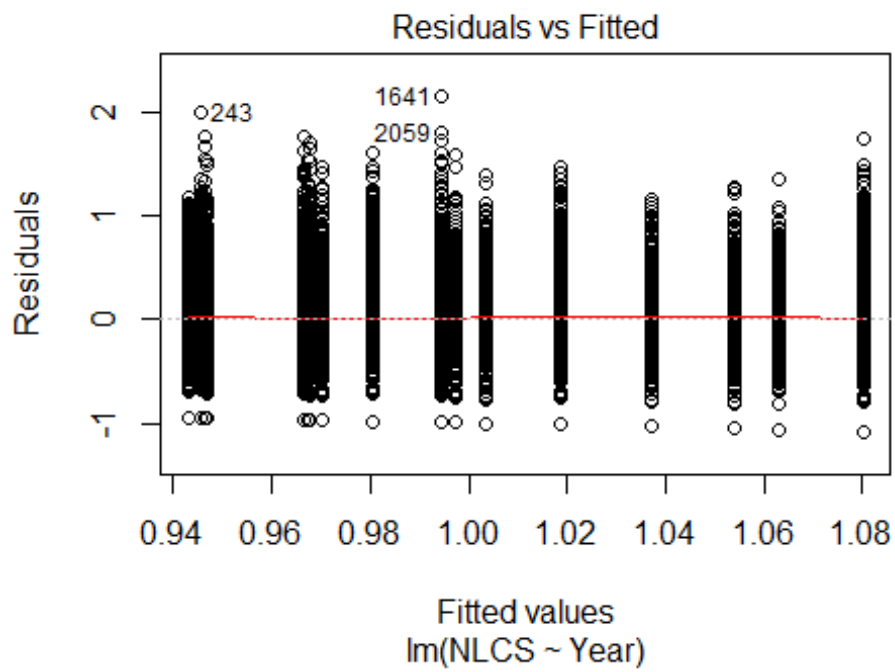
```
## Year2007            0.00317    0.02654    0.12  0.90501
## Year2008           -0.02635    0.02740   -0.96  0.33612
## Year2009           -0.03633    0.02760   -1.32  0.18811
## Year2010           -0.00747    0.02686   -0.28  0.78103
## Year2011           -0.05339    0.02687   -1.99  0.04699 *
## Year2012           -0.10462    0.02779   -3.77  0.00017 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.395
## Multiple R-squared:  0.0101, Adjusted R-squared:  0.00836
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  observation 8501 is an outlier with |weight| = 0 ( < 1e-05);
##  786 weights are ~= 1. The remaining 8841 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.029   0.861   0.951   0.891   0.986   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.04e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                    psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##                   GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.016  1           1.008
## Year             1.016 16           1.001
```



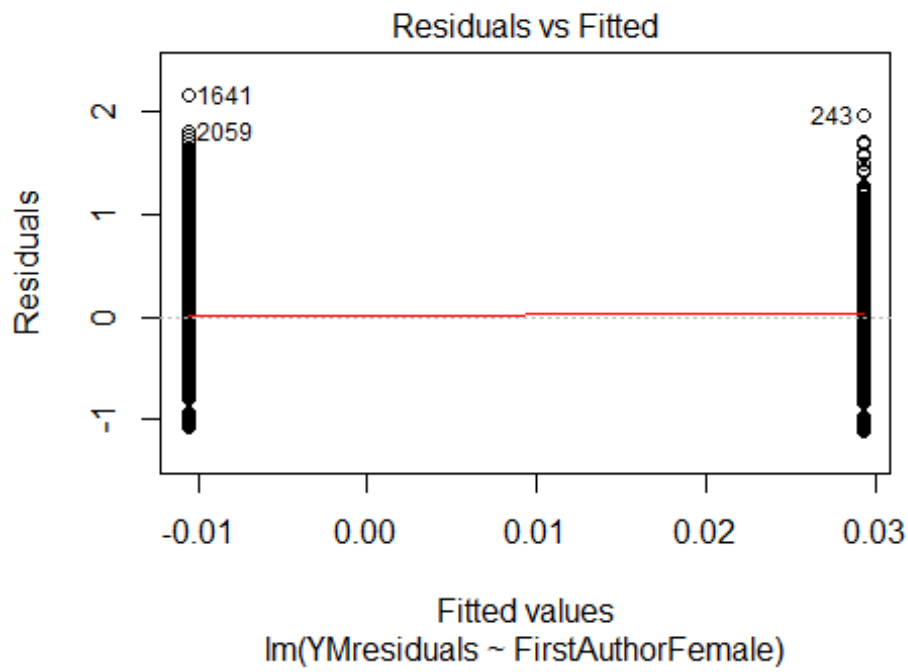

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.19913 -0.26798  0.00808  0.26677  2.09673
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.12669    0.02199   51.24  < 2e-16 ***
## LastAuthorFemale1 -0.01175    0.01107   -1.06  0.28854
## Year1997            0.04988    0.03012    1.66  0.09769 .
## Year1998            0.00995    0.02807    0.35  0.72288
## Year1999            0.01759    0.02901    0.61  0.54436
## Year2000            0.07244    0.03303    2.19  0.02833 *
## Year2001            0.01660    0.03161    0.53  0.59955
## Year2002           -0.00206    0.02934   -0.07  0.94414
## Year2003            0.03524    0.02857    1.23  0.21745
## Year2004            0.01489    0.02873    0.52  0.60418
## Year2005           -0.03400    0.02718   -1.25  0.21107
## Year2006           -0.00398    0.02796   -0.14  0.88676
```

```
## Year2007           0.00283    0.02655    0.11  0.91517
## Year2008          -0.02667    0.02738   -0.97  0.33013
## Year2009          -0.03675    0.02758   -1.33  0.18282
## Year2010          -0.00757    0.02687   -0.28  0.77828
## Year2011          -0.05342    0.02689   -1.99  0.04698 *
## Year2012          -0.10471    0.02781   -3.77  0.00017 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.395
## Multiple R-squared:  0.0101, Adjusted R-squared:  0.00838
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  observation 8501 is an outlier with |weight| = 0 ( < 1e-05);
##  794 weights are ~= 1. The remaining 8833 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0289  0.8610  0.9510  0.8910  0.9860  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.04e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Sample size for the above analysis:  9628"
## [1] ""
## [1] ""
## [1] "####################################"
## [1] "Analysis of AJSC 2310"
## [1] "####################################"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 1092  990 1010  983 1036 1018  937  799  895  874  973 1094 1144 1473 1330
## 2011 2012
## 1436 1395
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  542  558  505  506  500  480  531  479  525  520  626  734  736  958  976
```

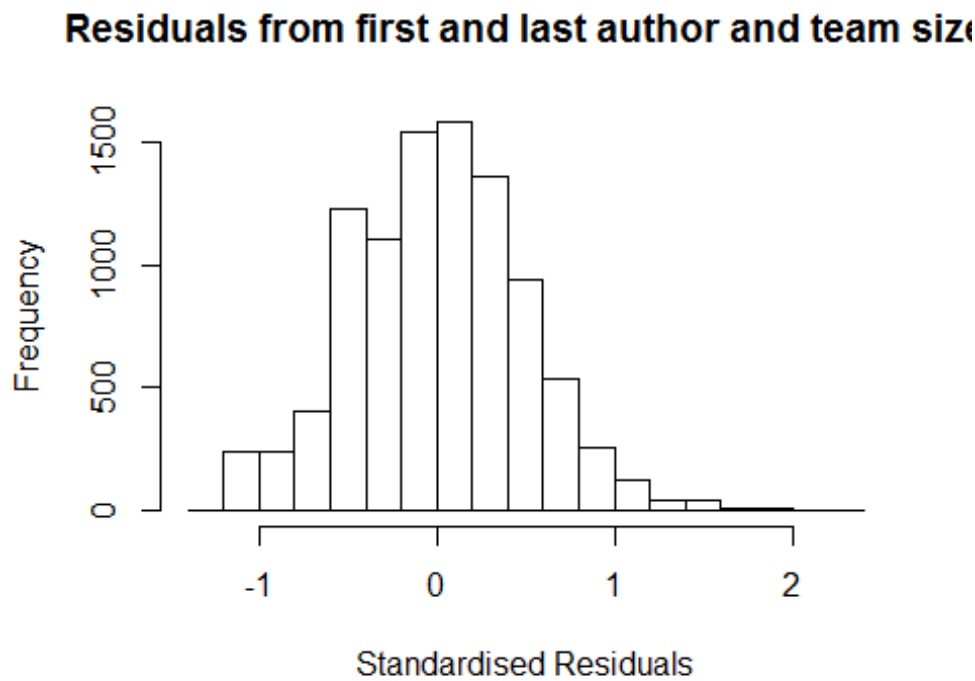
```
## 2011 2012
##  971  940
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  480  495  443  442  448  422  468  416  448  443  553  623  642  844  864
## 2011 2012
##  830  806
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
##  Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 150, df = 16, p-value <2e-16
```




```
##
##  Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 7.1, df = 1, p-value = 0.008
```

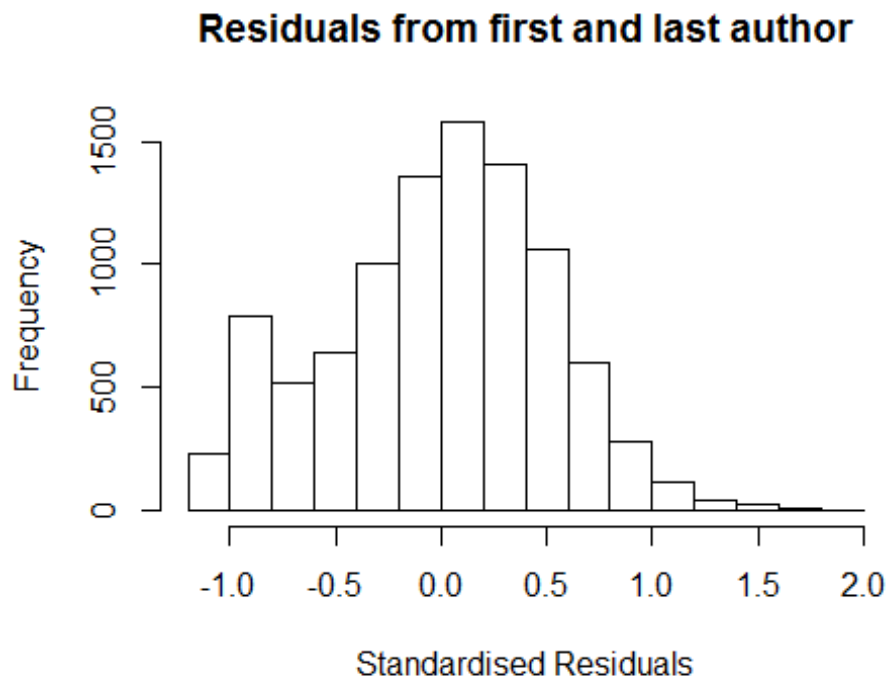



```
## [1] "Female first author team size 2018 geometric mean: 4.10060349801275"
## [1] "Male first author team size 2018 geometric mean: 3.81278058955068"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 120000, p-value = 0.2
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 4.02001753692774"
## [1] "Male last author team size 2018 geometric mean: 3.8772585455803"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 97000, p-value = 0.4
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.045  1           1.022
## LastAuthorFemale  1.030  1           1.015
## UniqueAuthors     1.072  4           1.009
## Year              1.088 16           1.003
```



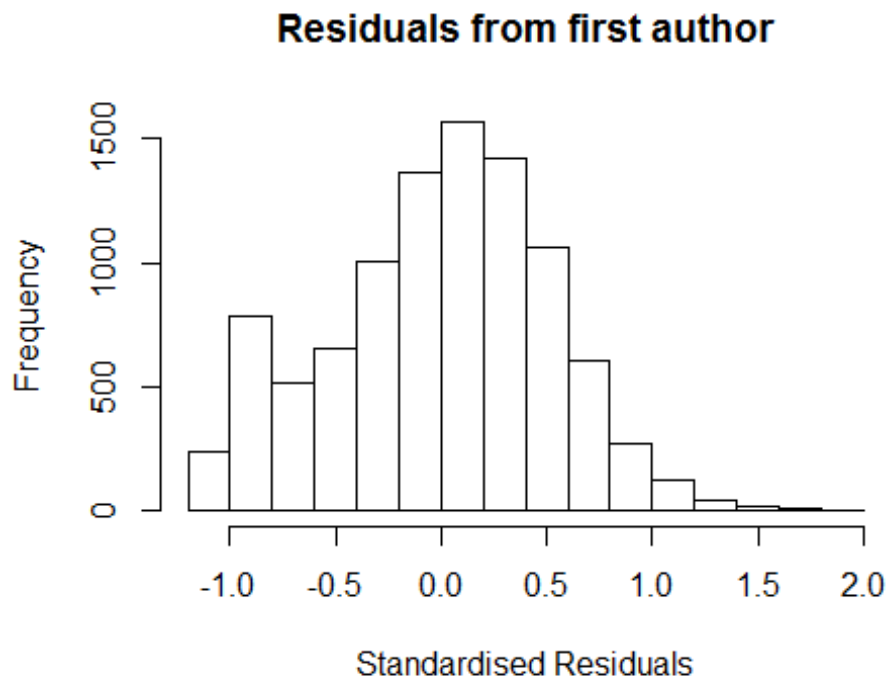

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
UniqueAuthors +
##     Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##     k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.2347 -0.3270  0.0097  0.3231  2.2039
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         0.52663    0.03412   15.44   <2e-16 ***
## FirstAuthorFemale1  0.01932    0.01137    1.70    0.089 .
## LastAuthorFemale1  -0.03342    0.01260   -2.65    0.008 **
## UniqueAuthors2      0.45786    0.01920   23.84   <2e-16 ***
## UniqueAuthors3      0.56790    0.01904   29.83   <2e-16 ***
## UniqueAuthors4      0.57335    0.01956   29.32   <2e-16 ***
## UniqueAuthors5      0.64982    0.01828   35.55   <2e-16 ***
## Year1997            0.04001    0.04207    0.95    0.342
## Year1998            0.01467    0.04141    0.35    0.723
## Year1999           -0.00956    0.04261   -0.22    0.822
```

```
## Year2000            0.05346    0.04088    1.31    0.191
## Year2001            0.00572    0.04003    0.14    0.886
## Year2002            0.02240    0.03900    0.57    0.566
## Year2003            0.05829    0.03894    1.50    0.134
## Year2004            0.05144    0.03745    1.37    0.170
## Year2005            0.03240    0.03726    0.87    0.385
## Year2006           -0.04235    0.03665   -1.16    0.248
## Year2007           -0.06368    0.03608   -1.76    0.078 .
## Year2008           -0.04430    0.03630   -1.22    0.222
## Year2009           -0.06156    0.03461   -1.78    0.075 .
## Year2010           -0.03653    0.03536   -1.03    0.302
## Year2011           -0.02663    0.03599   -0.74    0.459
## Year2012            0.03530    0.03562    0.99    0.322
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.456
## Multiple R-squared:  0.186,  Adjusted R-squared:  0.184
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
##  2 observations c(769,1624) are outliers with |weight| = 0 ( < 1e-05);
##  786 weights are ~= 1. The remaining 8879 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0143  0.8670  0.9430  0.8960  0.9840  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.03e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                    psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.026  1           1.013
## LastAuthorFemale  1.012  1           1.006
## Year              1.037 16           1.001
```



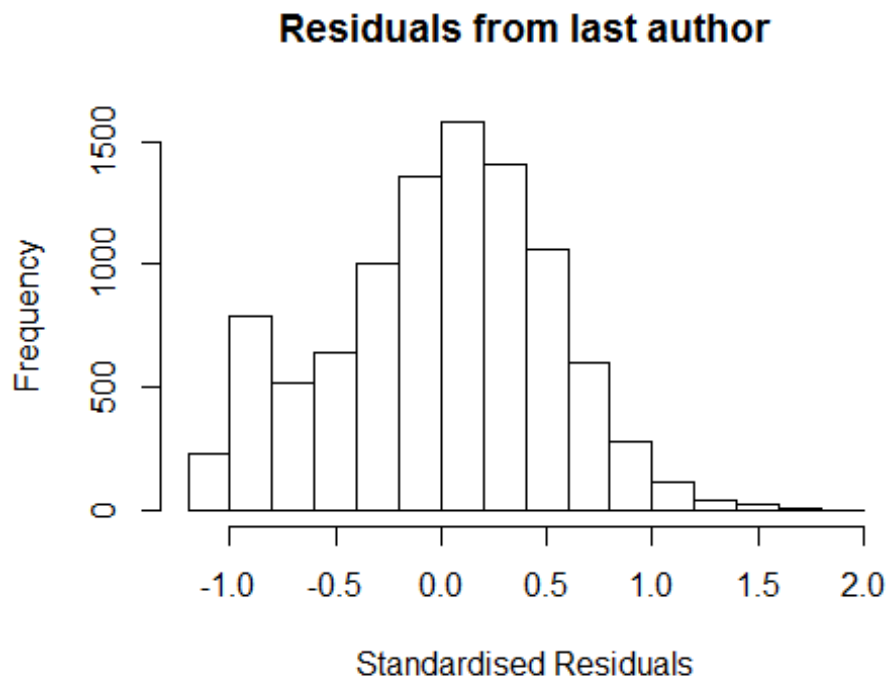

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##    Min     1Q Median     3Q    Max
## -1.121 -0.350  0.036  0.354  1.980
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)         0.926724   0.031903   29.05  < 2e-16 ***
## FirstAuthorFemale1  0.059014   0.012310    4.79  1.7e-06 ***
## LastAuthorFemale1  -0.030554   0.013974   -2.19  0.02881 *
## Year1997            0.044380   0.044590    1.00  0.31962
## Year1998            0.007152   0.044156    0.16  0.87134
## Year1999            0.011693   0.044411    0.26  0.79233
## Year2000            0.069831   0.042509    1.64  0.10047
## Year2001            0.009062   0.042856    0.21  0.83253
## Year2002            0.074216   0.040882    1.82  0.06950 .
## Year2003            0.129730   0.039493    3.28  0.00102 **
## Year2004            0.117564   0.039258    2.99  0.00275 **
## Year2005            0.103551   0.038957    2.66  0.00787 **
```

```
## Year2006            0.022813   0.039161    0.58  0.56023
## Year2007           -0.000397   0.038729   -0.01  0.99182
## Year2008            0.032765   0.037995    0.86  0.38851
## Year2009            0.018458   0.036233    0.51  0.61048
## Year2010            0.031566   0.037464    0.84  0.39950
## Year2011            0.066373   0.037084    1.79  0.07352 .
## Year2012            0.135235   0.037004    3.65  0.00026 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.509
## Multiple R-squared:  0.00956,    Adjusted R-squared:  0.00771
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
##  819 weights are ~= 1. The remaining 8848 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.096   0.852   0.947   0.900   0.985   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.03e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.028  1           1.014
## Year              1.028 16           1.001
```




```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.1112 -0.3484  0.0368  0.3530  1.9579
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)          0.92286    0.03190   28.93  < 2e-16 ***
## FirstAuthorFemale1   0.05425    0.01251    4.34  1.5e-05 ***
## Year1997             0.04414    0.04457    0.99  0.32199
## Year1998             0.00621    0.04417    0.14  0.88820
## Year1999             0.01182    0.04441    0.27  0.79015
## Year2000             0.06898    0.04252    1.62  0.10475
## Year2001             0.00750    0.04284    0.18  0.86103
## Year2002             0.07314    0.04085    1.79  0.07340 .
## Year2003             0.12831    0.03945    3.25  0.00115 **
## Year2004             0.11559    0.03917    2.95  0.00318 **
## Year2005             0.10290    0.03896    2.64  0.00828 **
## Year2006             0.02196    0.03914    0.56  0.57475
```

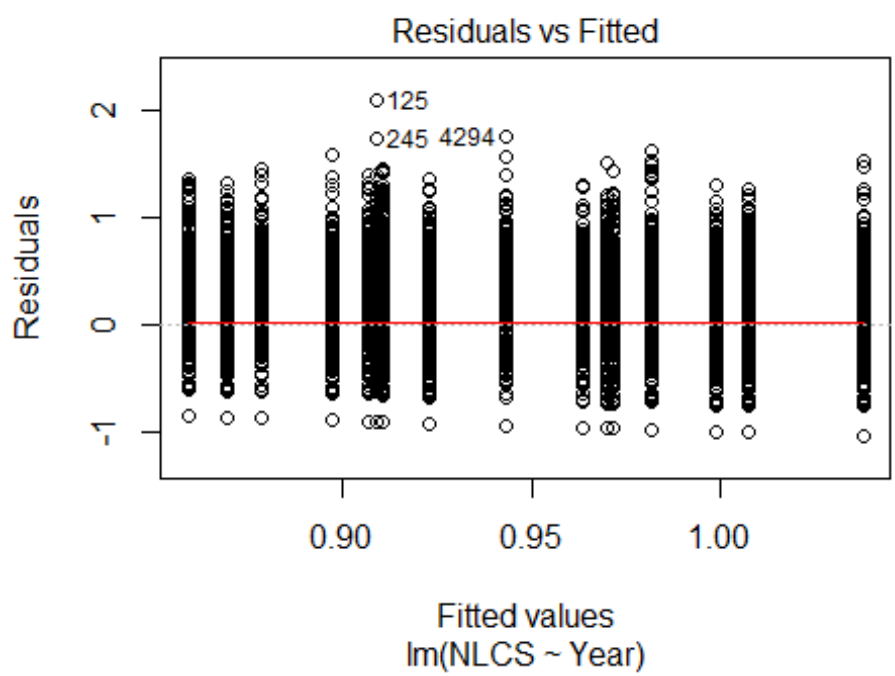
```
## Year2007           -0.00138    0.03867   -0.04  0.97156
## Year2008            0.03184    0.03797    0.84  0.40168
## Year2009            0.01708    0.03616    0.47  0.63664
## Year2010            0.03092    0.03744    0.83  0.40892
## Year2011            0.06445    0.03700    1.74  0.08160 .
## Year2012            0.13412    0.03698    3.63  0.00029 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.509
## Multiple R-squared:  0.00911,    Adjusted R-squared:  0.00737
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
##  820 weights are ~= 1. The remaining 8847 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.106   0.854   0.947   0.900   0.985   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.03e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##                   GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.013  1           1.006
## Year             1.013 16           1.000
```



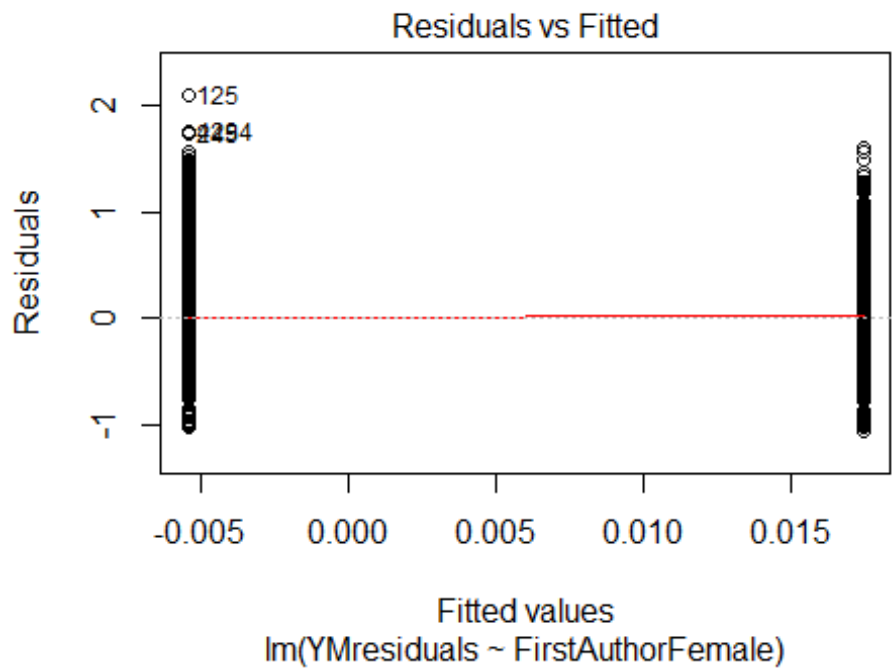

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.0785 -0.3493  0.0402  0.3541  2.0195
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         0.93475    0.03168   29.51  < 2e-16 ***
## LastAuthorFemale1 -0.01921    0.01402   -1.37  0.17053
## Year1997            0.04676    0.04433    1.05  0.29158
## Year1998            0.00885    0.04389    0.20  0.84018
## Year1999            0.01302    0.04432    0.29  0.76887
## Year2000            0.07709    0.04239    1.82  0.06899 .
## Year2001            0.01149    0.04278    0.27  0.78821
## Year2002            0.07936    0.04066    1.95  0.05102 .
## Year2003            0.13144    0.03929    3.35  0.00083 ***
## Year2004            0.12195    0.03905    3.12  0.00180 **
## Year2005            0.10988    0.03877    2.83  0.00460 **
## Year2006            0.03017    0.03897    0.77  0.43883
```

```
## Year2007           0.00915    0.03850    0.24  0.81205
## Year2008           0.04032    0.03775    1.07  0.28549
## Year2009           0.02552    0.03596    0.71  0.47794
## Year2010           0.03863    0.03730    1.04  0.30043
## Year2011           0.07406    0.03685    2.01  0.04447 *
## Year2012           0.14372    0.03676    3.91  9.3e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.509
## Multiple R-squared:  0.00734,    Adjusted R-squared:  0.00559
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
##  805 weights are ~= 1. The remaining 8862 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0804  0.8530  0.9480  0.9010  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.03e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Sample size for the above analysis:  9667"
## [1] ""
## [1] ""
## [1] "####################################"
## [1] "Analysis of AJSC 2311"
## [1] "####################################"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  673  533  601  643  733  583  562  575  508  599  667  669  768  989 1027
## 2011 2012
## 1102 1034
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  416  327  334  345  421  286  347  360  313  400  458  454  514  660  699
## 2011 2012
```

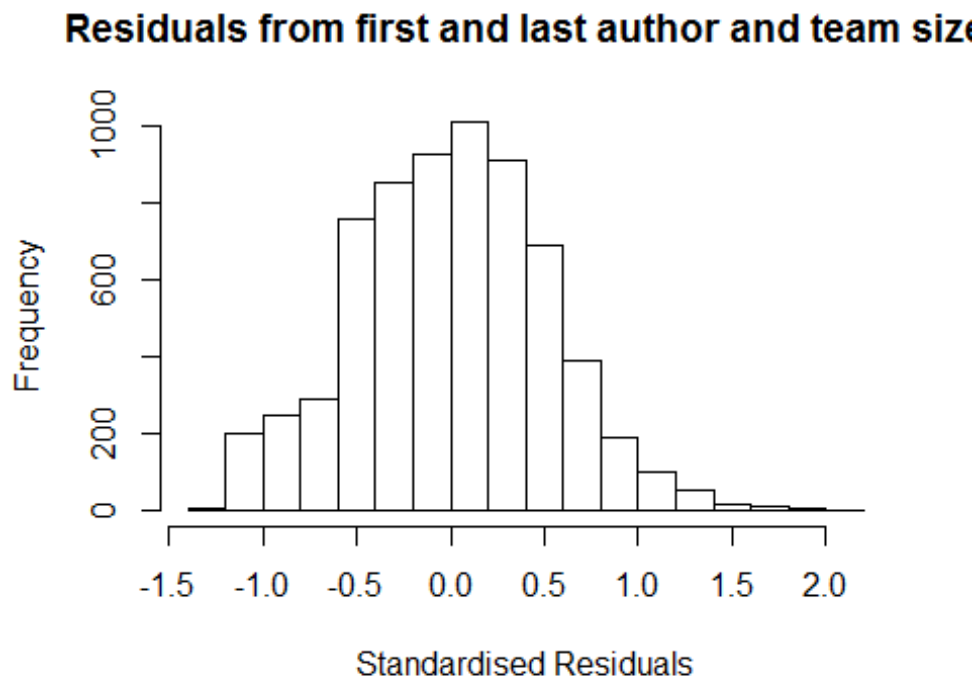
```
##  692  699
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  374  292  289  303  383  252  306  309  275  346  397  384  454  572  592
## 2011 2012
##  566  573
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
##  Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 65, df = 16, p-value = 8e-08
```




```
##
##  Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 1.3, df = 1, p-value = 0.3
```

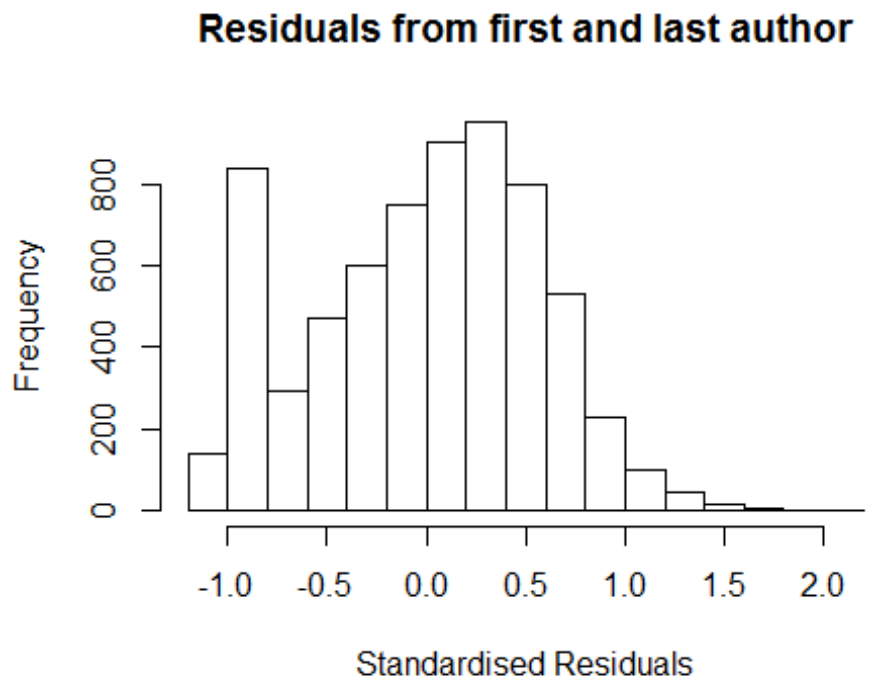



```
## [1] "Female first author team size 2018 geometric mean: 3.98637130127571"
## [1] "Male first author team size 2018 geometric mean: 3.67650204680711"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 46000, p-value = 0.2
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 3.75609276056981"
## [1] "Male last author team size 2018 geometric mean: 3.79409518381926"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 38000, p-value = 0.8
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.080  1           1.039
## LastAuthorFemale  1.075  1           1.037
## UniqueAuthors     1.112  4           1.013
## Year              1.125 16           1.004
```



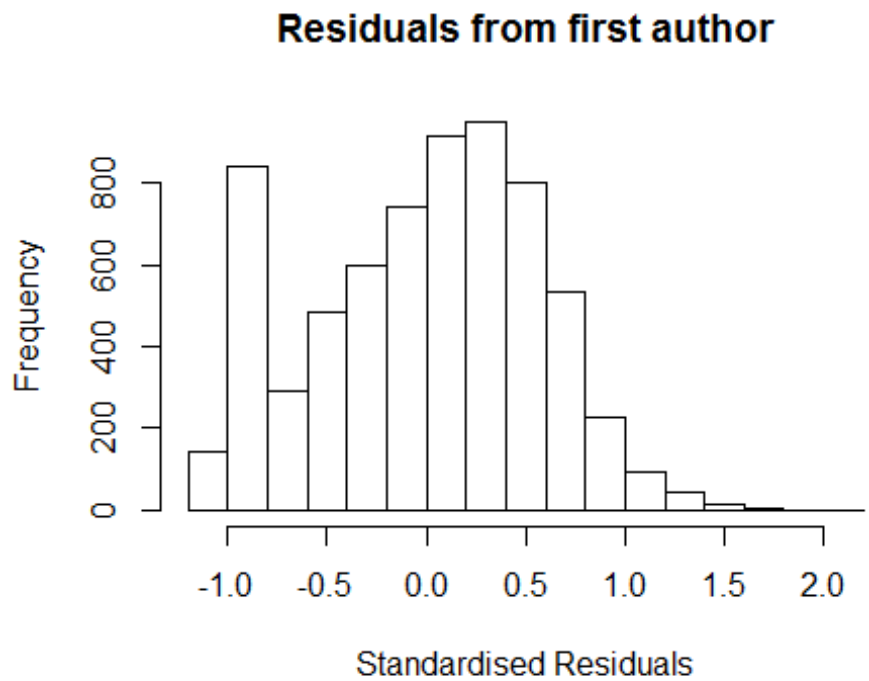

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
UniqueAuthors +
##     Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##     k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.2394 -0.3727  0.0102  0.3535  2.1856
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         0.44643    0.03842   11.62   <2e-16 ***
## FirstAuthorFemale1  0.00550    0.01543    0.36    0.722
## LastAuthorFemale1  -0.02859    0.01721   -1.66    0.097 .
## UniqueAuthors2      0.53083    0.02239   23.71   <2e-16 ***
## UniqueAuthors3      0.63933    0.02158   29.62   <2e-16 ***
## UniqueAuthors4      0.66624    0.02320   28.72   <2e-16 ***
## UniqueAuthors5      0.72869    0.02092   34.84   <2e-16 ***
## Year1997            0.05604    0.05259    1.07    0.287
## Year1998            0.01370    0.04997    0.27    0.784
## Year1999           -0.02773    0.04937   -0.56    0.574
```

```
## Year2000            0.06431    0.04713    1.36    0.172
## Year2001           -0.00485    0.05054   -0.10    0.924
## Year2002           -0.00811    0.04897   -0.17    0.868
## Year2003            0.04228    0.04704    0.90    0.369
## Year2004            0.00553    0.04668    0.12    0.906
## Year2005           -0.01289    0.04401   -0.29    0.770
## Year2006           -0.04711    0.04452   -1.06    0.290
## Year2007           -0.10360    0.04417   -2.35    0.019 *
## Year2008           -0.10610    0.04242   -2.50    0.012 *
## Year2009           -0.04219    0.04124   -1.02    0.306
## Year2010           -0.07370    0.04153   -1.77    0.076 .
## Year2011           -0.00575    0.04204   -0.14    0.891
## Year2012            0.02843    0.04080    0.70    0.486
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.476
## Multiple R-squared:  0.227,  Adjusted R-squared:  0.224
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
##  509 weights are ~= 1. The remaining 6158 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0016  0.8640  0.9400  0.8940  0.9820  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.50e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.026  1           1.013
## LastAuthorFemale  1.019  1           1.009
## Year              1.033 16           1.001
```



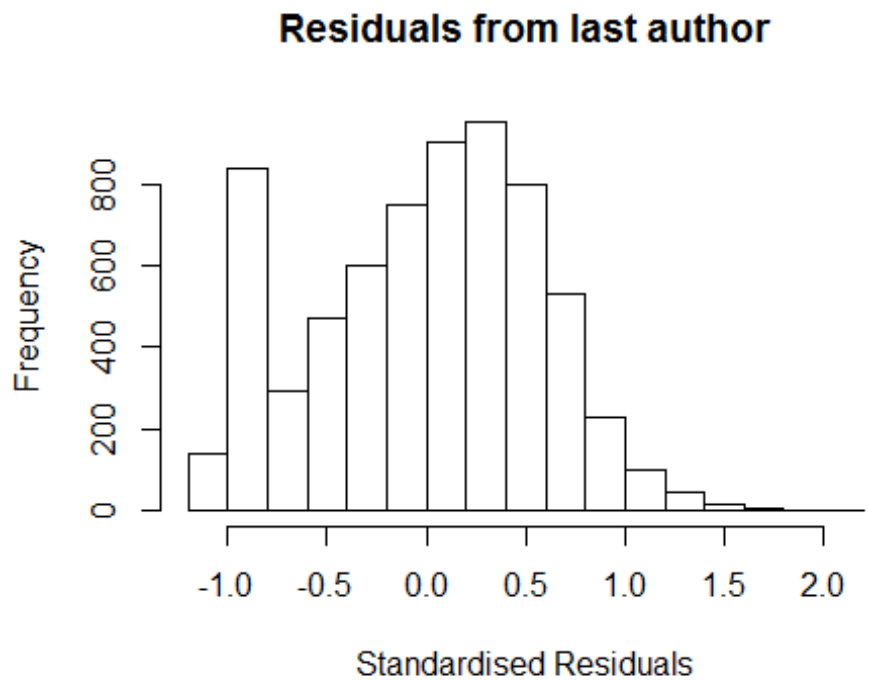

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##    Min     1Q Median     3Q    Max
## -1.068 -0.433  0.055  0.412  2.102
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)          0.89343    0.03532   25.29  < 2e-16 ***
## FirstAuthorFemale1   0.03541    0.01706    2.08  0.03796 *
## LastAuthorFemale1   -0.04747    0.01957   -2.43  0.01533 *
## Year1997             0.05525    0.05461    1.01  0.31167
## Year1998             0.01697    0.05209    0.33  0.74465
## Year1999             0.00326    0.05130    0.06  0.94938
## Year2000             0.07645    0.04738    1.61  0.10665
## Year2001            -0.03127    0.05416   -0.58  0.56370
## Year2002             0.04332    0.05198    0.83  0.40464
## Year2003             0.12508    0.04744    2.64  0.00840 **
## Year2004             0.06301    0.04915    1.28  0.19984
## Year2005             0.05638    0.04550    1.24  0.21537
```

```
## Year2006            0.00698    0.04728    0.15  0.88269
## Year2007           -0.05849    0.04867   -1.20  0.22955
## Year2008           -0.05129    0.04437   -1.16  0.24768
## Year2009            0.02140    0.04209    0.51  0.61112
## Year2010           -0.02298    0.04412   -0.52  0.60248
## Year2011            0.11384    0.04287    2.66  0.00795 **
## Year2012            0.13955    0.04183    3.34  0.00085 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.588
## Multiple R-squared:  0.0119, Adjusted R-squared:  0.00925
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  522 weights are ~= 1. The remaining 6145 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.175   0.860   0.946   0.914   0.984   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.50e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.022  1           1.011
## Year              1.022 16           1.001
```




```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.0517 -0.4378  0.0545  0.4135  2.1077
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)          0.88726    0.03537   25.09  < 2e-16 ***
## FirstAuthorFemale1   0.02591    0.01738    1.49  0.13598
## Year1997             0.05492    0.05466    1.00  0.31500
## Year1998             0.01885    0.05229    0.36  0.71851
## Year1999             0.00246    0.05141    0.05  0.96181
## Year2000             0.07817    0.04745    1.65  0.09956 .
## Year2001            -0.03317    0.05425   -0.61  0.54087
## Year2002             0.04341    0.05206    0.83  0.40444
## Year2003             0.12450    0.04749    2.62  0.00877 **
## Year2004             0.06243    0.04915    1.27  0.20406
## Year2005             0.05716    0.04563    1.25  0.21033
## Year2006             0.00625    0.04729    0.13  0.89484
```

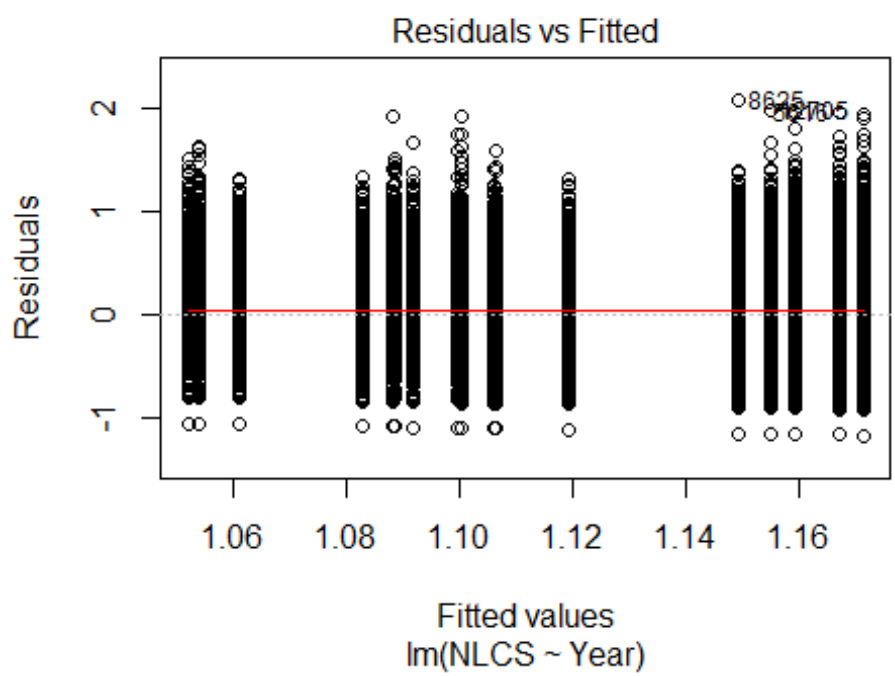
```
## Year2007           -0.05689    0.04875   -1.17  0.24327
## Year2008           -0.05170    0.04444   -1.16  0.24468
## Year2009            0.01966    0.04208    0.47  0.64031
## Year2010           -0.02272    0.04422   -0.51  0.60733
## Year2011            0.11275    0.04290    2.63  0.00861 **
## Year2012            0.13853    0.04188    3.31  0.00095 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.588
## Multiple R-squared:  0.011,  Adjusted R-squared:  0.00852
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  521 weights are ~= 1. The remaining 6146 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.172   0.860   0.946   0.914   0.983   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.50e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##                   GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.013  1           1.006
## Year             1.013 16           1.000
```



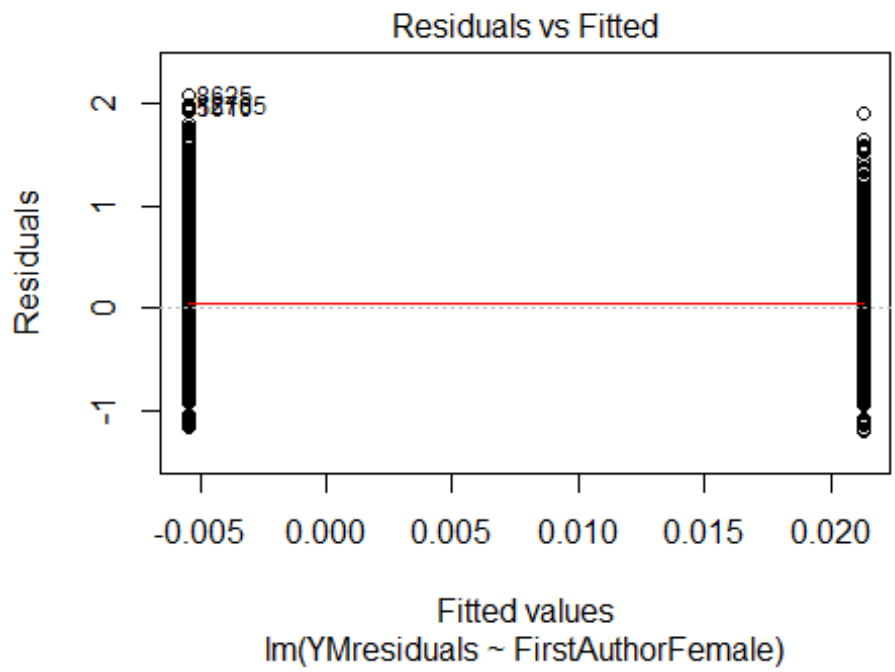

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.0419 -0.4368  0.0543  0.4149  2.0965
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         0.89853    0.03507   25.62  < 2e-16 ***
## LastAuthorFemale1 -0.03874    0.01954   -1.98  0.04744 *
## Year1997            0.05397    0.05436    0.99  0.32081
## Year1998            0.01760    0.05189    0.34  0.73450
## Year1999            0.00272    0.05120    0.05  0.95757
## Year2000            0.07828    0.04720    1.66  0.09724 .
## Year2001           -0.03169    0.05400   -0.59  0.55730
## Year2002            0.04420    0.05179    0.85  0.39341
## Year2003            0.12555    0.04728    2.66  0.00794 **
## Year2004            0.06557    0.04905    1.34  0.18130
## Year2005            0.05898    0.04539    1.30  0.19391
## Year2006            0.00874    0.04719    0.19  0.85305
```

```
## Year2007          -0.05586    0.04860   -1.15  0.25046
## Year2008          -0.04791    0.04417   -1.08  0.27814
## Year2009           0.02310    0.04192    0.55  0.58160
## Year2010          -0.02044    0.04402   -0.46  0.64238
## Year2011           0.11732    0.04271    2.75  0.00603 **
## Year2012           0.14333    0.04165    3.44  0.00058 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.589
## Multiple R-squared:  0.0113, Adjusted R-squared:  0.00881
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  519 weights are ~= 1. The remaining 6148 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.179   0.860   0.945   0.914   0.984   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.50e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Sample size for the above analysis:  6667"
## [1] ""
## [1] ""
## [1] "####################################"
## [1] "Analysis of AJSC 2312"
## [1] "####################################"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 1989 2193 2175 2011 2062 2236 2003 2162 1855 1884 2124 2104 2340 2276 2272
## 2011 2012
## 2273 2411
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 1157 1231 1237 1100  986 1044 1256 1296 1100 1138 1369 1283 1401 1416 1362
## 2011 2012
```

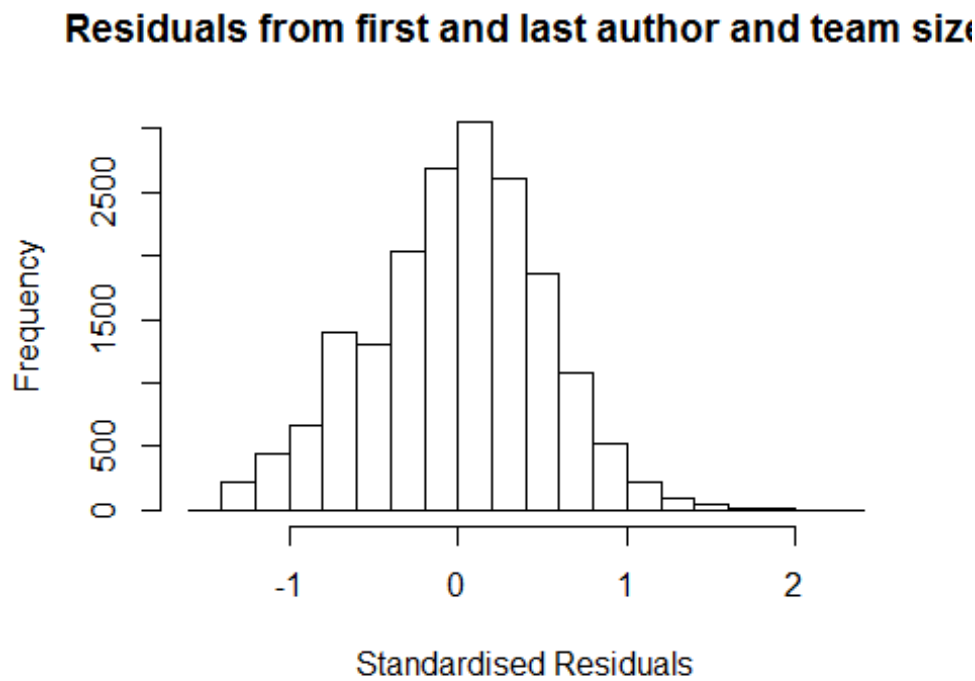
```
## 1331 1446
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 1016 1113 1077  968  870  898 1102 1129  965  974 1147 1072 1192 1212 1159
## 2011 2012
## 1145 1205
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
##  Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 260, df = 16, p-value <2e-16
```




```
##
##  Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 16, df = 1, p-value = 6e-05
```

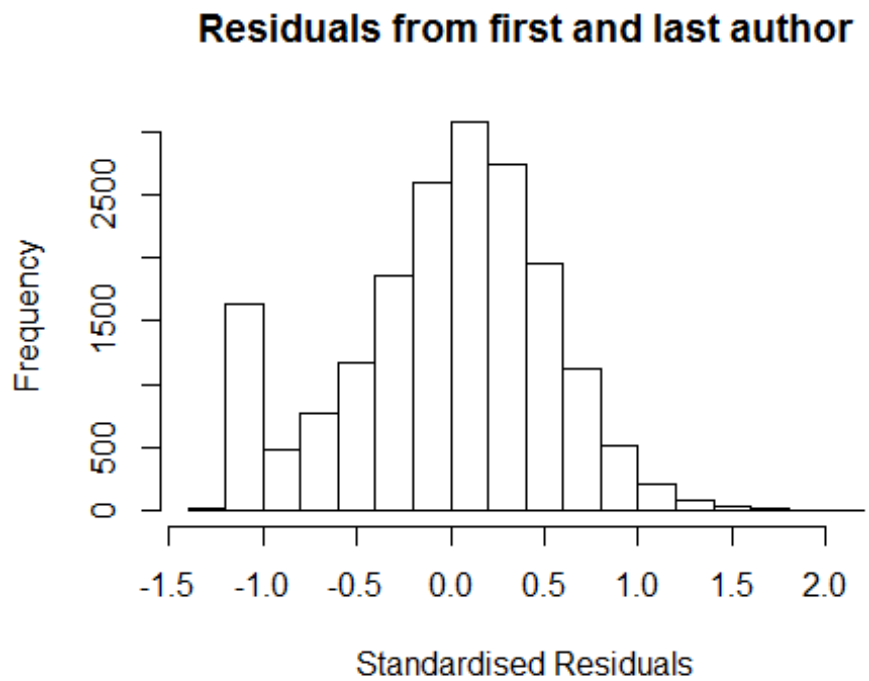



```
## [1] "Female first author team size 2018 geometric mean: 3.53978450962624"
## [1] "Male first author team size 2018 geometric mean: 3.20828370896282"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 240000, p-value = 0.09
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 3.4527075246852"
## [1] "Male last author team size 2018 geometric mean: 3.26234388568244"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 2e+05, p-value = 0.1
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##                     GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.027  1           1.014
## LastAuthorFemale  1.014  1           1.007
## UniqueAuthors     1.052  4           1.006
## Year              1.056 16           1.002
```



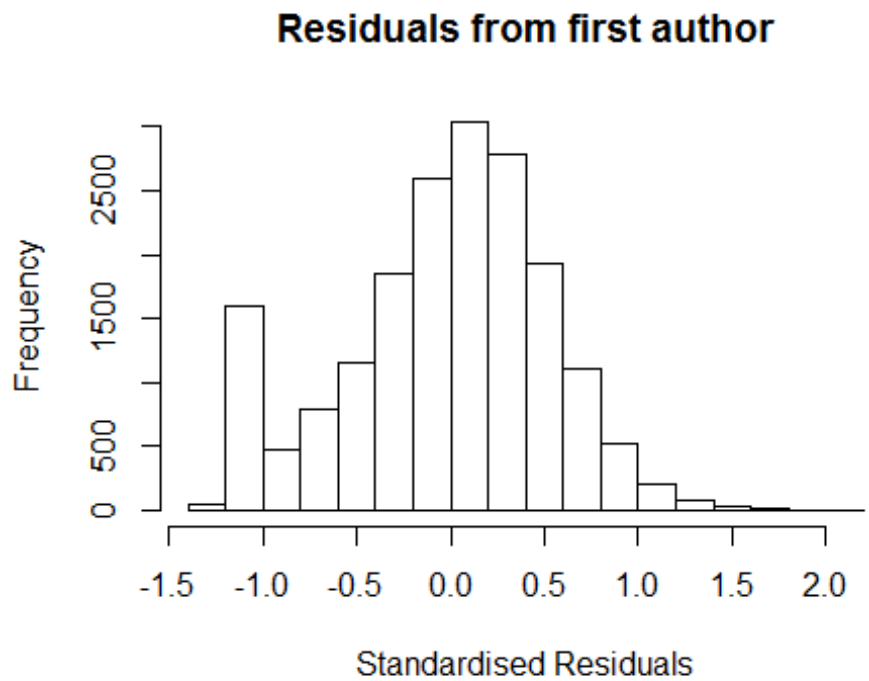

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
UniqueAuthors +
##     Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##     k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.4054 -0.3402  0.0242  0.3372  2.2789
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         0.86834    0.02368   36.68  < 2e-16 ***
## FirstAuthorFemale1 -0.00537    0.00927   -0.58  0.56286
## LastAuthorFemale1  -0.03689    0.01090   -3.38  0.00072 ***
## UniqueAuthors2      0.39478    0.01386   28.49  < 2e-16 ***
## UniqueAuthors3      0.44706    0.01406   31.79  < 2e-16 ***
## UniqueAuthors4      0.48576    0.01505   32.29  < 2e-16 ***
## UniqueAuthors5      0.52675    0.01436   36.67  < 2e-16 ***
## Year1997            0.01030    0.02797    0.37  0.71276
## Year1998           -0.02394    0.02779   -0.86  0.38901
## Year1999           -0.04863    0.02790   -1.74  0.08131 .
```

```
## Year2000           -0.07515    0.02875   -2.61  0.00897 **
## Year2001           -0.06409    0.02820   -2.27  0.02303 *
## Year2002           -0.15231    0.02791   -5.46  4.9e-08 ***
## Year2003           -0.11058    0.02568   -4.31  1.7e-05 ***
## Year2004           -0.13425    0.02653   -5.06  4.2e-07 ***
## Year2005           -0.11692    0.02614   -4.47  7.7e-06 ***
## Year2006           -0.17819    0.02601   -6.85  7.5e-12 ***
## Year2007           -0.17120    0.02594   -6.60  4.3e-11 ***
## Year2008           -0.17859    0.02565   -6.96  3.5e-12 ***
## Year2009           -0.15535    0.02532   -6.14  8.7e-10 ***
## Year2010           -0.13140    0.02571   -5.11  3.2e-07 ***
## Year2011           -0.14737    0.02547   -5.79  7.3e-09 ***
## Year2012           -0.16440    0.02556   -6.43  1.3e-10 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.493
## Multiple R-squared:  0.12,   Adjusted R-squared:  0.119
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
##  1593 weights are ~= 1. The remaining 16651 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0007  0.8490  0.9470  0.8980  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          5.48e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.011  1           1.005
## LastAuthorFemale  1.007  1           1.004
## Year              1.018 16           1.001
```



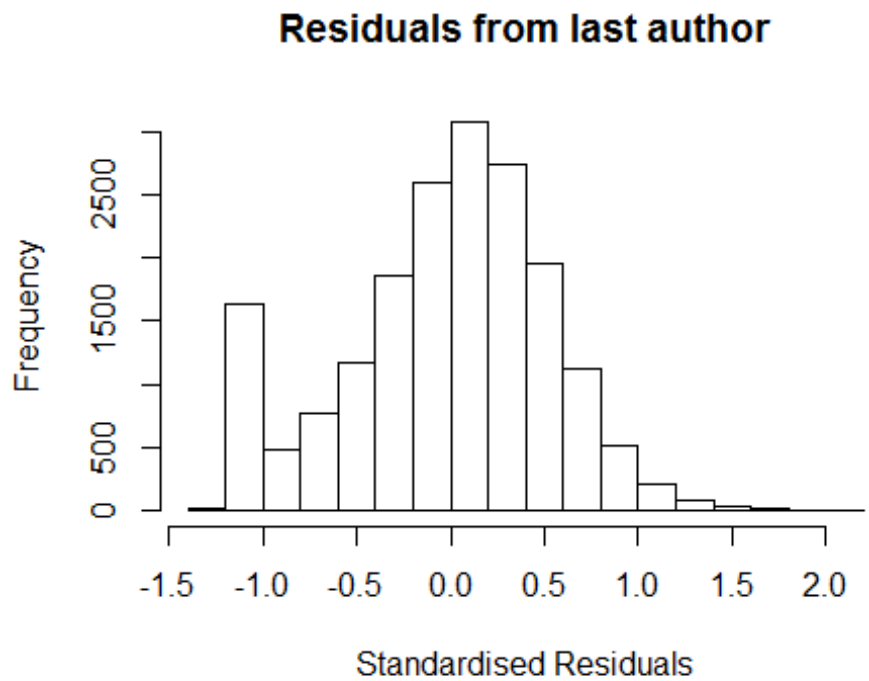

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.2138 -0.3396  0.0383  0.3469  2.0722
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)          1.18384    0.02139   55.35  < 2e-16 ***
## FirstAuthorFemale1  0.03000    0.00978    3.07  0.00215 **
## LastAuthorFemale1  -0.02913    0.01181   -2.47  0.01362 *
## Year1997           -0.00653    0.02871   -0.23  0.81996
## Year1998           -0.00461    0.02845   -0.16  0.87139
## Year1999           -0.03708    0.02836   -1.31  0.19098
## Year2000           -0.06825    0.03050   -2.24  0.02525 *
## Year2001           -0.02082    0.02953   -0.71  0.48073
## Year2002           -0.10826    0.02920   -3.71  0.00021 ***
## Year2003           -0.06514    0.02631   -2.48  0.01328 *
## Year2004           -0.08456    0.02705   -3.13  0.00178 **
## Year2005           -0.06571    0.02699   -2.43  0.01492 *
```

```
## Year2006           -0.12253    0.02698   -4.54  5.6e-06 ***
## Year2007           -0.09974    0.02677   -3.73  0.00020 ***
## Year2008           -0.11620    0.02636   -4.41  1.0e-05 ***
## Year2009           -0.08488    0.02599   -3.27  0.00110 **
## Year2010           -0.05587    0.02608   -2.14  0.03216 *
## Year2011           -0.07299    0.02593   -2.81  0.00489 **
## Year2012           -0.08638    0.02618   -3.30  0.00097 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.5
## Multiple R-squared:  0.00555,    Adjusted R-squared:  0.00456
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
##  1535 weights are ~= 1. The remaining 16709 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0472  0.8500  0.9480  0.8910  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          5.48e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.011  1           1.005
## Year              1.011 16           1.000
```




```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.2068 -0.3401  0.0381  0.3459  2.0759
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)          1.18081    0.02134   55.34  < 2e-16 ***
## FirstAuthorFemale1  0.02597    0.00992    2.62  0.00885 **
## Year1997           -0.00608    0.02870   -0.21  0.83219
## Year1998           -0.00488    0.02843   -0.17  0.86382
## Year1999           -0.03769    0.02834   -1.33  0.18348
## Year2000           -0.06889    0.03050   -2.26  0.02392 *
## Year2001           -0.02119    0.02951   -0.72  0.47259
## Year2002           -0.10851    0.02919   -3.72  0.00020 ***
## Year2003           -0.06568    0.02630   -2.50  0.01253 *
## Year2004           -0.08528    0.02705   -3.15  0.00162 **
## Year2005           -0.06610    0.02696   -2.45  0.01421 *
## Year2006           -0.12325    0.02696   -4.57  4.9e-06 ***
```

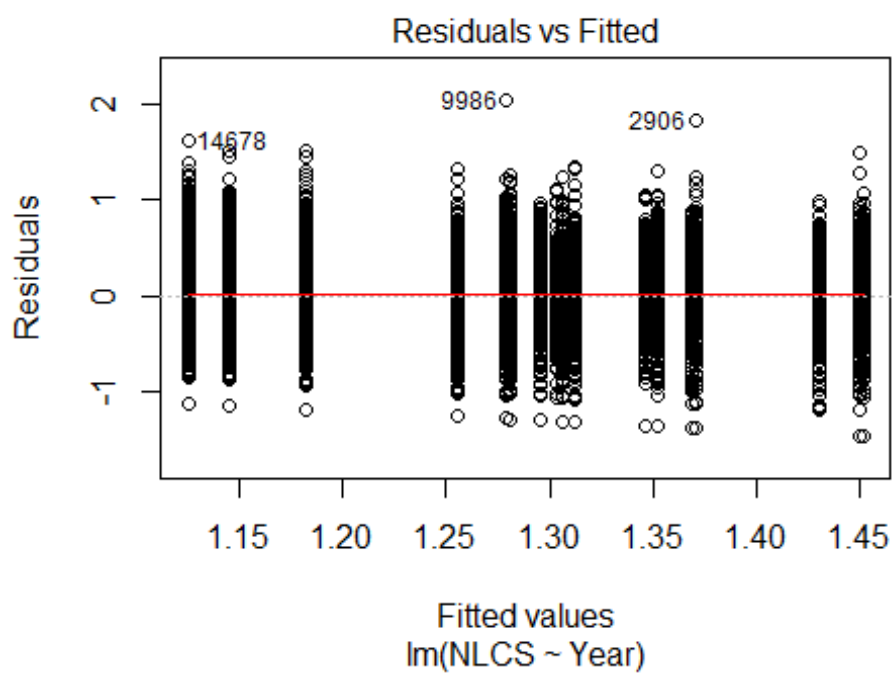
```
## Year2007           -0.09986    0.02677   -3.73  0.00019 ***
## Year2008           -0.11774    0.02633   -4.47  7.8e-06 ***
## Year2009           -0.08561    0.02598   -3.29  0.00099 ***
## Year2010           -0.05707    0.02605   -2.19  0.02846 *
## Year2011           -0.07412    0.02590   -2.86  0.00422 **
## Year2012           -0.08832    0.02613   -3.38  0.00073 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.5
## Multiple R-squared:  0.0052, Adjusted R-squared:  0.00427
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
##  1513 weights are ~= 1. The remaining 16731 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.046   0.850   0.948   0.891   0.985   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          5.48e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##                   GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.007  1           1.004
## Year             1.007 16           1.000
```



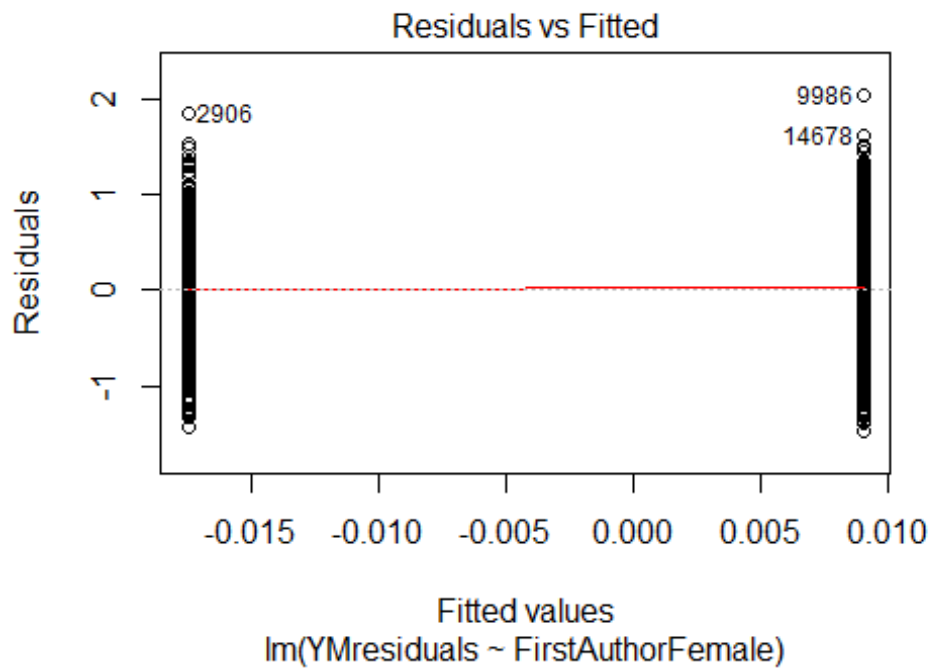

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.1873 -0.3380  0.0393  0.3470  2.0675
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.18730    0.02134   55.63  < 2e-16 ***
## LastAuthorFemale1 -0.02380    0.01185   -2.01  0.04468 *
## Year1997          -0.00579    0.02871   -0.20  0.84025
## Year1998          -0.00392    0.02843   -0.14  0.89030
## Year1999          -0.03578    0.02832   -1.26  0.20644
## Year2000          -0.06657    0.03050   -2.18  0.02905 *
## Year2001          -0.01926    0.02953   -0.65  0.51438
## Year2002          -0.10602    0.02918   -3.63  0.00028 ***
## Year2003          -0.06342    0.02628   -2.41  0.01580 *
## Year2004          -0.08267    0.02705   -3.06  0.00224 **
## Year2005          -0.06289    0.02697   -2.33  0.01971 *
## Year2006          -0.12035    0.02696   -4.46  8.1e-06 ***
```

```
## Year2007          -0.09710    0.02675   -3.63  0.00028 ***
## Year2008          -0.11302    0.02630   -4.30  1.7e-05 ***
## Year2009          -0.08210    0.02595   -3.16  0.00156 **
## Year2010          -0.05327    0.02605   -2.05  0.04087 *
## Year2011          -0.06932    0.02587   -2.68  0.00737 **
## Year2012          -0.08349    0.02614   -3.19  0.00141 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.5
## Multiple R-squared:  0.00507,    Adjusted R-squared:  0.00415
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
##  1539 weights are ~= 1. The remaining 16705 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0488  0.8500  0.9480  0.8900  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          5.48e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Sample size for the above analysis:  18244"
## [1] ""
## [1] ""
## [1] "####################################"
## [1] "Analysis of AJSC 2400"
## [1] "####################################"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  501  496  505  509  598  576  614  587  591  536  651  885  714 1100 1266
## 2011 2012
## 1431 1411
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  278  277  275  315  308  226  372  393  394  347  442  602  467  749  855
## 2011 2012
```

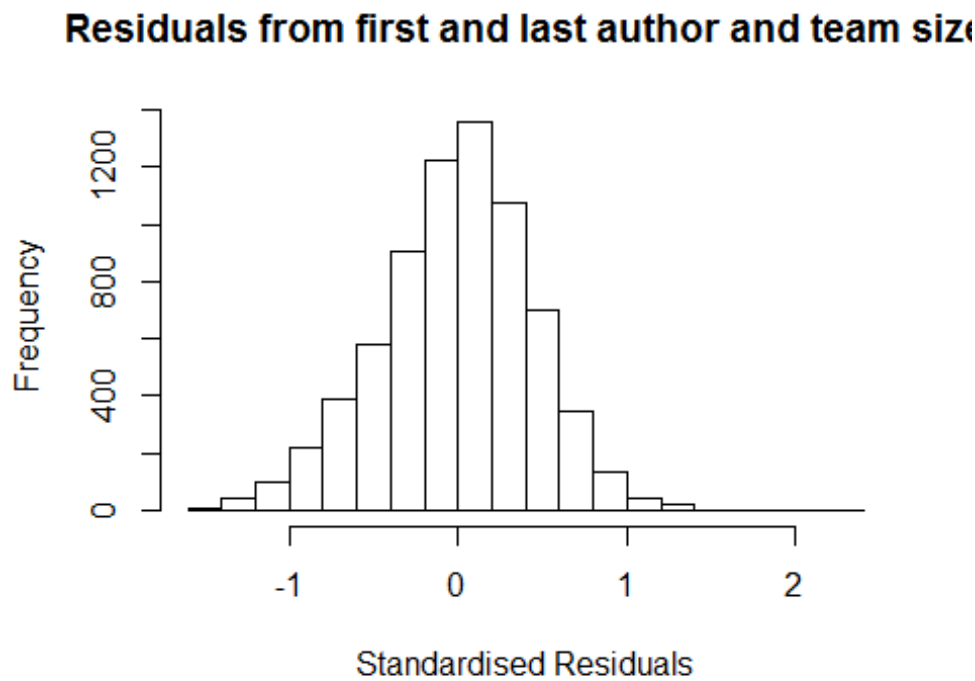
```
##  989  984
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  240  238  247  285  266  199  328  341  346  293  383  507  398  645  729
## 2011 2012
##  842  844
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
##  Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 110, df = 16, p-value <2e-16
```




```
##
##  Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 10, df = 1, p-value = 0.001
```

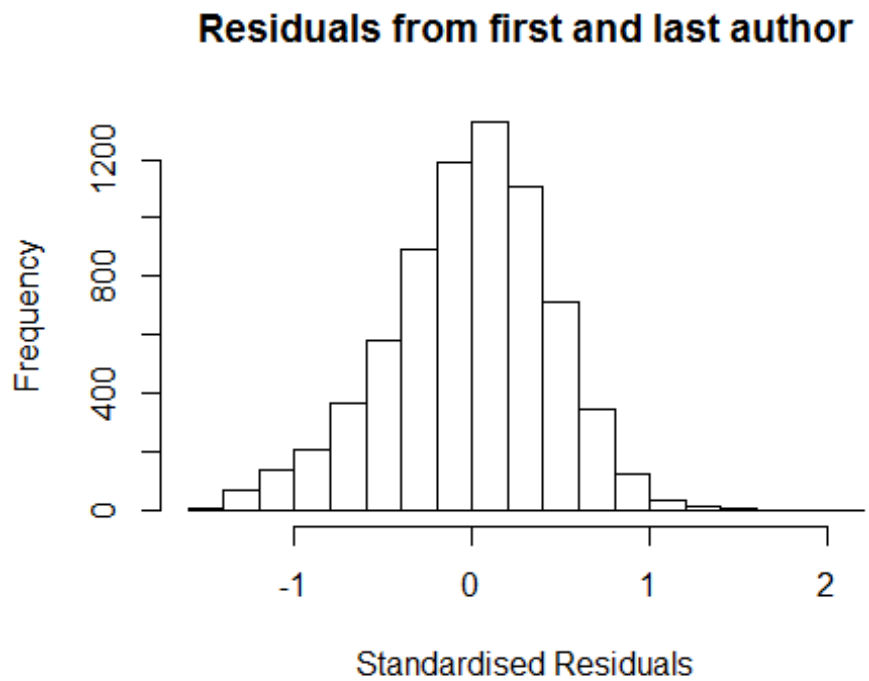



```
## [1] "Female first author team size 2018 geometric mean: 4.5373402676608"
## [1] "Male first author team size 2018 geometric mean: 3.97440787786449"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 180000, p-value = 0.002
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 4.05990496458048"
## [1] "Male last author team size 2018 geometric mean: 4.25677144862631"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 130000, p-value = 0.2
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.068  1           1.034
## LastAuthorFemale  1.030  1           1.015
## UniqueAuthors     1.108  4           1.013
## Year              1.109 16           1.003
```



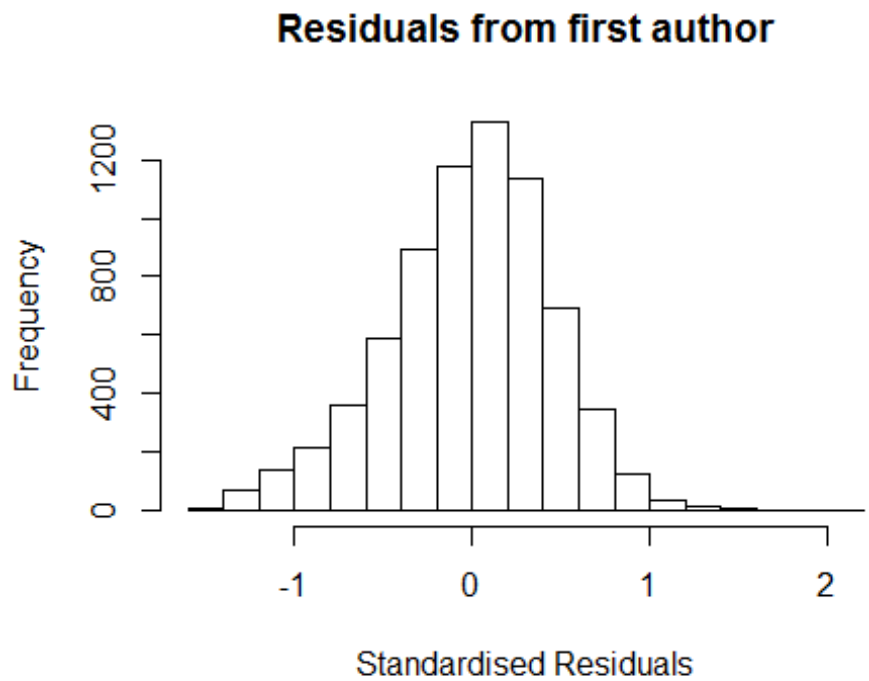

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
UniqueAuthors +
##     Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##     k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.5176 -0.2906  0.0149  0.2926  2.2286
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)          1.3150     0.0374   35.16  < 2e-16 ***
## FirstAuthorFemale1  -0.0479     0.0115   -4.18  2.9e-05 ***
## LastAuthorFemale1   -0.0281     0.0129   -2.18  0.02960 *
## UniqueAuthors2       0.1507     0.0239    6.31  2.9e-10 ***
## UniqueAuthors3       0.2026     0.0240    8.43  < 2e-16 ***
## UniqueAuthors4       0.2293     0.0244    9.41  < 2e-16 ***
## UniqueAuthors5       0.2957     0.0223   13.25  < 2e-16 ***
## Year1997            -0.0285     0.0445   -0.64  0.52146
## Year1998            -0.0341     0.0419   -0.81  0.41584
## Year1999            -0.1200     0.0412   -2.91  0.00362 **
```

```
## Year2000            -0.0937     0.0417   -2.25  0.02453 *
## Year2001            -0.1836     0.0415   -4.42  9.9e-06 ***
## Year2002            -0.1679     0.0401   -4.18  2.9e-05 ***
## Year2003            -0.1691     0.0396   -4.27  2.0e-05 ***
## Year2004            -0.0986     0.0390   -2.52  0.01160 *
## Year2005            -0.1398     0.0406   -3.45  0.00057 ***
## Year2006            -0.1755     0.0396   -4.43  9.4e-06 ***
## Year2007            -0.2023     0.0388   -5.21  2.0e-07 ***
## Year2008            -0.1955     0.0399   -4.90  9.6e-07 ***
## Year2009            -0.2226     0.0364   -6.11  1.0e-09 ***
## Year2010            -0.2980     0.0374   -7.98  1.8e-15 ***
## Year2011            -0.3363     0.0371   -9.06  < 2e-16 ***
## Year2012            -0.3465     0.0372   -9.30  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.431
## Multiple R-squared:  0.0867, Adjusted R-squared:  0.0838
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
##  observation 4461 is an outlier with |weight| = 0 ( < 1.4e-05);
##  653 weights are ~= 1. The remaining 6477 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0162  0.8630  0.9480  0.8990  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.40e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                    psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.041  1           1.020
## LastAuthorFemale  1.018  1           1.009
## Year              1.040 16           1.001
```



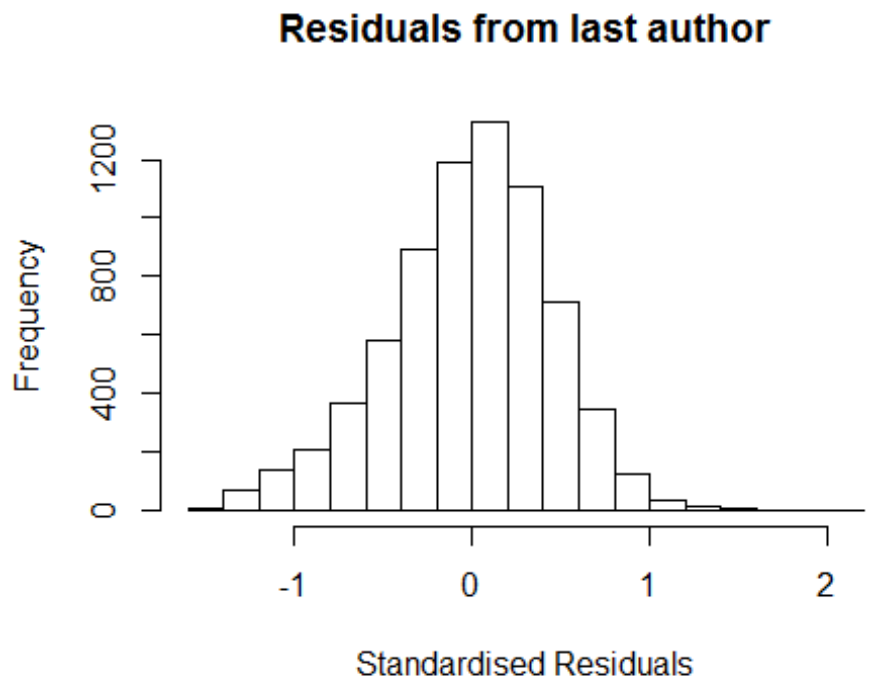

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.4854 -0.3019  0.0176  0.2984  2.0300
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)          1.48536    0.03334   44.55  < 2e-16 ***
## FirstAuthorFemale1 -0.02574    0.01149   -2.24  0.02510 *
## LastAuthorFemale1  -0.02836    0.01309   -2.17  0.03034 *
## Year1997           -0.00208    0.04522   -0.05  0.96329
## Year1998           -0.01784    0.04244   -0.42  0.67431
## Year1999           -0.10979    0.04124   -2.66  0.00778 **
## Year2000           -0.08938    0.04188   -2.13  0.03288 *
## Year2001           -0.16269    0.04225   -3.85  0.00012 ***
## Year2002           -0.14499    0.04010   -3.62  0.00030 ***
## Year2003           -0.14685    0.03982   -3.69  0.00023 ***
## Year2004           -0.07407    0.03937   -1.88  0.05995 .
## Year2005           -0.11830    0.04097   -2.89  0.00390 **
```

```
## Year2006           -0.13984    0.03977   -3.52  0.00044 ***
## Year2007           -0.16823    0.03938   -4.27  2.0e-05 ***
## Year2008           -0.16909    0.04059   -4.17  3.1e-05 ***
## Year2009           -0.19434    0.03680   -5.28  1.3e-07 ***
## Year2010           -0.26015    0.03781   -6.88  6.4e-12 ***
## Year2011           -0.30536    0.03763   -8.11  5.7e-16 ***
## Year2012           -0.31218    0.03786   -8.25  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.436
## Multiple R-squared:  0.0463, Adjusted R-squared:  0.0438
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  614 weights are ~= 1. The remaining 6517 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0001  0.8660  0.9490  0.8980  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.40e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.031  1           1.015
## Year              1.031 16           1.001
```




```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.4818 -0.2982  0.0174  0.2965  2.0353
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.48180    0.03331   44.48  < 2e-16 ***
## FirstAuthorFemale1 -0.02928    0.01145   -2.56  0.01057 *
## Year1997           -0.00248    0.04525   -0.05  0.95627
## Year1998           -0.01928    0.04250   -0.45  0.65008
## Year1999           -0.10957    0.04130   -2.65  0.00799 **
## Year2000           -0.09091    0.04196   -2.17  0.03029 *
## Year2001           -0.16552    0.04228   -3.91  9.1e-05 ***
## Year2002           -0.14555    0.04013   -3.63  0.00029 ***
## Year2003           -0.14984    0.03987   -3.76  0.00017 ***
## Year2004           -0.07481    0.03939   -1.90  0.05759 .
## Year2005           -0.12021    0.04104   -2.93  0.00341 **
## Year2006           -0.14106    0.03984   -3.54  0.00040 ***
```

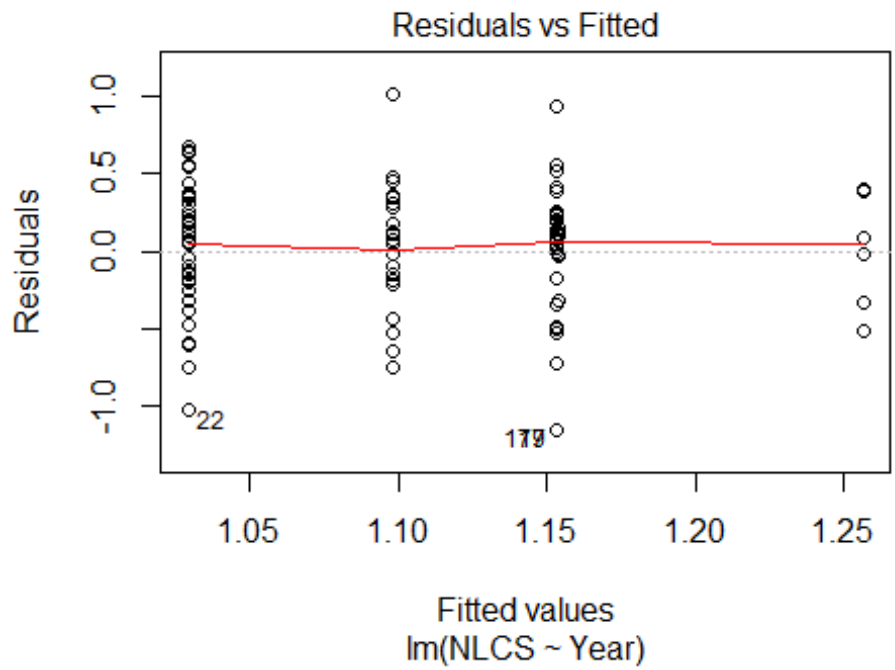
```
## Year2007           -0.17002    0.03944   -4.31  1.6e-05 ***
## Year2008           -0.17121    0.04069   -4.21  2.6e-05 ***
## Year2009           -0.19612    0.03686   -5.32  1.1e-07 ***
## Year2010           -0.26200    0.03787   -6.92  5.0e-12 ***
## Year2011           -0.30772    0.03768   -8.17  3.7e-16 ***
## Year2012           -0.31505    0.03792   -8.31  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.436
## Multiple R-squared:  0.0455, Adjusted R-squared:  0.0432
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  624 weights are ~= 1. The remaining 6507 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.000   0.866   0.949   0.898   0.985   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.40e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##                   GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.009  1           1.004
## Year             1.009 16           1.000
```




```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.4778 -0.3024  0.0166  0.2985  2.0391
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.477818   0.033071   44.69  < 2e-16 ***
## LastAuthorFemale1 -0.032459   0.013058   -2.49  0.01295 *
## Year1997          -0.000634   0.045182   -0.01  0.98881
## Year1998          -0.016432   0.042353   -0.39  0.69805
## Year1999          -0.108419   0.041200   -2.63  0.00852 **
## Year2000          -0.089146   0.041835   -2.13  0.03313 *
## Year2001          -0.160890   0.042181   -3.81  0.00014 ***
## Year2002          -0.145425   0.040071   -3.63  0.00029 ***
## Year2003          -0.145385   0.039745   -3.66  0.00026 ***
## Year2004          -0.073314   0.039333   -1.86  0.06237 .
## Year2005          -0.118391   0.040915   -2.89  0.00382 **
## Year2006          -0.138948   0.039735   -3.50  0.00047 ***
```

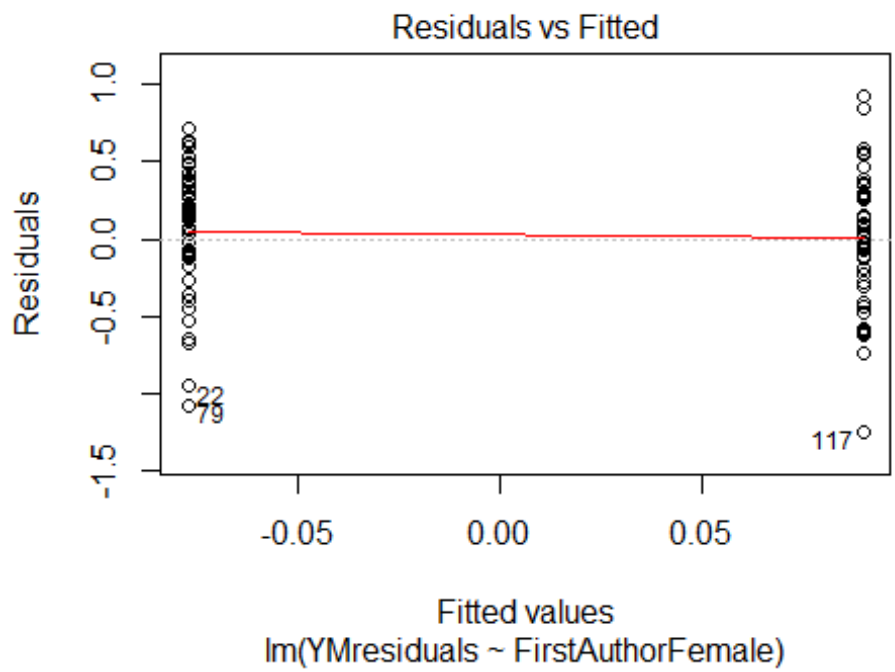
```
## Year2007          -0.169076   0.039357   -4.30  1.8e-05 ***
## Year2008          -0.169302   0.040588   -4.17  3.1e-05 ***
## Year2009          -0.195890   0.036755   -5.33  1.0e-07 ***
## Year2010          -0.262174   0.037762   -6.94  4.2e-12 ***
## Year2011          -0.307613   0.037586   -8.18  3.2e-16 ***
## Year2012          -0.314168   0.037839   -8.30  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.436
## Multiple R-squared:  0.0456, Adjusted R-squared:  0.0433
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  599 weights are ~= 1. The remaining 6532 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.000   0.866   0.949   0.898   0.985   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.40e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Sample size for the above analysis:  7131"
## [1] ""
## [1] ""
## [1] "####################################"
## [1] "Analysis of AJSC 2401"
## [1] "####################################"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 2008 2009 2010 2011 2012
##    6   10   50   41   36
##
## 2008 2009 2010 2011 2012
##    6    6   40   29   27
##
## 2008 2009 2010 2011 2012
##    5    6   37   25   24
```

```
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
##  Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 6.1, df = 4, p-value = 0.2
```

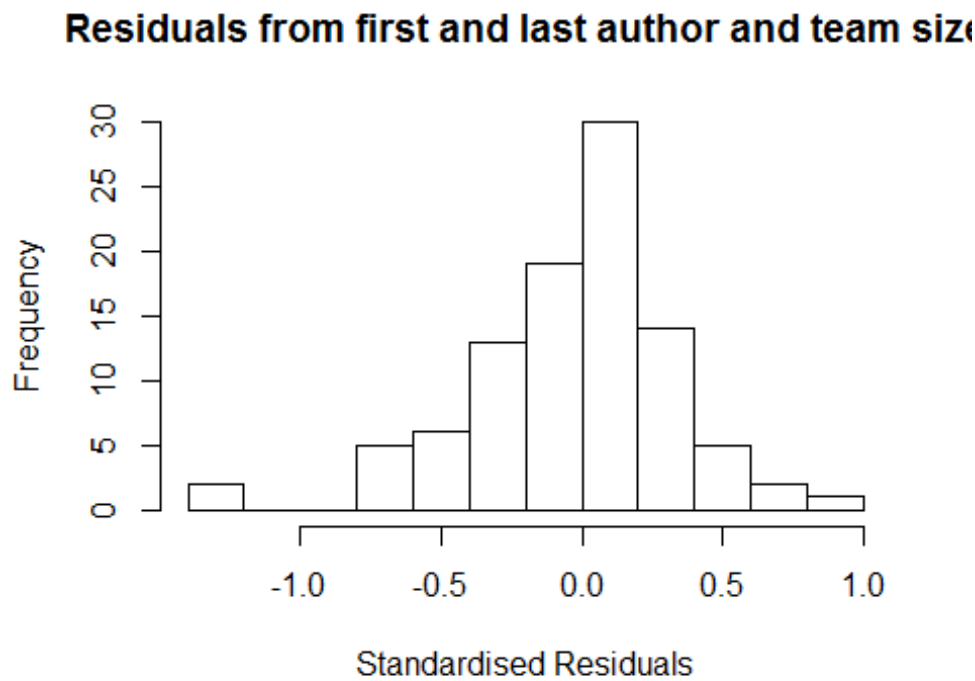



```
##
##  Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 0.016, df = 1, p-value = 0.9

## [1] "Female first author team size 2018 geometric mean: 5.46749528332211"
## [1] "Male first author team size 2018 geometric mean: 5.36770038167723"

## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties

##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 110, p-value = 0.9
## alternative hypothesis: true location shift is not equal to 0
##
```

```
## [1] "Female last author team size 2018 geometric mean: 4.50018397420559"
## [1] "Male last author team size 2018 geometric mean: 6.38670231555652"

## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties
```

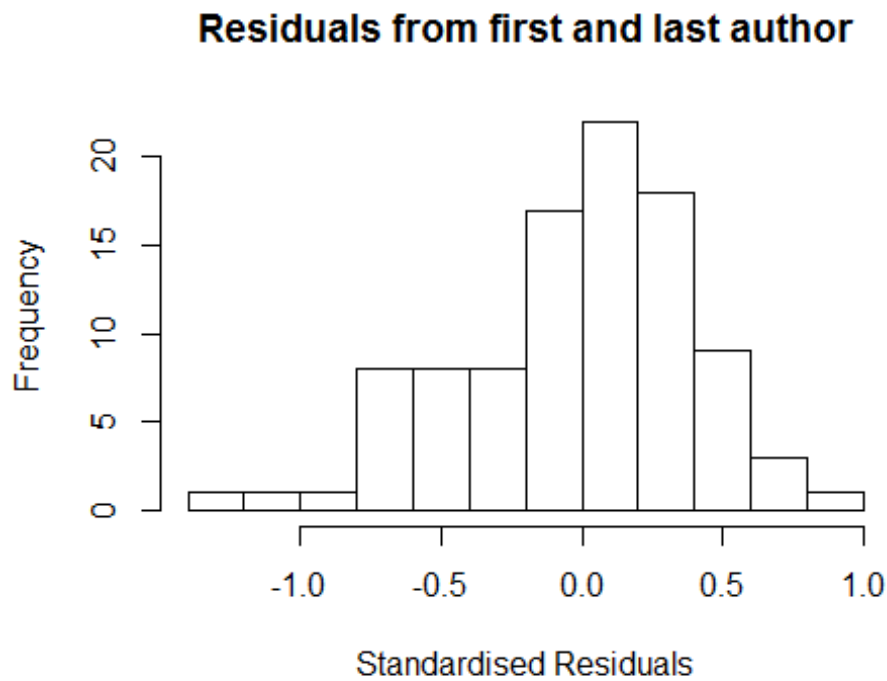



```
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 79, p-value = 0.2
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##                     GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.998  1           1.414
## LastAuthorFemale  1.765  1           1.329
## UniqueAuthors     5.267  4           1.231
## Year              5.282  4           1.231
```



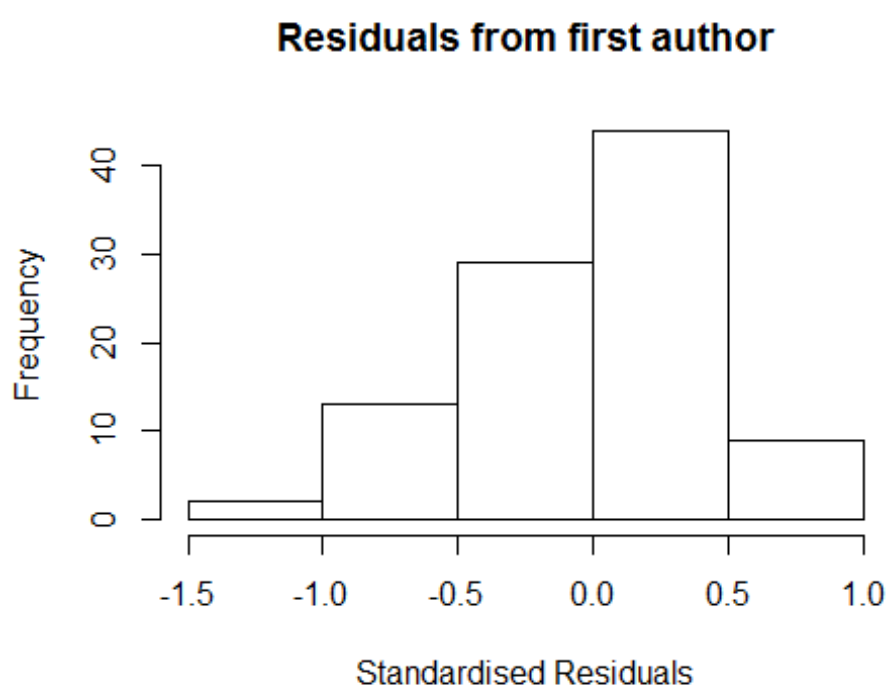

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
UniqueAuthors +
##     Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##     k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.3059 -0.2724  0.0168  0.1860  0.8668
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)          0.6725     0.1706    3.94  0.00016 ***
## FirstAuthorFemale1   0.0514     0.0877    0.59  0.55949
## LastAuthorFemale1   -0.0914     0.0874   -1.05  0.29869
## UniqueAuthors2       0.5383     0.1303    4.13  8.3e-05 ***
## UniqueAuthors3       0.5838     0.2382    2.45  0.01625 *
## UniqueAuthors4       0.5921     0.1232    4.81  6.5e-06 ***
## UniqueAuthors5       0.6928     0.1348    5.14  1.7e-06 ***
## Year2009            -0.1237     0.2014   -0.61  0.54057
## Year2010            -0.0948     0.1700   -0.56  0.57846
## Year2011            -0.0594     0.1684   -0.35  0.72501
```

```
## Year2012            -0.1726     0.1756   -0.98  0.32828
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.323
## Multiple R-squared:  0.351,  Adjusted R-squared:  0.276
## Convergence in 16 IRWLS iterations
##
## Robustness weights:
##  11 weights are ~= 1. The remaining 86 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0655  0.8680  0.9420  0.8820  0.9890  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.03e-03          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                    psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.105  1           1.051
## LastAuthorFemale  1.158  1           1.076
## Year              1.105  4           1.013
```



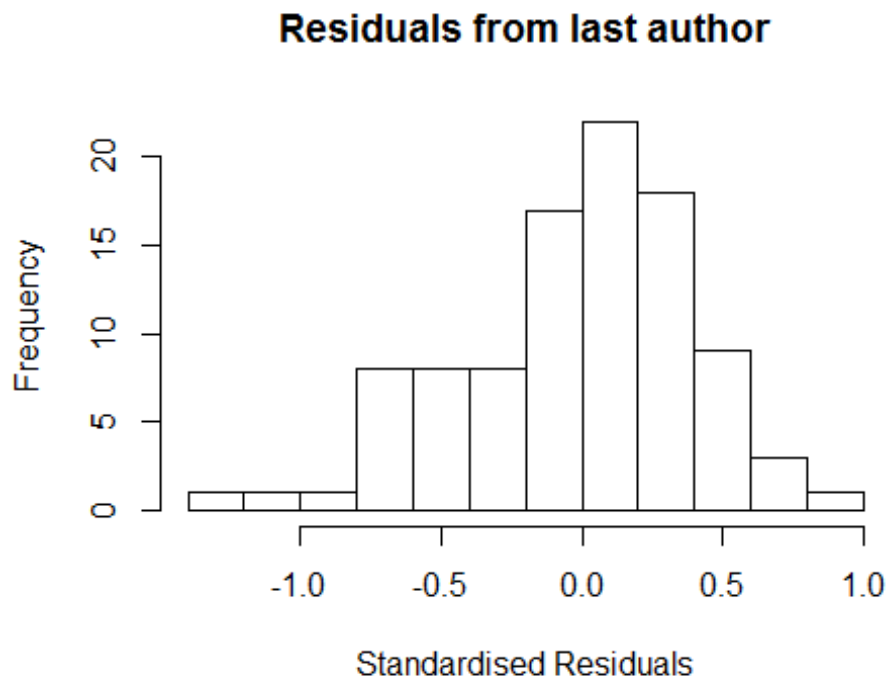

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##    Min     1Q Median     3Q    Max
## -1.271 -0.267  0.023  0.245  0.893
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)           1.0650     0.1648    6.46  5.2e-09 ***
## FirstAuthorFemale1    0.2184     0.0882    2.48    0.015 *
## LastAuthorFemale1    -0.0682     0.0939   -0.73    0.470
## Year2009              0.0176     0.1861    0.09    0.925
## Year2010             -0.1293     0.1782   -0.73    0.470
## Year2011              0.0554     0.1856    0.30    0.766
## Year2012             -0.0650     0.1876   -0.35    0.730
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.402
## Multiple R-squared:  0.0914, Adjusted R-squared:  0.0308
```

```
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  8 weights are ~= 1. The remaining 89 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.296   0.860   0.952   0.903   0.988   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.03e-03          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                  psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.025  1           1.013
## Year              1.025  4           1.003
```



```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.3150 -0.2823  0.0401  0.2491  0.9257
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.06039    0.15690    6.76  1.3e-09 ***
## FirstAuthorFemale1  0.20512    0.08578    2.39    0.019 *
## Year2009           -0.00335    0.17768   -0.02    0.985
## Year2010           -0.13085    0.16968   -0.77    0.443
## Year2011            0.04950    0.17594    0.28    0.779
## Year2012           -0.08019    0.17535   -0.46    0.649
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.395
## Multiple R-squared:  0.0886, Adjusted R-squared:  0.0386
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  9 weights are ~= 1. The remaining 88 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.245   0.853   0.954   0.898   0.988   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.03e-03          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##                   GVIF Df GVIF^(1/(2*Df))
```

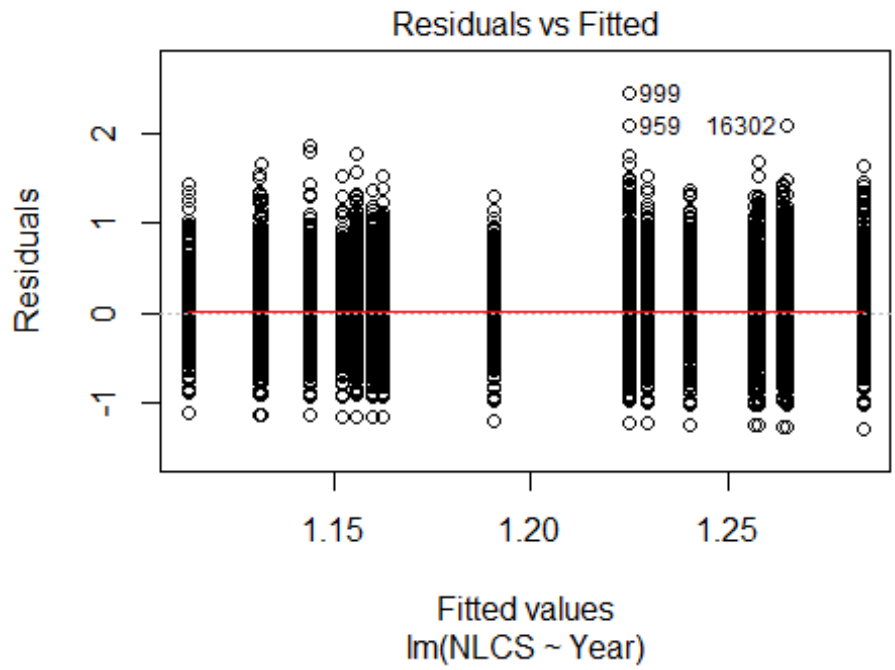
```
## LastAuthorFemale 1.101  1           1.049
## Year             1.101  4           1.012
```



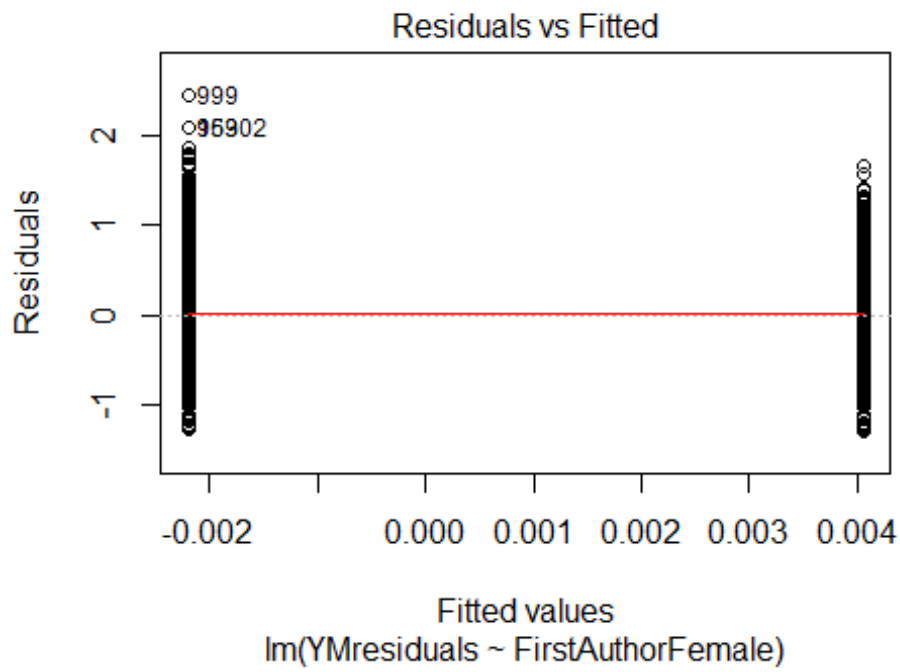

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.1969 -0.2269  0.0531  0.2161  0.9991
##
## Coefficients:
##                   Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.1772     0.1535    7.67  1.8e-11 ***
## LastAuthorFemale1  -0.0148     0.0904   -0.16     0.87
## Year2009           -0.0136     0.1723   -0.08     0.94
## Year2010           -0.1553     0.1774   -0.88     0.38
## Year2011            0.0197     0.1785    0.11     0.91
## Year2012           -0.0652     0.1790   -0.36     0.72
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
```

```
## Robust residual standard error: 0.4
## Multiple R-squared:  0.0292, Adjusted R-squared:  -0.0242
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  8 weights are ~= 1. The remaining 89 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.350   0.850   0.962   0.896   0.988   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.03e-03          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Sample size for the above analysis:  97"
## [1] ""
## [1] ""
## [1] "##################################"
## [1] "Analysis of AJSC 2402"
## [1] "##################################"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  941  995 1020 1015 1043  953 1004  855  952 1005 1028 1147 1109  975 1000
## 2011 2012
## 1030 1076
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  505  542  478  571  421  352  602  512  596  632  635  726  677  606  610
## 2011 2012
##  634  638
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  452  490  426  506  371  314  525  448  527  544  561  633  584  506  486
## 2011 2012
##  517  549
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
```

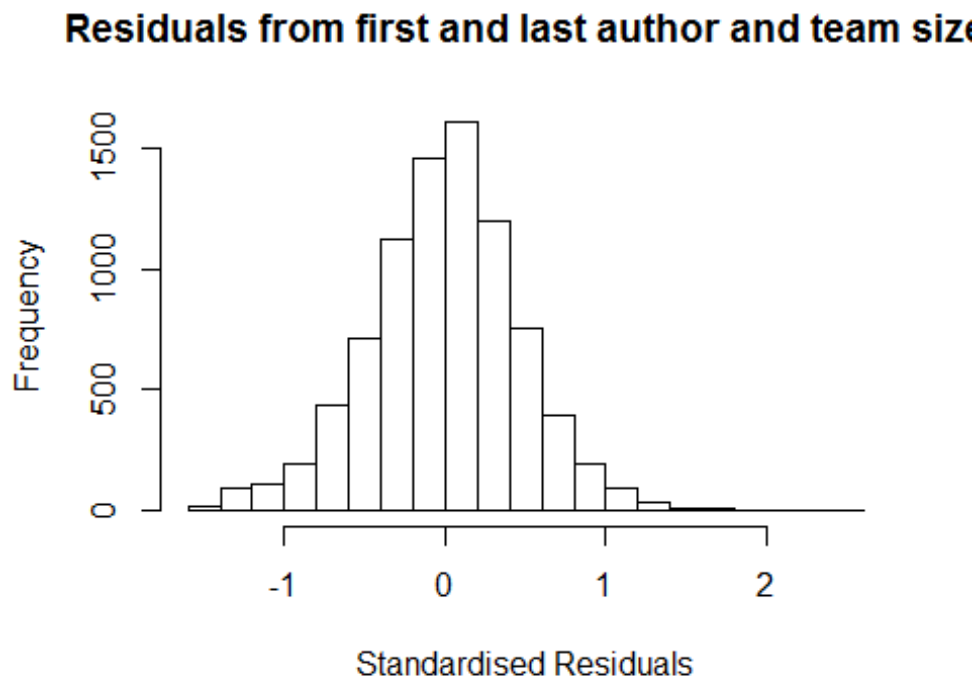
```
##  Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 280, df = 16, p-value <2e-16
```




```
##
##  Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 63, df = 1, p-value = 2e-15
```

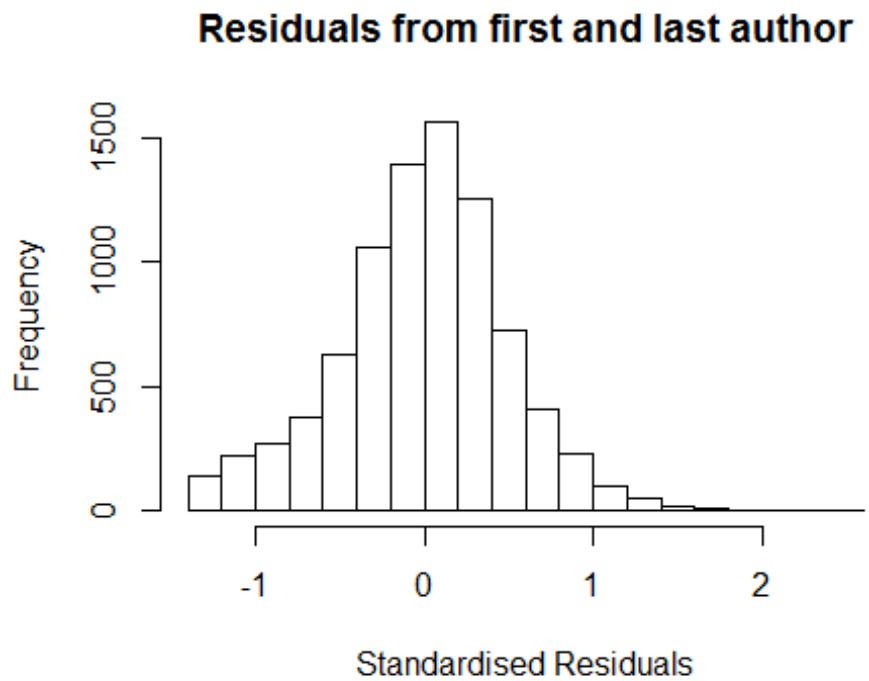



```
## [1] "Female first author team size 2018 geometric mean: 4.54938151518464"
## [1] "Male first author team size 2018 geometric mean: 4.66407708123406"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 21000, p-value = 0.5
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 4.13301543727582"
## [1] "Male last author team size 2018 geometric mean: 4.79893562153624"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 15000, p-value = 0.08
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.034  1           1.017
## LastAuthorFemale  1.035  1           1.017
## UniqueAuthors     1.126  4           1.015
## Year              1.144 16           1.004
```



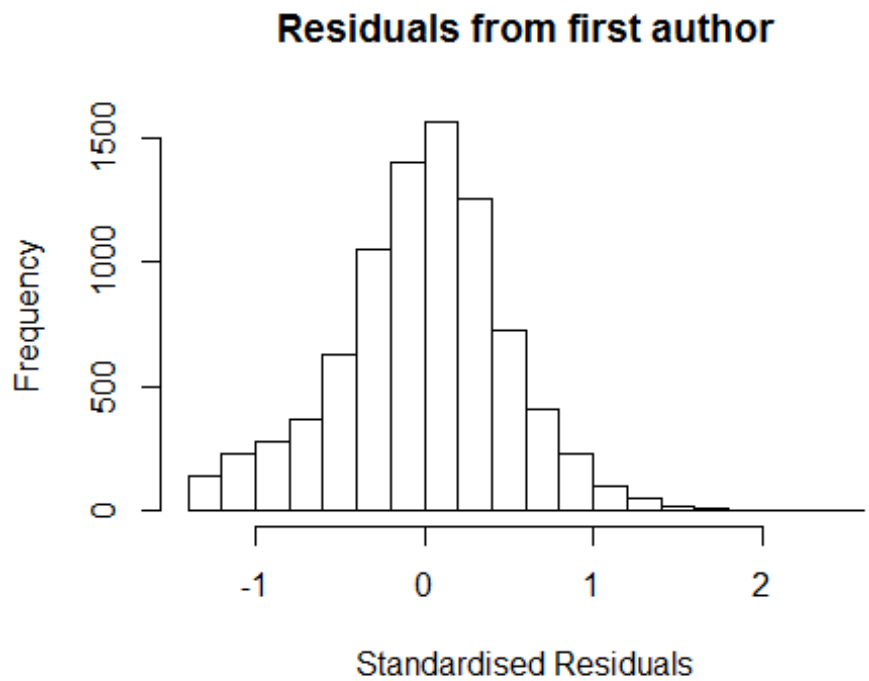

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
UniqueAuthors +
##     Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##     k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.4821 -0.2983  0.0108  0.2885  2.4260
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         0.74419    0.03724   19.98  < 2e-16 ***
## FirstAuthorFemale1 -0.01701    0.01035   -1.64   0.1002
## LastAuthorFemale1  -0.01643    0.01236   -1.33   0.1838
## UniqueAuthors2      0.50477    0.02618   19.28  < 2e-16 ***
## UniqueAuthors3      0.57459    0.02591   22.18  < 2e-16 ***
## UniqueAuthors4      0.61708    0.02657   23.22  < 2e-16 ***
## UniqueAuthors5      0.73791    0.02588   28.51  < 2e-16 ***
## Year1997           -0.00993    0.04312   -0.23   0.8179
## Year1998           -0.02242    0.03865   -0.58   0.5618
## Year1999           -0.09349    0.03772   -2.48   0.0132 *
```

```
## Year2000           -0.03807    0.04054   -0.94   0.3476
## Year2001           -0.10169    0.04133   -2.46   0.0139 *
## Year2002           -0.12391    0.03804   -3.26   0.0011 **
## Year2003           -0.10005    0.03746   -2.67   0.0076 **
## Year2004           -0.15296    0.03649   -4.19  2.8e-05 ***
## Year2005           -0.19227    0.03629   -5.30  1.2e-07 ***
## Year2006           -0.19062    0.03569   -5.34  9.5e-08 ***
## Year2007           -0.17856    0.03609   -4.95  7.7e-07 ***
## Year2008           -0.08933    0.03705   -2.41   0.0159 *
## Year2009           -0.04516    0.03786   -1.19   0.2329
## Year2010           -0.07868    0.03806   -2.07   0.0387 *
## Year2011           -0.06616    0.03881   -1.70   0.0883 .
## Year2012           -0.07413    0.03797   -1.95   0.0509 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.43
## Multiple R-squared:  0.165,  Adjusted R-squared:  0.163
## Convergence in 15 IRWLS iterations
##
## Robustness weights:
##  2 observations c(420,2543) are outliers with |weight| = 0 ( < 1.2e-05);
##  709 weights are ~= 1. The remaining 7728 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0125  0.8590  0.9490  0.8930  0.9860  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.18e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.014  1           1.007
## LastAuthorFemale  1.021  1           1.011
## Year              1.031 16           1.001
```



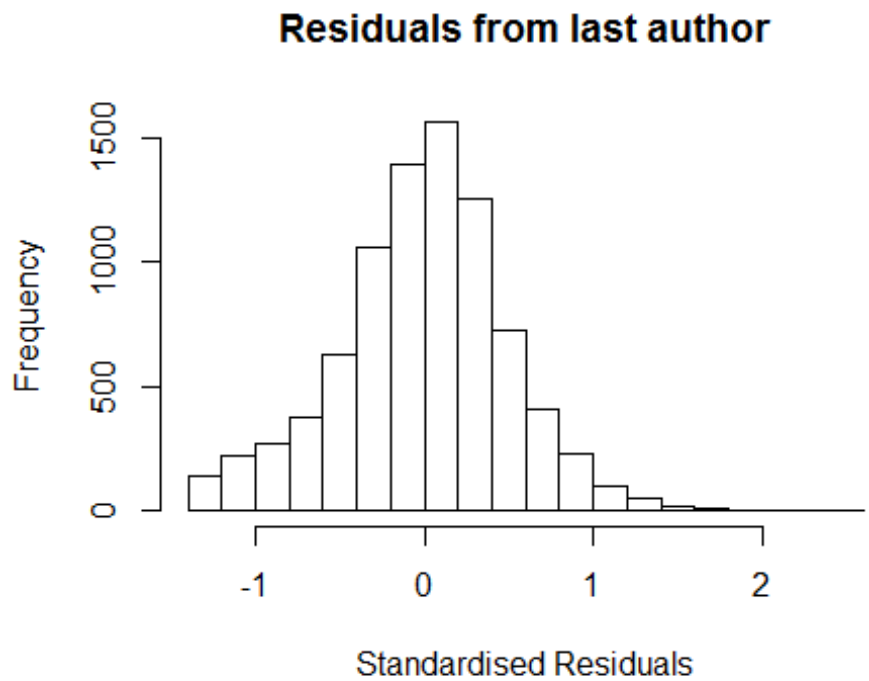

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.2951 -0.3038  0.0154  0.2990  2.4512
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)          1.22380    0.03780   32.38   <2e-16 ***
## FirstAuthorFemale1   0.01288    0.01093    1.18    0.238
## LastAuthorFemale1   -0.01866    0.01344   -1.39    0.165
## Year1997            -0.04292    0.05113   -0.84    0.401
## Year1998             0.03518    0.04386    0.80    0.422
## Year1999            -0.03431    0.04375   -0.78    0.433
## Year2000             0.00367    0.04548    0.08    0.936
## Year2001            -0.06205    0.04732   -1.31    0.190
## Year2002            -0.05423    0.04348   -1.25    0.212
## Year2003            -0.01113    0.04277   -0.26    0.795
## Year2004            -0.06783    0.04215   -1.61    0.108
## Year2005            -0.09403    0.04211   -2.23    0.026 *
```

```
## Year2006           -0.09723    0.04153   -2.34    0.019 *
## Year2007           -0.09035    0.04151   -2.18    0.030 *
## Year2008            0.01396    0.04234    0.33    0.742
## Year2009            0.05323    0.04324    1.23    0.218
## Year2010            0.05121    0.04322    1.18    0.236
## Year2011            0.05846    0.04449    1.31    0.189
## Year2012            0.05594    0.04353    1.29    0.199
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.448
## Multiple R-squared:  0.0139, Adjusted R-squared:  0.0118
## Convergence in 18 IRWLS iterations
##
## Robustness weights:
##  2 observations c(393,420) are outliers with |weight| = 0 ( < 1.2e-05);
##  754 weights are ~= 1. The remaining 7683 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.046   0.853   0.950   0.886   0.986   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.18e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.012  1           1.006
## Year              1.012 16           1.000
```




```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.2903 -0.3027  0.0154  0.3003  2.4529
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.22213    0.03787   32.27   <2e-16 ***
## FirstAuthorFemale1  0.01111    0.01094    1.02    0.310
## Year1997           -0.04377    0.05115   -0.86    0.392
## Year1998            0.03477    0.04389    0.79    0.428
## Year1999           -0.03560    0.04375   -0.81    0.416
## Year2000            0.00272    0.04549    0.06    0.952
## Year2001           -0.06371    0.04732   -1.35    0.178
## Year2002           -0.05576    0.04347   -1.28    0.200
## Year2003           -0.01241    0.04279   -0.29    0.772
## Year2004           -0.06974    0.04211   -1.66    0.098 .
## Year2005           -0.09530    0.04210   -2.26    0.024 *
## Year2006           -0.09870    0.04152   -2.38    0.017 *
```

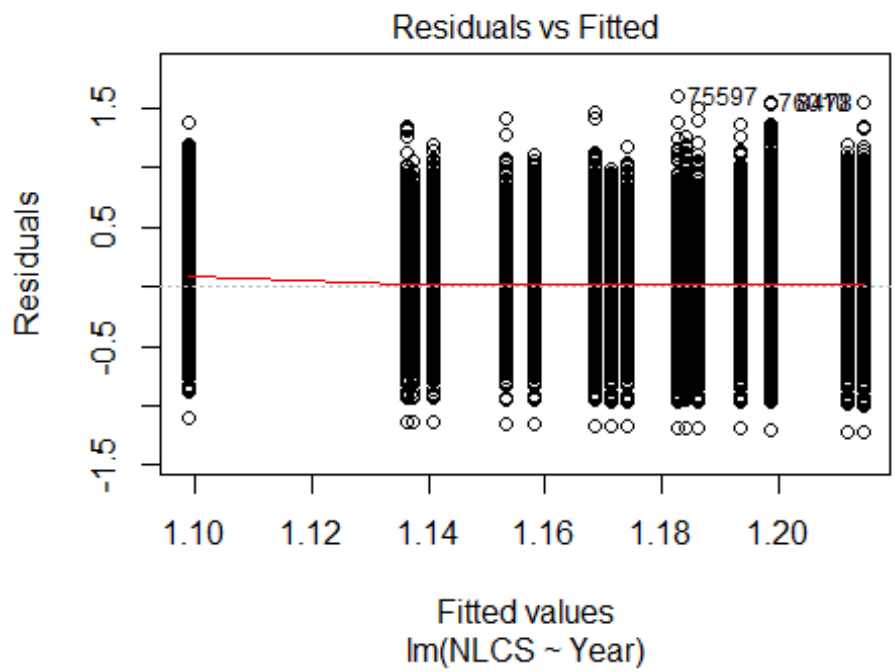
```
## Year2007           -0.09254    0.04145   -2.23    0.026 *
## Year2008            0.01176    0.04227    0.28    0.781
## Year2009            0.05203    0.04323    1.20    0.229
## Year2010            0.04917    0.04317    1.14    0.255
## Year2011            0.05708    0.04449    1.28    0.199
## Year2012            0.05332    0.04345    1.23    0.220
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.448
## Multiple R-squared:  0.0136, Adjusted R-squared:  0.0116
## Convergence in 17 IRWLS iterations
##
## Robustness weights:
##  2 observations c(393,420) are outliers with |weight| = 0 ( < 1.2e-05);
##  751 weights are ~= 1. The remaining 7686 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0449  0.8530  0.9500  0.8860  0.9860  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.18e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##                   GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.018  1           1.009
## Year             1.018 16           1.001
```



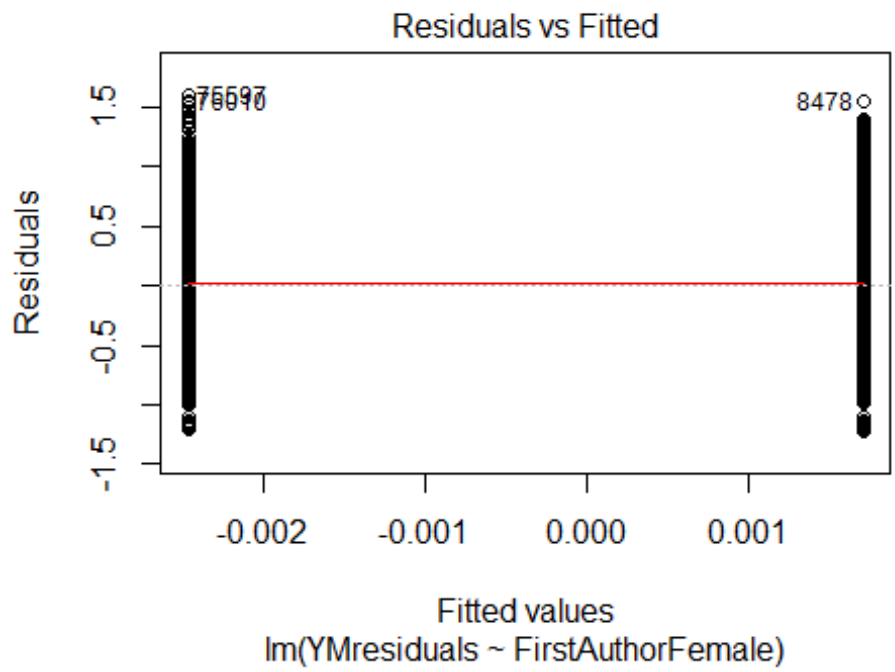

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.2869 -0.3034  0.0152  0.2993  2.4476
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.22739    0.03765   32.60   <2e-16 ***
## LastAuthorFemale1 -0.01690    0.01344   -1.26    0.209
## Year1997          -0.04287    0.05118   -0.84    0.402
## Year1998           0.03536    0.04389    0.81    0.420
## Year1999          -0.03404    0.04378   -0.78    0.437
## Year2000           0.00369    0.04551    0.08    0.935
## Year2001          -0.06175    0.04736   -1.30    0.192
## Year2002          -0.05320    0.04351   -1.22    0.222
## Year2003          -0.01061    0.04282   -0.25    0.804
## Year2004          -0.06670    0.04221   -1.58    0.114
## Year2005          -0.09315    0.04215   -2.21    0.027 *
## Year2006          -0.09598    0.04157   -2.31    0.021 *
```

```
## Year2007          -0.08892    0.04154   -2.14    0.032 *
## Year2008           0.01496    0.04237    0.35    0.724
## Year2009           0.05442    0.04328    1.26    0.209
## Year2010           0.05188    0.04326    1.20    0.231
## Year2011           0.05953    0.04454    1.34    0.181
## Year2012           0.05737    0.04354    1.32    0.188
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.447
## Multiple R-squared:  0.0137, Adjusted R-squared:  0.0117
## Convergence in 18 IRWLS iterations
##
## Robustness weights:
##  2 observations c(393,420) are outliers with |weight| = 0 ( < 1.2e-05);
##  739 weights are ~= 1. The remaining 7698 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0473  0.8530  0.9500  0.8870  0.9860  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.18e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Sample size for the above analysis:  8439"
## [1] ""
## [1] ""
## [1] "####################################"
## [1] "Analysis of AJSC 2403"
## [1] "####################################"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 4923 4970 5020 4915 5767 5378 5100 4228 4535 4764 4756 5103 5051 4699 4412
## 2011 2012
## 4750 4728
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 3128 3050 3116 3372 2783 1465 3380 2916 3097 3310 3188 3448 3339 3211 3068
```

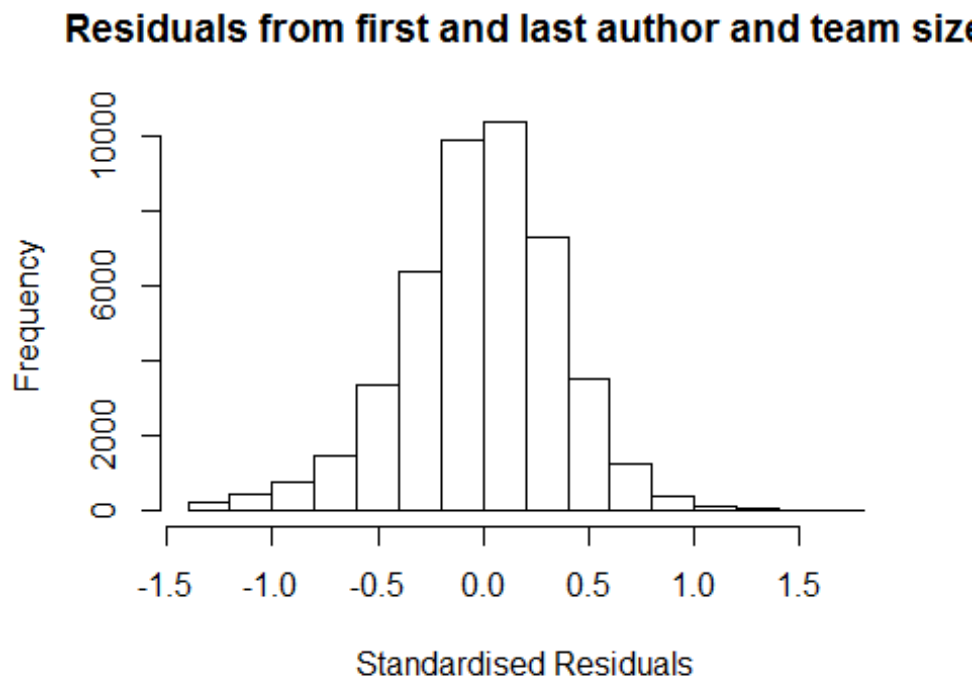
```
## 2011 2012
## 3175 3207
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 2755 2677 2726 2923 2450 1250 2976 2552 2706 2878 2755 2979 2869 2746 2637
## 2011 2012
## 2746 2759
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
##  Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 1100, df = 16, p-value <2e-16
```




```
##
##  Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 120, df = 1, p-value <2e-16
```

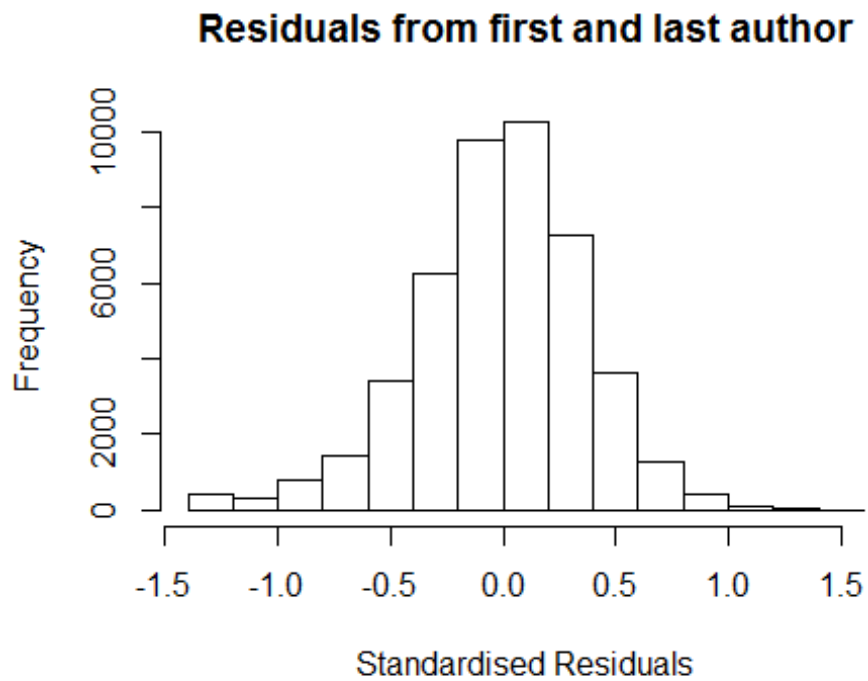



```
## [1] "Female first author team size 2018 geometric mean: 5.91117040641259"
## [1] "Male first author team size 2018 geometric mean: 5.65427333967026"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 670000, p-value = 0.4
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 5.93349634708842"
## [1] "Male last author team size 2018 geometric mean: 5.71316342469773"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 590000, p-value = 0.3
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##                     GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.022  1           1.011
## LastAuthorFemale  1.013  1           1.007
## UniqueAuthors     1.045  4           1.006
## Year              1.054 16           1.002
```



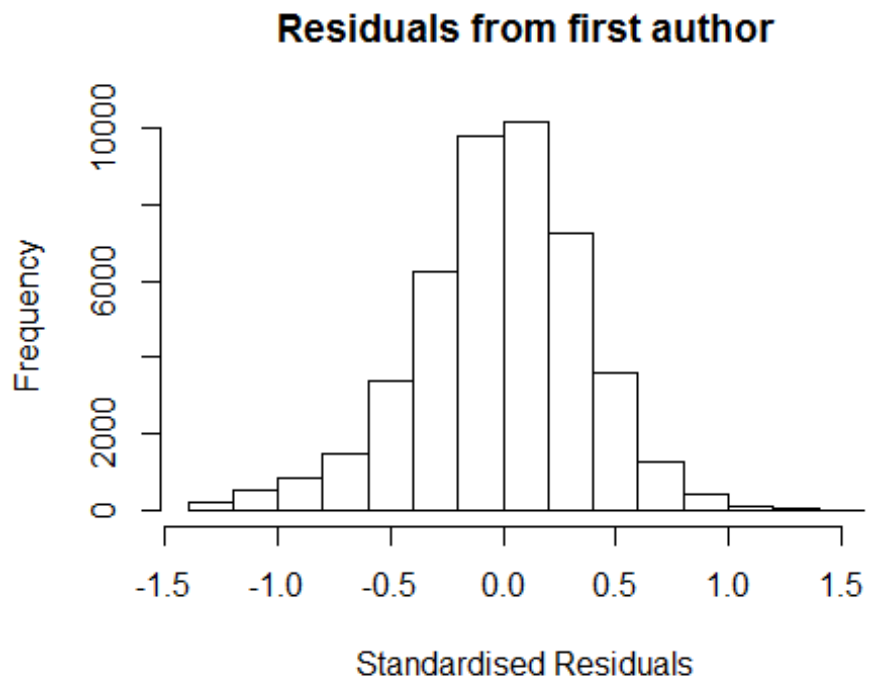

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
UniqueAuthors +
##     Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##     k.max = 1000))
##  \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.32290 -0.22980  0.00426  0.22714  1.78874
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         0.95271    0.01565   60.88  < 2e-16 ***
## FirstAuthorFemale1 -0.00762    0.00344   -2.21  0.02687 *
## LastAuthorFemale1  -0.03430    0.00404   -8.48  < 2e-16 ***
## UniqueAuthors2      0.20038    0.01468   13.65  < 2e-16 ***
## UniqueAuthors3      0.22294    0.01449   15.39  < 2e-16 ***
## UniqueAuthors4      0.25746    0.01449   17.77  < 2e-16 ***
## UniqueAuthors5      0.34564    0.01409   24.53  < 2e-16 ***
## Year1997            0.02456    0.01094    2.24  0.02480 *
## Year1998           -0.01442    0.01036   -1.39  0.16396
## Year1999           -0.02043    0.01001   -2.04  0.04134 *
```

```
## Year2000           -0.01105    0.01228   -0.90  0.36811
## Year2001           -0.02420    0.01373   -1.76  0.07795 .
## Year2002           -0.04976    0.00985   -5.05  4.4e-07 ***
## Year2003           -0.05364    0.01001   -5.36  8.4e-08 ***
## Year2004           -0.07349    0.00991   -7.41  1.3e-13 ***
## Year2005           -0.07719    0.00989   -7.80  6.2e-15 ***
## Year2006           -0.07483    0.01016   -7.37  1.8e-13 ***
## Year2007           -0.05118    0.01008   -5.08  3.8e-07 ***
## Year2008           -0.04218    0.01015   -4.16  3.2e-05 ***
## Year2009           -0.02911    0.01037   -2.81  0.00500 **
## Year2010           -0.01457    0.01068   -1.36  0.17247
## Year2011           -0.04045    0.01052   -3.84  0.00012 ***
## Year2012           -0.02150    0.01109   -1.94  0.05244 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.34
## Multiple R-squared:  0.0563, Adjusted R-squared:  0.0559
## Convergence in 15 IRWLS iterations
##
## Robustness weights:
##  2 observations c(2748,4054) are outliers with |weight| = 0 ( < 2.2e-06);
##  3898 weights are ~= 1. The remaining 41484 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0046  0.8640  0.9500  0.8930  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          2.20e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.016  1           1.008
## LastAuthorFemale  1.010  1           1.005
## Year              1.012 16           1.000
```



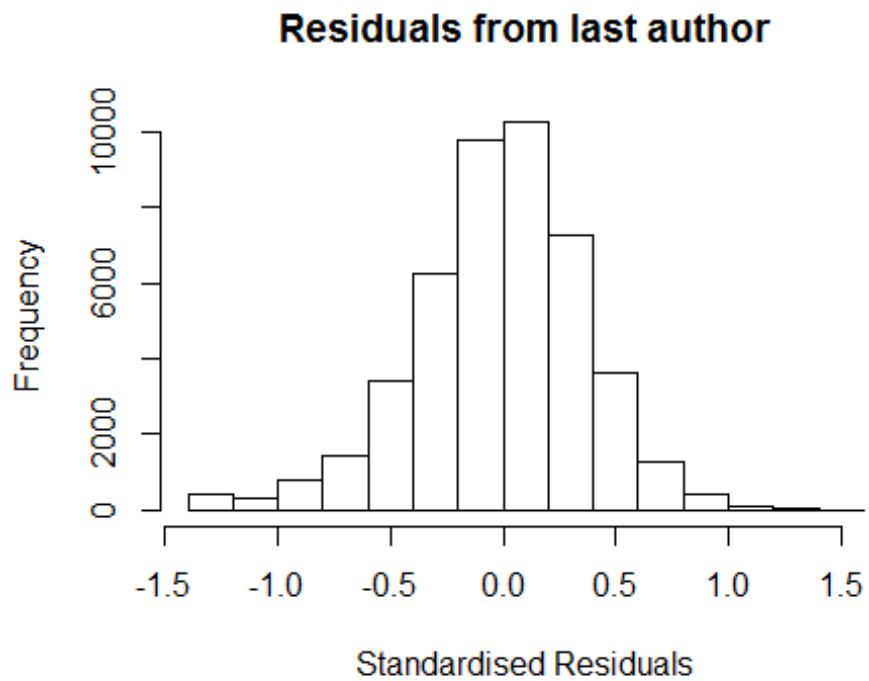

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.24513 -0.23412  0.00444  0.22941  1.57963
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.21666    0.00810  150.22  < 2e-16 ***
## FirstAuthorFemale1 -0.00434    0.00350   -1.24  0.21456
## LastAuthorFemale1  -0.03989    0.00412   -9.68  < 2e-16 ***
## Year1997            0.02847    0.01112    2.56  0.01045 *
## Year1998           -0.01200    0.01058   -1.13  0.25639
## Year1999           -0.01550    0.01013   -1.53  0.12621
## Year2000           -0.01050    0.01257   -0.83  0.40383
## Year2001           -0.00366    0.01396   -0.26  0.79306
## Year2002           -0.03284    0.01002   -3.28  0.00105 **
## Year2003           -0.03613    0.01022   -3.54  0.00041 ***
## Year2004           -0.05501    0.01014   -5.43  5.7e-08 ***
## Year2005           -0.05607    0.01011   -5.55  2.9e-08 ***
```

```
## Year2006           -0.05336    0.01035   -5.15  2.6e-07 ***
## Year2007           -0.02500    0.01028   -2.43  0.01501 *
## Year2008           -0.00995    0.01033   -0.96  0.33569
## Year2009            0.00338    0.01062    0.32  0.75010
## Year2010            0.01867    0.01085    1.72  0.08525 .
## Year2011           -0.00575    0.01073   -0.54  0.59228
## Year2012            0.01163    0.01127    1.03  0.30228
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.345
## Multiple R-squared:  0.00705,    Adjusted R-squared:  0.00666
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
##  3882 weights are ~= 1. The remaining 41502 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0022  0.8630  0.9500  0.8920  0.9860  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          2.20e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.009  1           1.005
## Year              1.009 16           1.000
```




```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.23790 -0.23440  0.00413  0.22939  1.59119
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.20938    0.00808  149.77  < 2e-16 ***
## FirstAuthorFemale1 -0.00739    0.00349   -2.12  0.03420 *
## Year1997            0.02852    0.01115    2.56  0.01050 *
## Year1998           -0.01218    0.01059   -1.15  0.25021
## Year1999           -0.01568    0.01015   -1.54  0.12248
## Year2000           -0.01102    0.01259   -0.88  0.38156
## Year2001           -0.00428    0.01396   -0.31  0.75898
## Year2002           -0.03282    0.01005   -3.27  0.00109 **
## Year2003           -0.03635    0.01024   -3.55  0.00039 ***
## Year2004           -0.05539    0.01016   -5.45  5.0e-08 ***
## Year2005           -0.05655    0.01013   -5.58  2.4e-08 ***
## Year2006           -0.05391    0.01037   -5.20  2.0e-07 ***
```

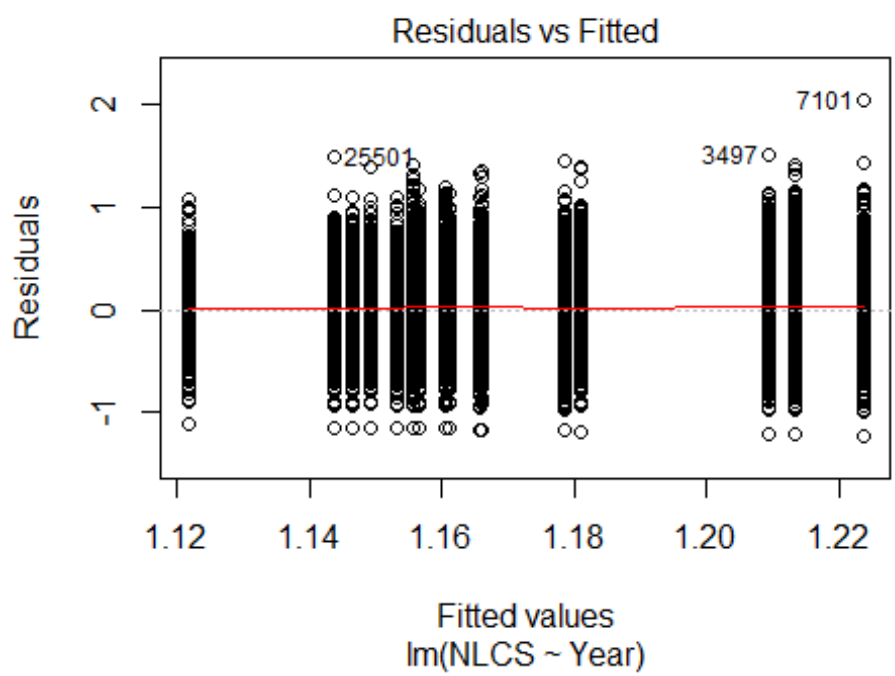
```
## Year2007           -0.02638    0.01030   -2.56  0.01043 *
## Year2008           -0.01118    0.01036   -1.08  0.28016
## Year2009            0.00249    0.01064    0.23  0.81483
## Year2010            0.01707    0.01086    1.57  0.11617
## Year2011           -0.00750    0.01075   -0.70  0.48523
## Year2012            0.00992    0.01129    0.88  0.37962
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.346
## Multiple R-squared:  0.00484,    Adjusted R-squared:  0.00447
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
##  3899 weights are ~= 1. The remaining 41485 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0012  0.8630  0.9500  0.8920  0.9860  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          2.20e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##                   GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.004  1           1.002
## Year             1.004 16           1.000
```



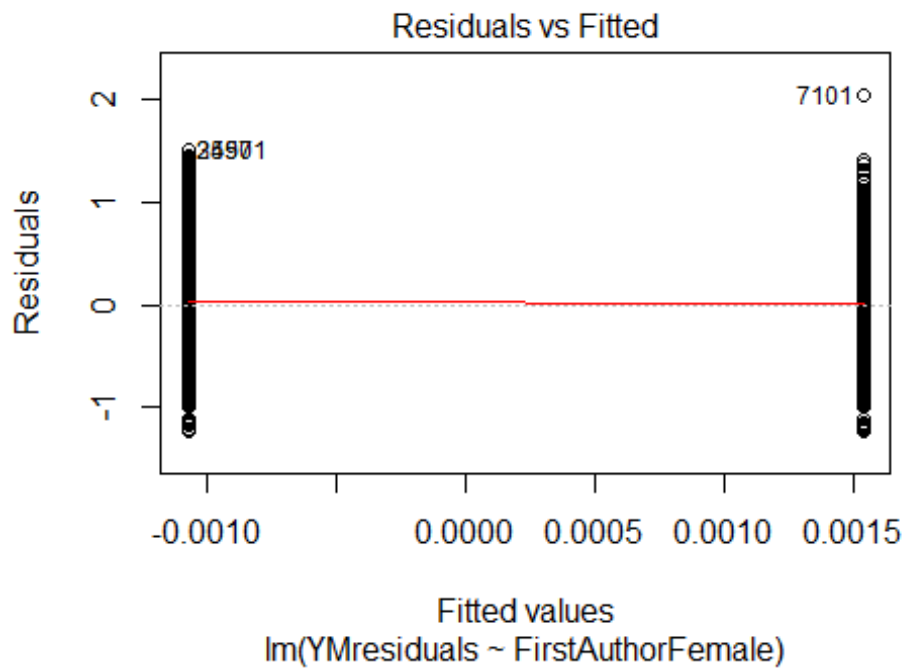

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.24365 -0.23415  0.00462  0.22948  1.57707
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.21522    0.00801  151.80  < 2e-16 ***
## LastAuthorFemale1 -0.04034    0.00411   -9.83  < 2e-16 ***
## Year1997            0.02843    0.01112    2.56  0.01057 *
## Year1998           -0.01208    0.01057   -1.14  0.25321
## Year1999           -0.01563    0.01013   -1.54  0.12307
## Year2000           -0.01075    0.01257   -0.86  0.39247
## Year2001           -0.00369    0.01396   -0.26  0.79155
## Year2002           -0.03306    0.01002   -3.30  0.00097 ***
## Year2003           -0.03638    0.01022   -3.56  0.00037 ***
## Year2004           -0.05533    0.01013   -5.46  4.8e-08 ***
## Year2005           -0.05637    0.01011   -5.58  2.4e-08 ***
## Year2006           -0.05373    0.01035   -5.19  2.1e-07 ***
```

```
## Year2007          -0.02544    0.01027   -2.48  0.01329 *
## Year2008          -0.01029    0.01033   -1.00  0.31910
## Year2009           0.00290    0.01061    0.27  0.78433
## Year2010           0.01817    0.01084    1.68  0.09373 .
## Year2011          -0.00627    0.01072   -0.58  0.55898
## Year2012           0.01111    0.01126    0.99  0.32388
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.345
## Multiple R-squared:  0.00701,    Adjusted R-squared:  0.00664
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
##  3906 weights are ~= 1. The remaining 41478 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0026  0.8630  0.9500  0.8920  0.9860  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          2.20e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Sample size for the above analysis:  45384"
## [1] ""
## [1] ""
## [1] "####################################"
## [1] "Analysis of AJSC 2404"
## [1] "####################################"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 2426 2425 2286 2202 2389 2368 2327 2165 2462 2633 2644 2733 2568 2549 2395
## 2011 2012
## 2643 2632
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 1497 1444 1312 1433 1100  555 1468 1427 1618 1742 1751 1849 1732 1772 1720
## 2011 2012
```

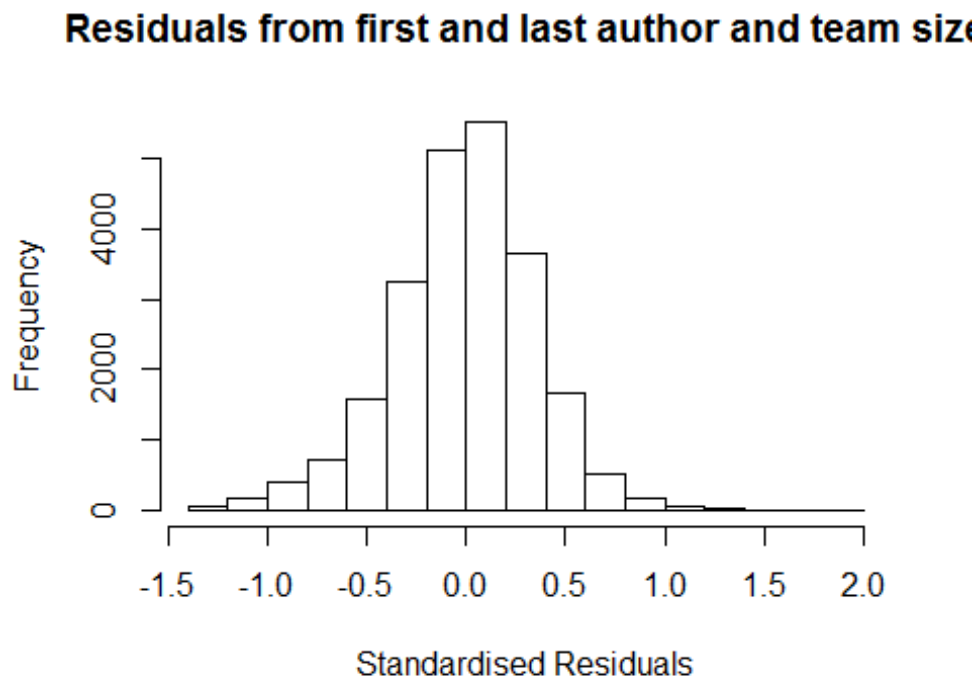
```
## 1835 1891
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 1313 1266 1147 1252  960  492 1296 1256 1437 1530 1539 1617 1517 1563 1470
## 2011 2012
## 1586 1650
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
##  Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 350, df = 16, p-value <2e-16
```




```
##
##  Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 71, df = 1, p-value <2e-16
```

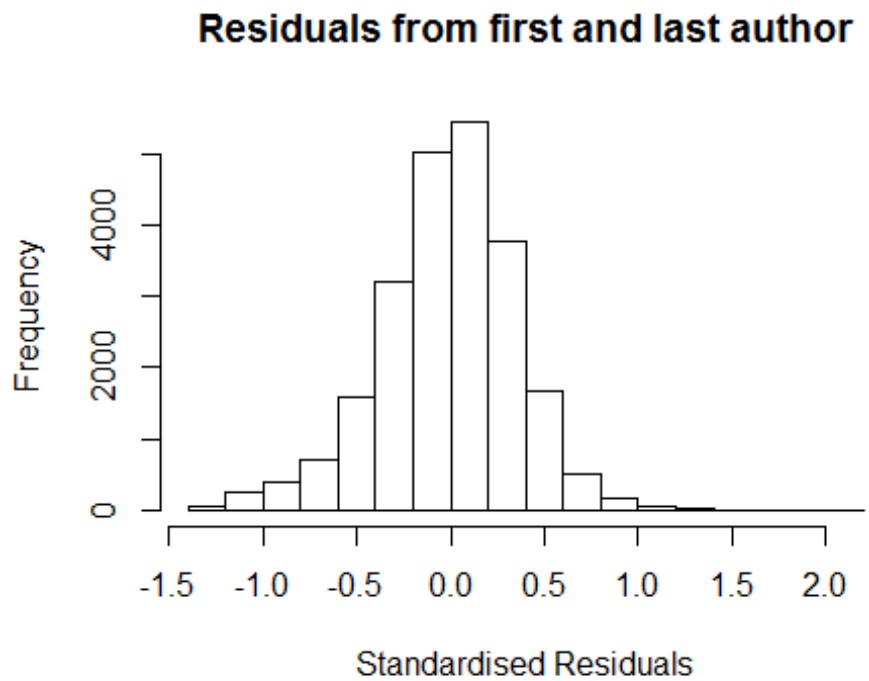



```
## [1] "Female first author team size 2018 geometric mean: 5.10008347885105"
## [1] "Male first author team size 2018 geometric mean: 4.75107950688753"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 390000, p-value = 0.03
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 4.80441855647366"
## [1] "Male last author team size 2018 geometric mean: 4.98545006088521"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 320000, p-value = 0.4
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.020  1           1.010
## LastAuthorFemale  1.014  1           1.007
## UniqueAuthors     1.048  4           1.006
## Year              1.059 16           1.002
```



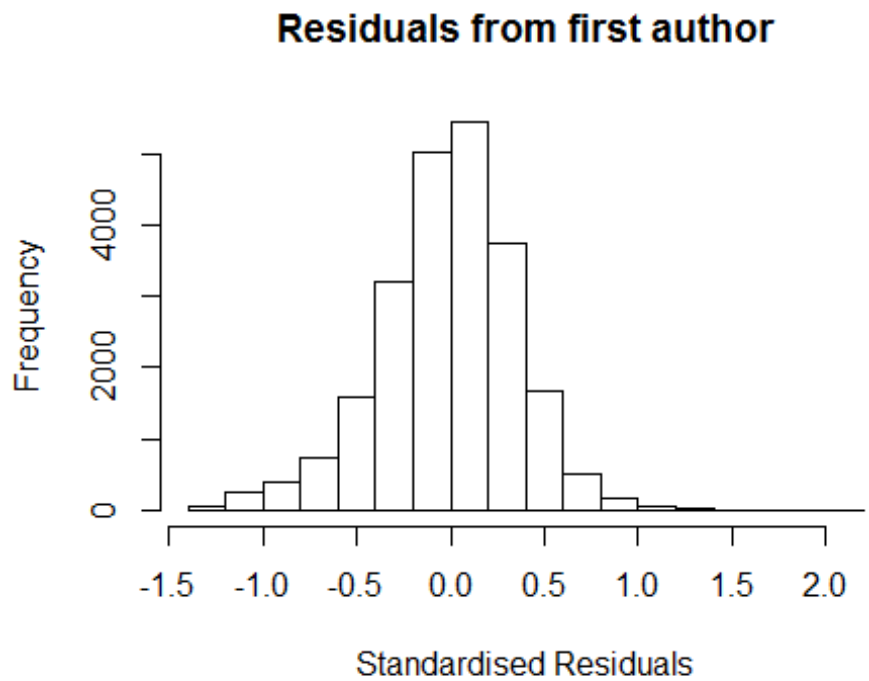

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
UniqueAuthors +
##     Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##     k.max = 1000))
##  \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.33924 -0.21956  0.00576  0.21624  1.93822
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         0.99187    0.02063   48.09  < 2e-16 ***
## FirstAuthorFemale1 -0.00596    0.00461   -1.29   0.1962
## LastAuthorFemale1   0.00346    0.00543    0.64   0.5231
## UniqueAuthors2      0.20199    0.01827   11.05  < 2e-16 ***
## UniqueAuthors3      0.23318    0.01814   12.85  < 2e-16 ***
## UniqueAuthors4      0.26746    0.01828   14.63  < 2e-16 ***
## UniqueAuthors5      0.34532    0.01784   19.35  < 2e-16 ***
## Year1997            0.00205    0.01562    0.13   0.8955
## Year1998           -0.00444    0.01530   -0.29   0.7714
## Year1999           -0.04540    0.01487   -3.05   0.0023 **
```

```
## Year2000           -0.06397    0.01597   -4.01  6.2e-05 ***
## Year2001           -0.09138    0.01965   -4.65  3.3e-06 ***
## Year2002           -0.08410    0.01424   -5.91  3.5e-09 ***
## Year2003           -0.07263    0.01428   -5.09  3.7e-07 ***
## Year2004           -0.10340    0.01398   -7.40  1.4e-13 ***
## Year2005           -0.10312    0.01374   -7.51  6.3e-14 ***
## Year2006           -0.09926    0.01393   -7.12  1.1e-12 ***
## Year2007           -0.09413    0.01386   -6.79  1.1e-11 ***
## Year2008           -0.09494    0.01400   -6.78  1.2e-11 ***
## Year2009           -0.09188    0.01403   -6.55  5.9e-11 ***
## Year2010           -0.06995    0.01460   -4.79  1.7e-06 ***
## Year2011           -0.08242    0.01446   -5.70  1.2e-08 ***
## Year2012           -0.09302    0.01472   -6.32  2.7e-10 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.324
## Multiple R-squared:  0.0607, Adjusted R-squared:  0.0598
## Convergence in 16 IRWLS iterations
##
## Robustness weights:
##  5 observations c(1097,1687,3391,11591,22658)
##   are outliers with |weight| = 0 ( < 4.4e-06);
##  1995 weights are ~= 1. The remaining 20891 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.005   0.864   0.950   0.892   0.985   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          4.37e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.014  1           1.007
## LastAuthorFemale  1.013  1           1.007
## Year              1.016 16           1.001
```



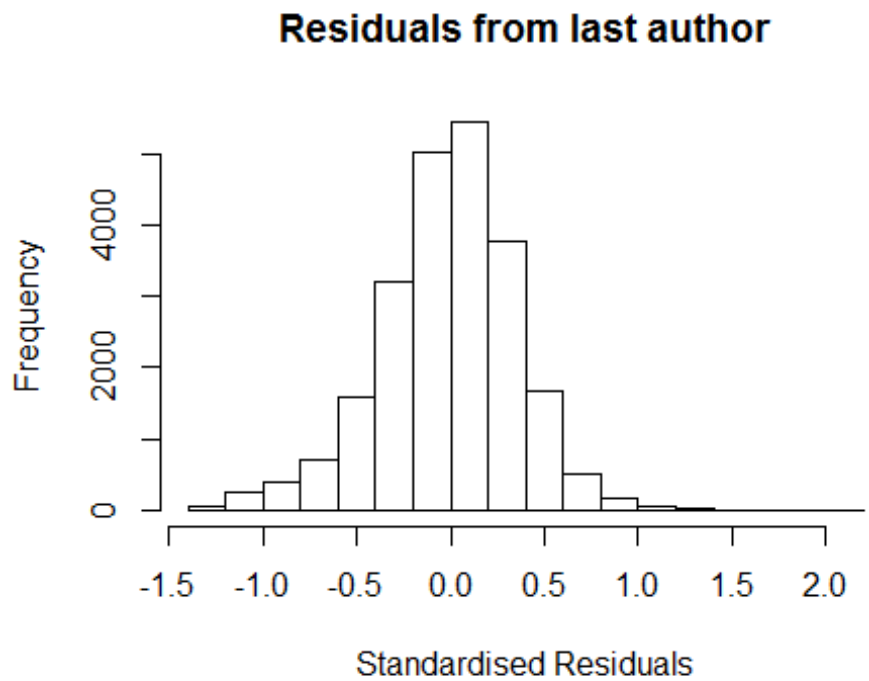

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.24109 -0.22456  0.00723  0.22000  2.02649
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.24109    0.01171  106.00  < 2e-16 ***
## FirstAuthorFemale1 -0.00091    0.00469   -0.19  0.84628
## LastAuthorFemale1  -0.00182    0.00548   -0.33  0.73938
## Year1997           -0.00213    0.01596   -0.13  0.89366
## Year1998           -0.00167    0.01565   -0.11  0.91521
## Year1999           -0.04253    0.01514   -2.81  0.00498 **
## Year2000           -0.05855    0.01635   -3.58  0.00034 ***
## Year2001           -0.08910    0.02009   -4.44  9.2e-06 ***
## Year2002           -0.06732    0.01450   -4.64  3.5e-06 ***
## Year2003           -0.05811    0.01461   -3.98  7.0e-05 ***
## Year2004           -0.08715    0.01437   -6.06  1.4e-09 ***
## Year2005           -0.08424    0.01405   -6.00  2.1e-09 ***
```

```
## Year2006           -0.07949    0.01427   -5.57  2.6e-08 ***
## Year2007           -0.06829    0.01419   -4.81  1.5e-06 ***
## Year2008           -0.06606    0.01435   -4.60  4.2e-06 ***
## Year2009           -0.06484    0.01436   -4.51  6.4e-06 ***
## Year2010           -0.04188    0.01502   -2.79  0.00529 **
## Year2011           -0.05103    0.01479   -3.45  0.00056 ***
## Year2012           -0.06035    0.01498   -4.03  5.6e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.331
## Multiple R-squared:  0.00611,    Adjusted R-squared:  0.00533
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
##  observation 3391 is an outlier with |weight| = 0 ( < 4.4e-06);
##  1935 weights are ~= 1. The remaining 20955 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0056  0.8650  0.9510  0.8910  0.9860  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          4.37e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.009  1           1.005
## Year              1.009 16           1.000
```




```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.24074 -0.22446  0.00683  0.21990  2.02694
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.24074    0.01164  106.63  < 2e-16 ***
## FirstAuthorFemale1 -0.00104    0.00468   -0.22  0.82419
## Year1997           -0.00215    0.01596   -0.13  0.89308
## Year1998           -0.00164    0.01565   -0.10  0.91646
## Year1999           -0.04249    0.01514   -2.81  0.00501 **
## Year2000           -0.05854    0.01635   -3.58  0.00034 ***
## Year2001           -0.08910    0.02009   -4.44  9.2e-06 ***
## Year2002           -0.06729    0.01450   -4.64  3.5e-06 ***
## Year2003           -0.05811    0.01462   -3.98  7.0e-05 ***
## Year2004           -0.08713    0.01437   -6.06  1.4e-09 ***
## Year2005           -0.08423    0.01405   -5.99  2.1e-09 ***
## Year2006           -0.07950    0.01428   -5.57  2.6e-08 ***
```

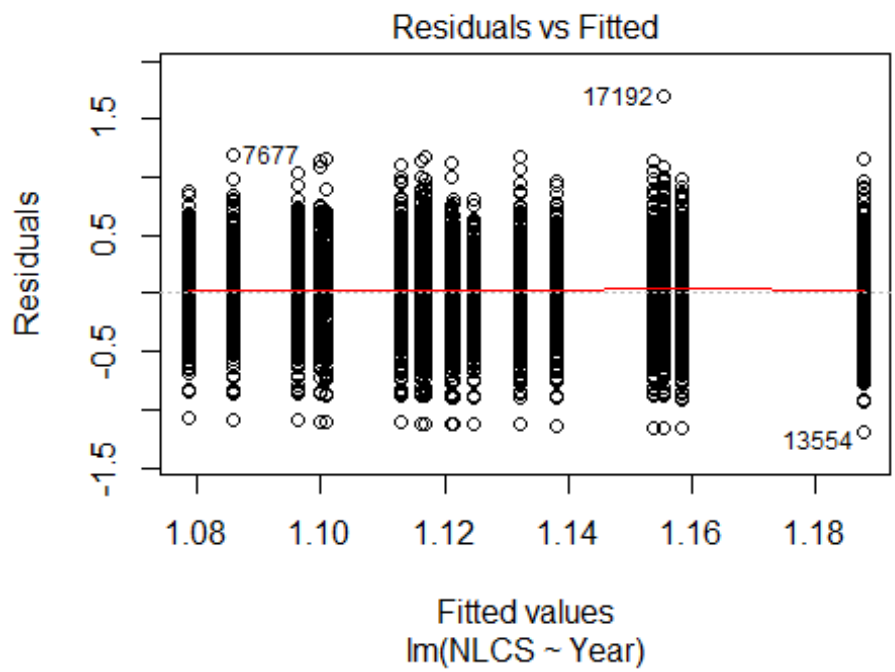
```
## Year2007           -0.06835    0.01419   -4.81  1.5e-06 ***
## Year2008           -0.06611    0.01435   -4.61  4.1e-06 ***
## Year2009           -0.06484    0.01437   -4.51  6.4e-06 ***
## Year2010           -0.04196    0.01502   -2.79  0.00521 **
## Year2011           -0.05111    0.01479   -3.46  0.00055 ***
## Year2012           -0.06046    0.01498   -4.04  5.5e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.331
## Multiple R-squared:  0.0061, Adjusted R-squared:  0.00537
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
##  observation 3391 is an outlier with |weight| = 0 ( < 4.4e-06);
##  1951 weights are ~= 1. The remaining 20939 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0059  0.8650  0.9510  0.8910  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          4.37e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##                   GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.008  1           1.004
## Year             1.008 16           1.000
```



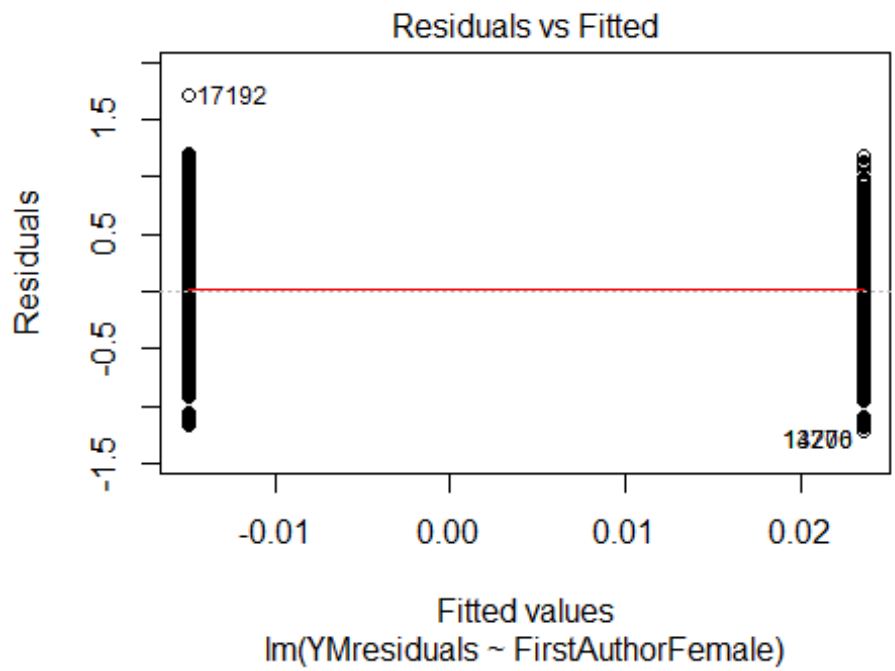

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.24078 -0.22443  0.00736  0.22036  2.02590
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.24078    0.01158  107.13  < 2e-16 ***
## LastAuthorFemale1 -0.00191    0.00546   -0.35  0.72647
## Year1997          -0.00217    0.01596   -0.14  0.89200
## Year1998          -0.00168    0.01565   -0.11  0.91456
## Year1999          -0.04254    0.01514   -2.81  0.00497 **
## Year2000          -0.05857    0.01635   -3.58  0.00034 ***
## Year2001          -0.08912    0.02009   -4.44  9.2e-06 ***
## Year2002          -0.06733    0.01450   -4.64  3.5e-06 ***
## Year2003          -0.05815    0.01461   -3.98  6.9e-05 ***
## Year2004          -0.08722    0.01437   -6.07  1.3e-09 ***
## Year2005          -0.08432    0.01404   -6.00  2.0e-09 ***
## Year2006          -0.07959    0.01427   -5.58  2.4e-08 ***
```

```
## Year2007          -0.06836    0.01419   -4.82  1.5e-06 ***
## Year2008          -0.06613    0.01435   -4.61  4.1e-06 ***
## Year2009          -0.06492    0.01436   -4.52  6.2e-06 ***
## Year2010          -0.04198    0.01500   -2.80  0.00514 **
## Year2011          -0.05112    0.01478   -3.46  0.00054 ***
## Year2012          -0.06048    0.01496   -4.04  5.3e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.331
## Multiple R-squared:  0.00611,    Adjusted R-squared:  0.00537
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
##  observation 3391 is an outlier with |weight| = 0 ( < 4.4e-06);
##  1941 weights are ~= 1. The remaining 20949 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0055  0.8650  0.9510  0.8910  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          4.37e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Sample size for the above analysis:  22891"
## [1] ""
## [1] ""
## [1] "####################################"
## [1] "Analysis of AJSC 2405"
## [1] "####################################"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  844  774  802  766  828  855  696  651  647  809  766  782  853  953  902
## 2011 2012
##  939  933
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  543  429  473  489  397  171  436  430  451  563  532  563  586  684  660
```

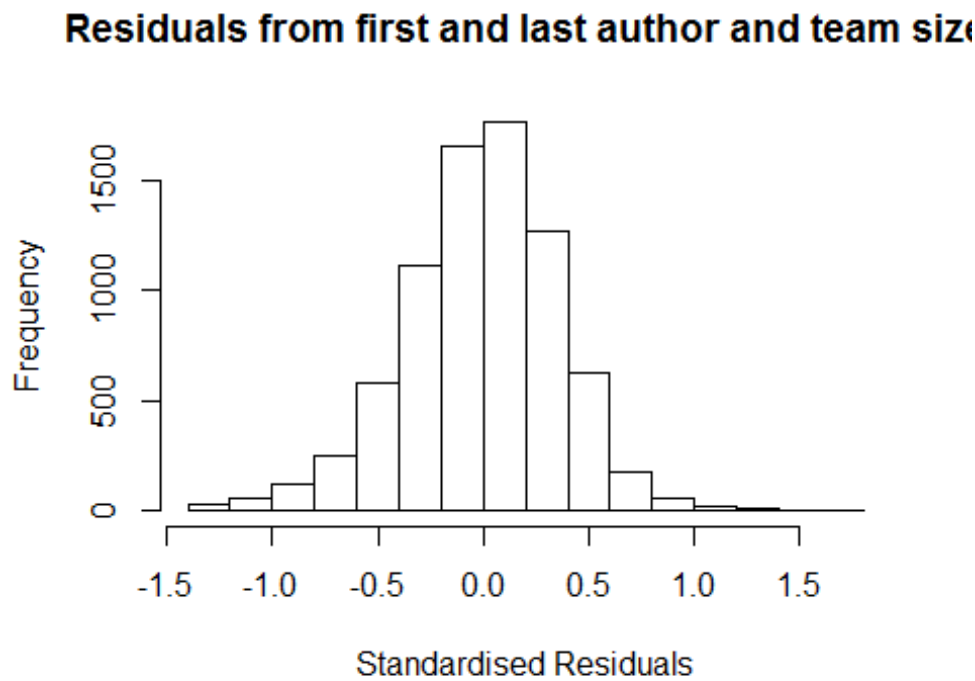
```
## 2011 2012
##  673  668
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  484  383  408  429  354  145  378  370  399  492  469  507  511  596  586
## 2011 2012
##  602  588
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
##  Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 140, df = 16, p-value <2e-16
```




```
##
##  Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 42, df = 1, p-value = 8e-11
```

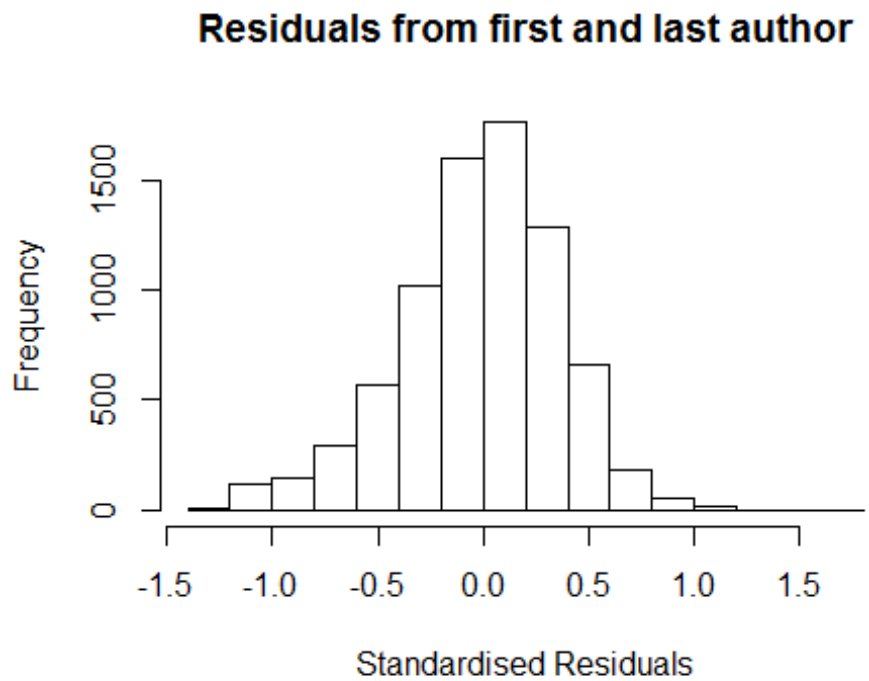



```
## [1] "Female first author team size 2018 geometric mean: 5.11387391984283"
## [1] "Male first author team size 2018 geometric mean: 4.71091895858354"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 38000, p-value = 0.4
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 5.05268808132344"
## [1] "Male last author team size 2018 geometric mean: 4.85078210518087"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 33000, p-value = 0.8
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##                     GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.040  1           1.020
## LastAuthorFemale  1.017  1           1.008
## UniqueAuthors     1.088  4           1.011
## Year              1.116 16           1.003
```



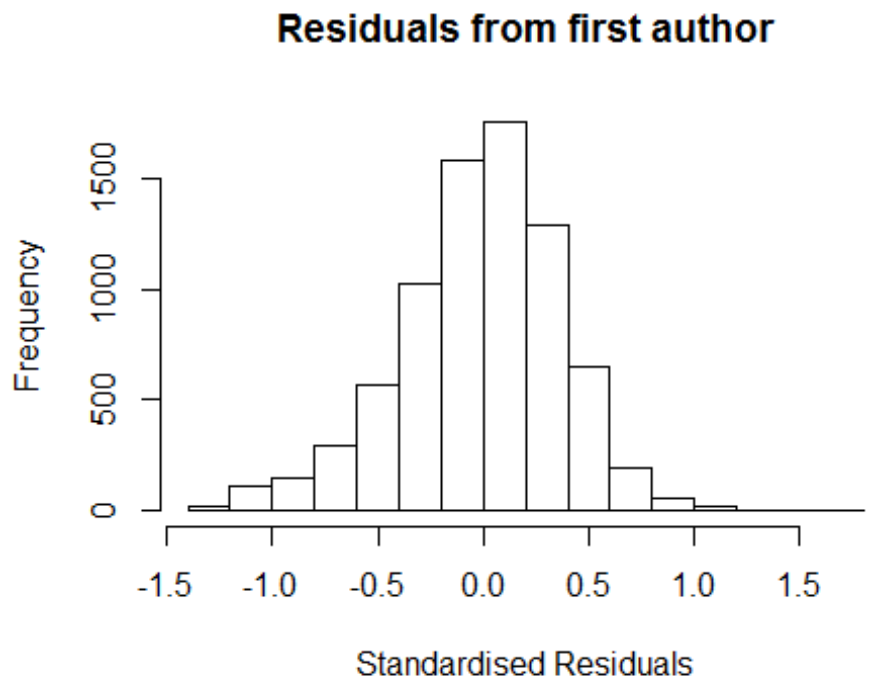

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
UniqueAuthors +
##     Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##     k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.2961 -0.2294  0.0064  0.2303  1.7968
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)         0.850829   0.031659   26.87  < 2e-16 ***
## FirstAuthorFemale1  0.025152   0.008308    3.03  0.00247 **
## LastAuthorFemale1  -0.017895   0.009818   -1.82  0.06839 .
## UniqueAuthors2      0.203602   0.029866    6.82    1e-11 ***
## UniqueAuthors3      0.260142   0.029394    8.85  < 2e-16 ***
## UniqueAuthors4      0.333232   0.029354   11.35  < 2e-16 ***
## UniqueAuthors5      0.419331   0.028405   14.76  < 2e-16 ***
## Year1997            0.014303   0.026050    0.55  0.58298
## Year1998           -0.006410   0.024810   -0.26  0.79613
## Year1999           -0.048420   0.024024   -2.02  0.04389 *
```

```
## Year2000           -0.022981   0.024968   -0.92  0.35737
## Year2001           -0.071334   0.036808   -1.94  0.05266 .
## Year2002           -0.069861   0.023154   -3.02  0.00256 **
## Year2003           -0.040737   0.023659   -1.72  0.08514 .
## Year2004           -0.063811   0.024437   -2.61  0.00904 **
## Year2005           -0.078799   0.023367   -3.37  0.00075 ***
## Year2006           -0.063723   0.023653   -2.69  0.00707 **
## Year2007           -0.060564   0.023491   -2.58  0.00995 **
## Year2008           -0.057013   0.023299   -2.45  0.01443 *
## Year2009           -0.016968   0.023109   -0.73  0.46282
## Year2010            0.019402   0.023184    0.84  0.40268
## Year2011           -0.015945   0.023662   -0.67  0.50040
## Year2012            0.000768   0.024783    0.03  0.97529
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.34
## Multiple R-squared:  0.102,  Adjusted R-squared:  0.0998
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
##  observation 7632 is an outlier with |weight| = 0 ( < 1.3e-05);
##  636 weights are ~= 1. The remaining 7064 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.115   0.869   0.951   0.899   0.986   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.30e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.030  1           1.015
## LastAuthorFemale  1.015  1           1.008
## Year              1.034 16           1.001
```



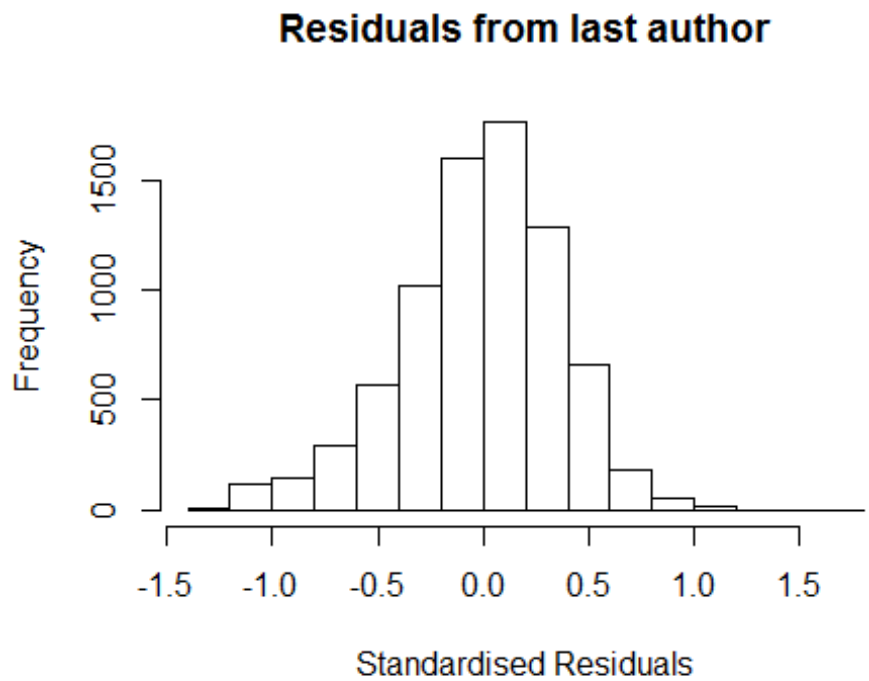

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.2327 -0.2407  0.0118  0.2370  1.6721
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)          1.12865    0.01852   60.93  < 2e-16 ***
## FirstAuthorFemale1   0.03866    0.00858    4.51  6.6e-06 ***
## LastAuthorFemale1   -0.02100    0.01011   -2.08   0.0378 *
## Year1997             0.02949    0.02709    1.09   0.2763
## Year1998             0.02448    0.02586    0.95   0.3438
## Year1999            -0.02600    0.02504   -1.04   0.2992
## Year2000            -0.00923    0.02597   -0.36   0.7224
## Year2001            -0.04698    0.03809   -1.23   0.2174
## Year2002            -0.01553    0.02408   -0.64   0.5190
## Year2003            -0.00809    0.02459   -0.33   0.7421
## Year2004            -0.03754    0.02518   -1.49   0.1359
## Year2005            -0.03855    0.02434   -1.58   0.1132
```

```
## Year2006           -0.02515    0.02500   -1.01   0.3144
## Year2007           -0.01998    0.02476   -0.81   0.4198
## Year2008           -0.00469    0.02473   -0.19   0.8496
## Year2009            0.03057    0.02391    1.28   0.2011
## Year2010            0.06538    0.02406    2.72   0.0066 **
## Year2011            0.04056    0.02480    1.64   0.1020
## Year2012            0.05127    0.02561    2.00   0.0453 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.35
## Multiple R-squared:  0.0118, Adjusted R-squared:  0.00946
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  observation 7632 is an outlier with |weight| = 0 ( < 1.3e-05);
##  631 weights are ~= 1. The remaining 7069 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.188   0.867   0.950   0.895   0.986   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.30e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.025  1           1.012
## Year              1.025 16           1.001
```




```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.2267 -0.2390  0.0113  0.2368  1.6771
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)          1.12507    0.01842   61.09  < 2e-16 ***
## FirstAuthorFemale1   0.03740    0.00856    4.37  1.3e-05 ***
## Year1997             0.02966    0.02707    1.10   0.2733
## Year1998             0.02422    0.02584    0.94   0.3486
## Year1999            -0.02605    0.02506   -1.04   0.2986
## Year2000            -0.00980    0.02596   -0.38   0.7057
## Year2001            -0.04591    0.03806   -1.21   0.2277
## Year2002            -0.01472    0.02409   -0.61   0.5413
## Year2003            -0.00827    0.02460   -0.34   0.7369
## Year2004            -0.03780    0.02520   -1.50   0.1336
## Year2005            -0.03888    0.02434   -1.60   0.1103
## Year2006            -0.02538    0.02502   -1.01   0.3104
```

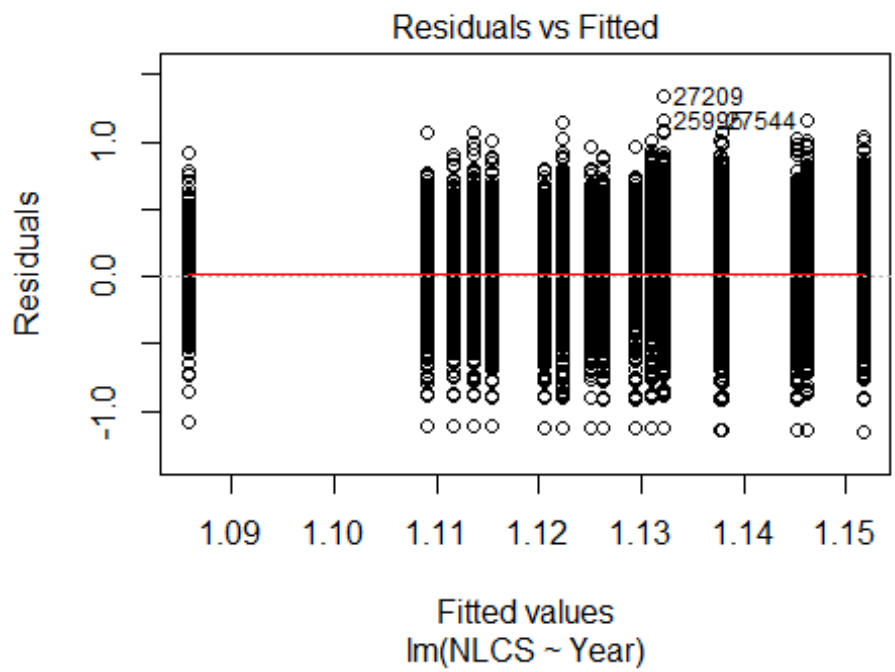
```
## Year2007           -0.02026    0.02476   -0.82   0.4132
## Year2008           -0.00470    0.02474   -0.19   0.8494
## Year2009            0.03035    0.02394    1.27   0.2048
## Year2010            0.06427    0.02407    2.67   0.0076 **
## Year2011            0.03957    0.02481    1.60   0.1107
## Year2012            0.04983    0.02562    1.95   0.0518 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.35
## Multiple R-squared:  0.0112, Adjusted R-squared:  0.00902
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  observation 7632 is an outlier with |weight| = 0 ( < 1.3e-05);
##  635 weights are ~= 1. The remaining 7065 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.193   0.867   0.949   0.895   0.986   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.30e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##                  GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.01  1           1.005
## Year             1.01 16           1.000
```



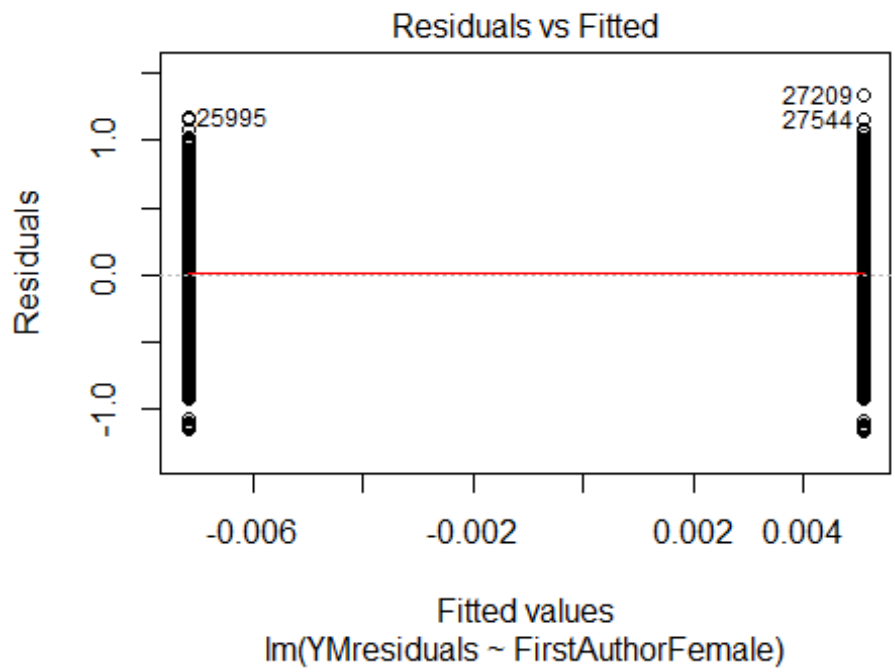

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.2111 -0.2378  0.0136  0.2345  1.6536
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.13839    0.01849   61.58   <2e-16 ***
## LastAuthorFemale1 -0.01761    0.01008   -1.75   0.0807 .
## Year1997            0.03317    0.02712    1.22   0.2214
## Year1998            0.02613    0.02587    1.01   0.3124
## Year1999           -0.02435    0.02512   -0.97   0.3324
## Year2000           -0.00462    0.02595   -0.18   0.8586
## Year2001           -0.04322    0.03788   -1.14   0.2539
## Year2002           -0.01187    0.02406   -0.49   0.6218
## Year2003           -0.00295    0.02455   -0.12   0.9045
## Year2004           -0.03172    0.02516   -1.26   0.2074
## Year2005           -0.03288    0.02428   -1.35   0.1757
## Year2006           -0.01879    0.02505   -0.75   0.4533
```

```
## Year2007          -0.01391    0.02480   -0.56   0.5750
## Year2008           0.00187    0.02475    0.08   0.9398
## Year2009           0.03840    0.02385    1.61   0.1074
## Year2010           0.07271    0.02398    3.03   0.0024 **
## Year2011           0.04831    0.02479    1.95   0.0514 .
## Year2012           0.05999    0.02559    2.34   0.0191 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.35
## Multiple R-squared:  0.00917,    Adjusted R-squared:  0.00698
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  observation 7632 is an outlier with |weight| = 0 ( < 1.3e-05);
##  622 weights are ~= 1. The remaining 7078 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.205   0.866   0.949   0.895   0.986   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.30e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                    psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Sample size for the above analysis:  7701"
## [1] ""
## [1] ""
## [1] "####################################"
## [1] "Analysis of AJSC 2406"
## [1] "####################################"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 1354 1344 1394 1221 1301 1265 1256 1082 1272 1324 1570 1700 1670 1674 1699
## 2011 2012
## 1898 1872
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  912  805  868  801  592  292  829  732  853  912 1088 1155 1132 1121 1218
```

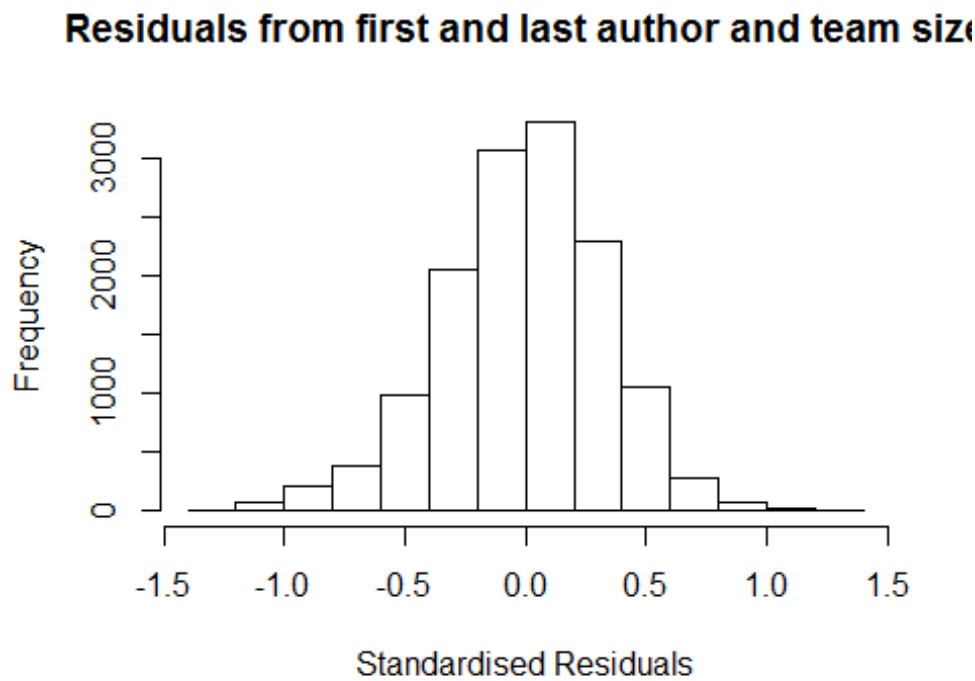
```
## 2011 2012
## 1314 1352
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  803  713  746  682  508  255  723  620  744  783  967  991  968  956 1059
## 2011 2012
## 1128 1166
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
##  Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 140, df = 16, p-value <2e-16
```




```
##
##  Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 3.6, df = 1, p-value = 0.06
```

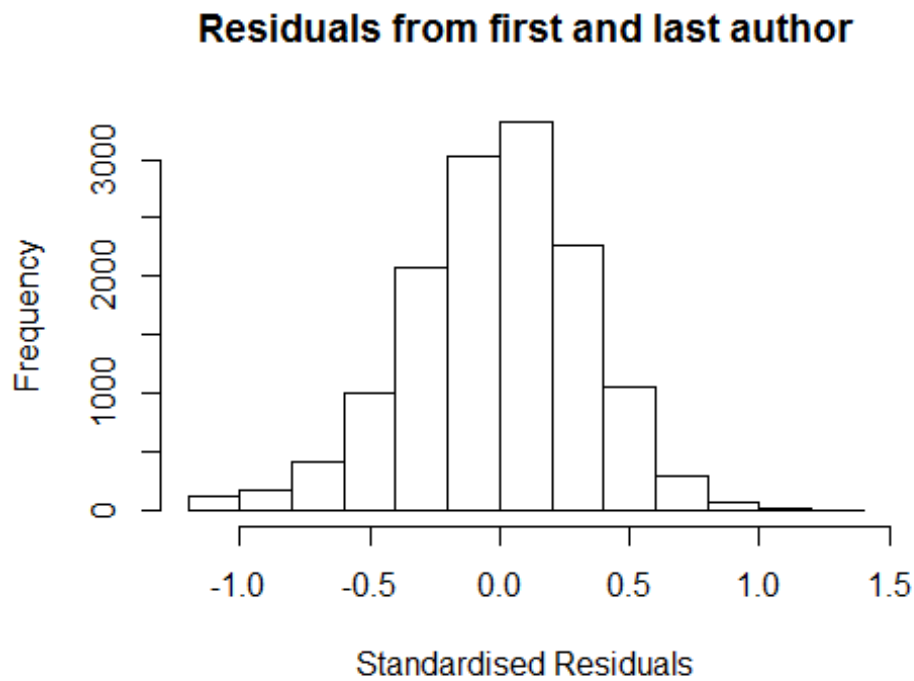



```
## [1] "Female first author team size 2018 geometric mean: 5.41579045631346"
## [1] "Male first author team size 2018 geometric mean: 5.17589709023764"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 110000, p-value = 0.5
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 5.16007638008773"
## [1] "Male last author team size 2018 geometric mean: 5.36143359597983"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 89000, p-value = 0.4
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.018  1           1.009
## LastAuthorFemale  1.012  1           1.006
## UniqueAuthors     1.043  4           1.005
## Year              1.052 16           1.002
```



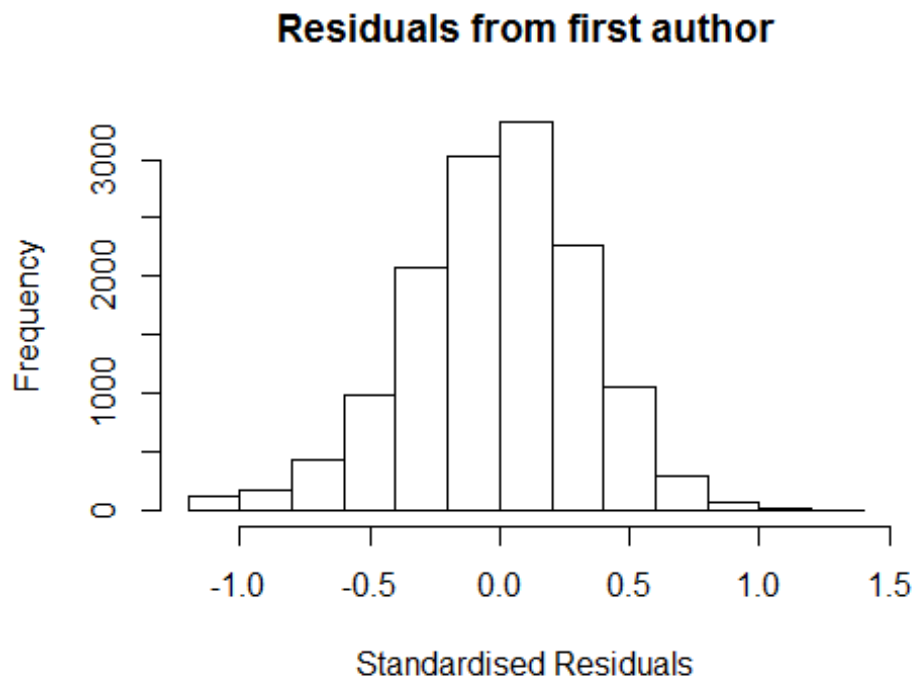

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
UniqueAuthors +
##     Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##     k.max = 1000))
##  \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.22732 -0.22004  0.00745  0.21518  1.25552
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         0.97151    0.02664   36.47  < 2e-16 ***
## FirstAuthorFemale1 -0.00816    0.00582   -1.40    0.161
## LastAuthorFemale1  -0.03867    0.00692   -5.59  2.3e-08 ***
## UniqueAuthors2      0.12214    0.02507    4.87  1.1e-06 ***
## UniqueAuthors3      0.14123    0.02478    5.70  1.2e-08 ***
## UniqueAuthors4      0.17040    0.02482    6.87  6.9e-12 ***
## UniqueAuthors5      0.24168    0.02421    9.98  < 2e-16 ***
## Year1997            0.01916    0.01837    1.04    0.297
## Year1998            0.00784    0.01774    0.44    0.658
## Year1999           -0.03152    0.01750   -1.80    0.072 .
```

```
## Year2000            0.00753    0.01925    0.39    0.696
## Year2001           -0.03278    0.02448   -1.34    0.181
## Year2002           -0.01740    0.01711   -1.02    0.309
## Year2003            0.00199    0.01801    0.11    0.912
## Year2004           -0.01986    0.01679   -1.18    0.237
## Year2005           -0.00520    0.01670   -0.31    0.756
## Year2006           -0.02446    0.01635   -1.50    0.135
## Year2007           -0.02157    0.01635   -1.32    0.187
## Year2008           -0.01565    0.01664   -0.94    0.347
## Year2009           -0.01073    0.01677   -0.64    0.523
## Year2010            0.01414    0.01706    0.83    0.407
## Year2011            0.00229    0.01683    0.14    0.892
## Year2012            0.00747    0.01699    0.44    0.660
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.323
## Multiple R-squared:  0.0369, Adjusted R-squared:  0.0354
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
##  1158 weights are ~= 1. The remaining 12654 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0964  0.8690  0.9500  0.9010  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          7.24e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.013  1           1.007
## LastAuthorFemale  1.007  1           1.004
## Year              1.016 16           1.000
```



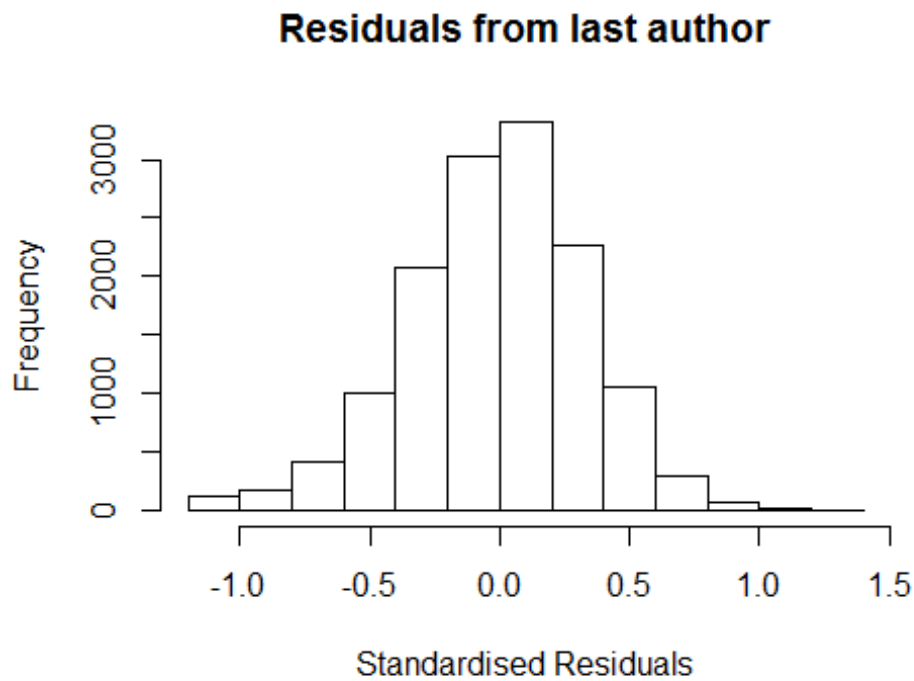

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.18048 -0.22373  0.00587  0.21654  1.30750
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.14731    0.01298   88.42  < 2e-16 ***
## FirstAuthorFemale1 -0.00721    0.00588   -1.23    0.220
## LastAuthorFemale1  -0.04386    0.00696   -6.30  3.1e-10 ***
## Year1997            0.01891    0.01848    1.02    0.306
## Year1998            0.00945    0.01779    0.53    0.595
## Year1999           -0.01729    0.01760   -0.98    0.326
## Year2000            0.01268    0.01935    0.66    0.512
## Year2001           -0.02300    0.02416   -0.95    0.341
## Year2002           -0.00118    0.01713   -0.07    0.945
## Year2003            0.01551    0.01809    0.86    0.391
## Year2004           -0.00564    0.01704   -0.33    0.741
## Year2005            0.01130    0.01685    0.67    0.502
```

```
## Year2006           -0.01196    0.01650   -0.72    0.469
## Year2007           -0.00715    0.01650   -0.43    0.665
## Year2008            0.00145    0.01679    0.09    0.931
## Year2009            0.01051    0.01678    0.63    0.531
## Year2010            0.03317    0.01717    1.93    0.053 .
## Year2011            0.01619    0.01692    0.96    0.338
## Year2012            0.02742    0.01701    1.61    0.107
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.326
## Multiple R-squared:  0.00503,    Adjusted R-squared:  0.00373
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  1173 weights are ~= 1. The remaining 12639 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0714  0.8690  0.9500  0.9000  0.9860  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          7.24e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.011  1           1.005
## Year              1.011 16           1.000
```




```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.17139 -0.22383  0.00539  0.21795  1.31655
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.138823   0.012942   87.99   <2e-16 ***
## FirstAuthorFemale1 -0.009559   0.005879   -1.63    0.104
## Year1997            0.019433   0.018503    1.05    0.294
## Year1998            0.009322   0.017835    0.52    0.601
## Year1999           -0.017708   0.017613   -1.01    0.315
## Year2000            0.012508   0.019402    0.64    0.519
## Year2001           -0.026124   0.024298   -1.08    0.282
## Year2002           -0.000882   0.017208   -0.05    0.959
## Year2003            0.015173   0.018153    0.84    0.403
## Year2004           -0.005823   0.017096   -0.34    0.733
## Year2005            0.010801   0.016920    0.64    0.523
## Year2006           -0.012046   0.016555   -0.73    0.467
```

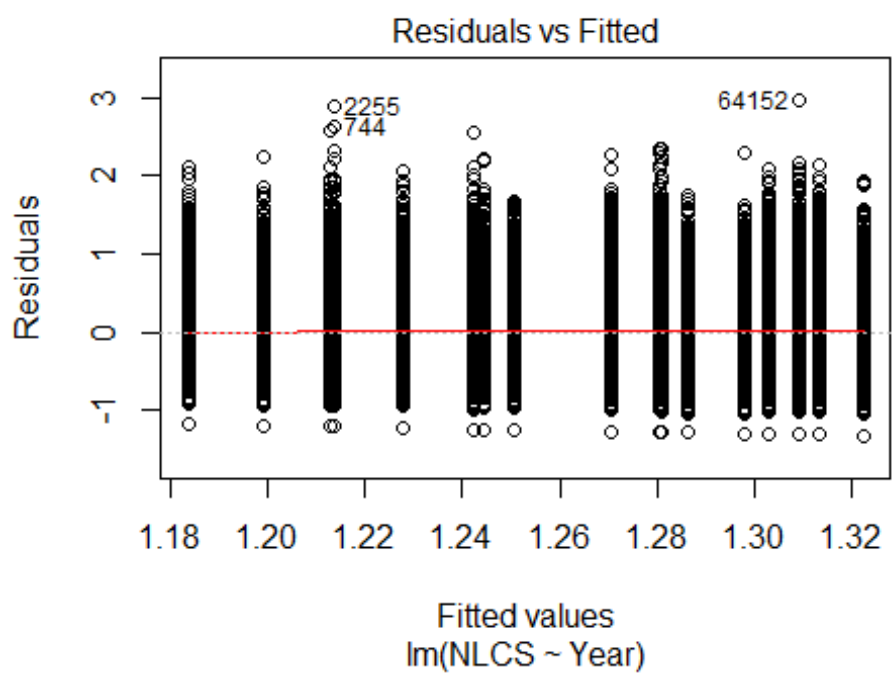
```
## Year2007           -0.005657   0.016554   -0.34    0.733
## Year2008            0.001149   0.016863    0.07    0.946
## Year2009            0.010404   0.016835    0.62    0.537
## Year2010            0.032566   0.017212    1.89    0.058 .
## Year2011            0.015629   0.016964    0.92    0.357
## Year2012            0.025713   0.017029    1.51    0.131
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.327
## Multiple R-squared:  0.00202,    Adjusted R-squared:  0.000795
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  1167 weights are ~= 1. The remaining 12645 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0673  0.8700  0.9500  0.9000  0.9860  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          7.24e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##                   GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.005  1           1.002
## Year             1.005 16           1.000
```



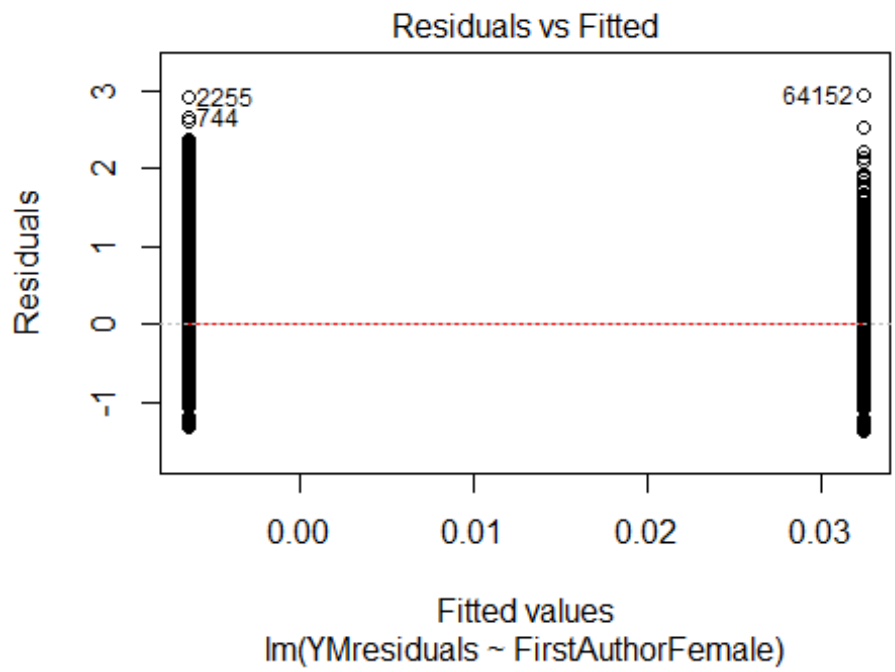

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.17711 -0.22330  0.00653  0.21664  1.31071
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)        1.145000   0.012844   89.15  < 2e-16 ***
## LastAuthorFemale1 -0.044409   0.006957   -6.38  1.8e-10 ***
## Year1997           0.018424   0.018470    1.00    0.319
## Year1998           0.009165   0.017787    0.52    0.606
## Year1999          -0.017788   0.017587   -1.01    0.312
## Year2000           0.012720   0.019358    0.66    0.511
## Year2001          -0.023508   0.024152   -0.97    0.330
## Year2002          -0.001623   0.017122   -0.09    0.924
## Year2003           0.015164   0.018086    0.84    0.402
## Year2004          -0.006418   0.017030   -0.38    0.706
## Year2005           0.010807   0.016836    0.64    0.521
## Year2006          -0.012637   0.016483   -0.77    0.443
```

```
## Year2007          -0.007724   0.016483   -0.47    0.639
## Year2008           0.000547   0.016765    0.03    0.974
## Year2009           0.009648   0.016752    0.58    0.565
## Year2010           0.032106   0.017131    1.87    0.061 .
## Year2011           0.015294   0.016903    0.90    0.366
## Year2012           0.026510   0.016989    1.56    0.119
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.326
## Multiple R-squared:  0.00491,    Adjusted R-squared:  0.00368
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  1182 weights are ~= 1. The remaining 12630 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0694  0.8690  0.9500  0.8990  0.9860  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          7.24e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Sample size for the above analysis:  13812"
## [1] ""
## [1] ""
## [1] "####################################"
## [1] "Analysis of AJSC 2500"
## [1] "####################################"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 3435 3388 3517 3305 3659 3727 3571 3367 3573 3855 4295 4516 4773 5319 5695
## 2011 2012
## 5795 5908
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 1126 1147 1217 1156 1304 1215 1410 1440 1524 1681 1933 2101 2168 2545 2849
## 2011 2012
```

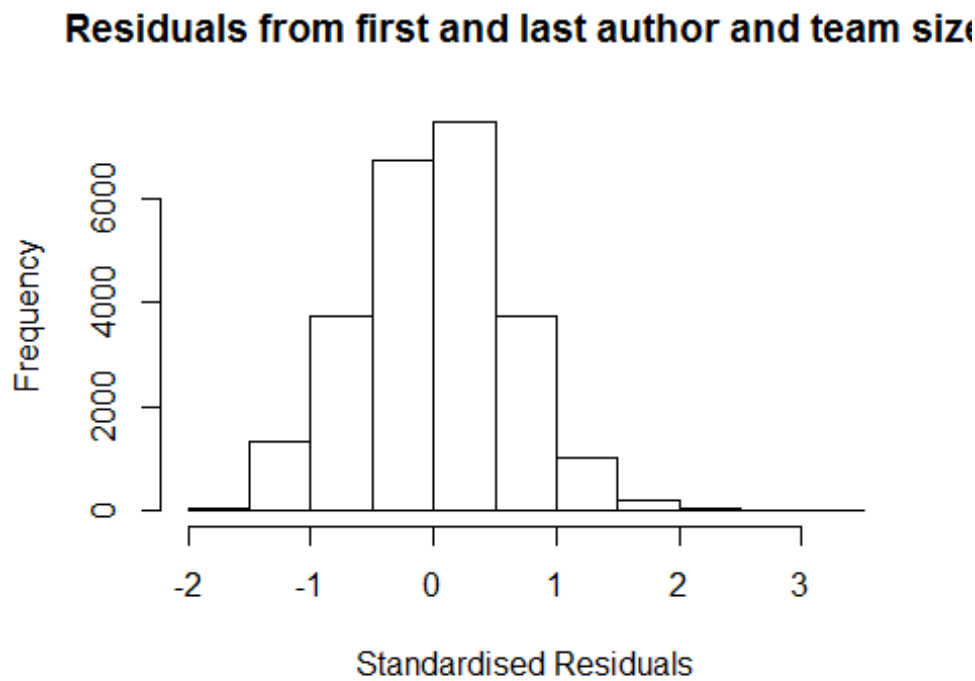
```
## 2949 2894
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  912  959  981  945 1067  974 1140 1114 1204 1323 1525 1673 1676 1965 2231
## 2011 2012
## 2277 2245
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
##  Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 160, df = 16, p-value <2e-16
```




```
##
##  Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 15, df = 1, p-value = 1e-04
```

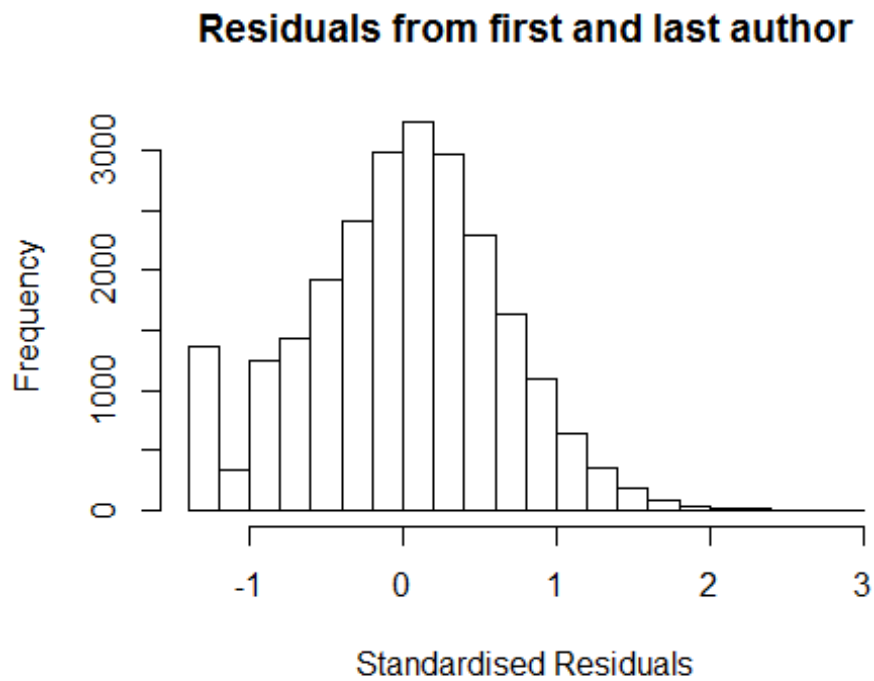



```
## [1] "Female first author team size 2018 geometric mean: 4.19835638960501"
## [1] "Male first author team size 2018 geometric mean: 3.76090601652609"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 620000, p-value = 4e-04
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 3.93733641857092"
## [1] "Male last author team size 2018 geometric mean: 3.82859646891818"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 480000, p-value = 0.5
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.046  1           1.023
## LastAuthorFemale  1.030  1           1.015
## UniqueAuthors     1.094  4           1.011
## Year              1.085 16           1.003
```



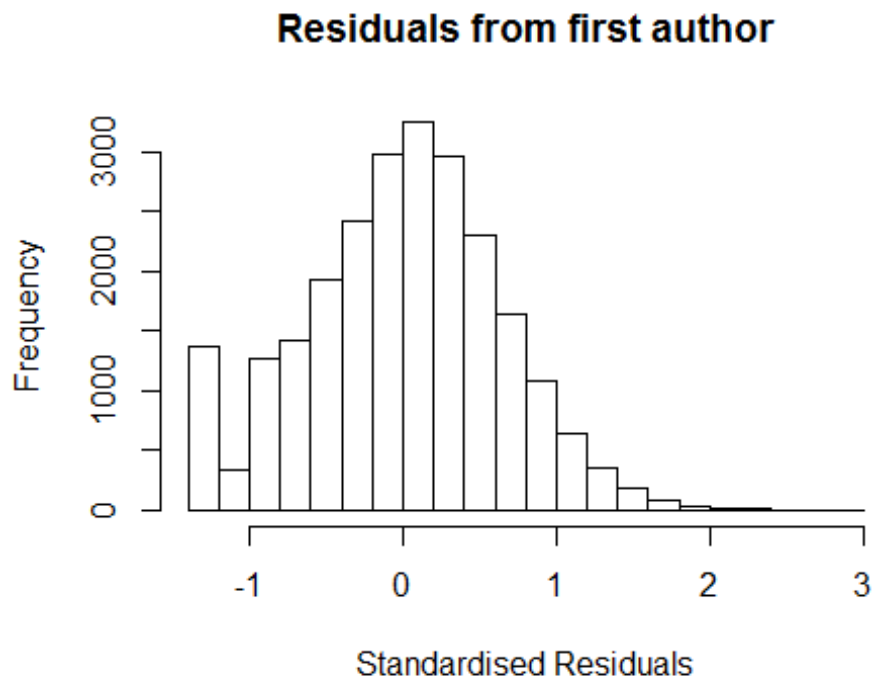

```
## [1] "List of  4 outliers with residuals above  2.5"
##          ScopusId  NLCS Year OneField Fields residuals
## 744    0030232761 3.847 1996     1502      6     2.630
## 2255   0029777844 4.113 1996     2210      3     3.158
## 9264   0031700973 3.543 1998     2210      3     2.533
## 64152 78650092372 4.275 2010     2200      3     2.801
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
UniqueAuthors +
##     Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##     k.max = 1000))
##  \--> method = "MM"
## Residuals:
##    Min     1Q Median     3Q    Max
## -1.518 -0.408  0.020  0.410  3.158
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)          0.95515    0.02657   35.95   <2e-16 ***
## FirstAuthorFemale1   0.00540    0.01087    0.50    0.619
## LastAuthorFemale1   -0.00757    0.01288   -0.59    0.556
## UniqueAuthors2       0.26136    0.01466   17.83   <2e-16 ***
## UniqueAuthors3       0.29961    0.01501   19.95   <2e-16 ***
## UniqueAuthors4       0.34074    0.01608   21.20   <2e-16 ***
## UniqueAuthors5       0.50828    0.01554   32.72   <2e-16 ***
```

```
## Year1997            0.03823    0.03263    1.17    0.241
## Year1998            0.05507    0.03467    1.59    0.112
## Year1999            0.05099    0.03302    1.54    0.122
## Year2000           -0.04165    0.03272   -1.27    0.203
## Year2001           -0.01317    0.03567   -0.37    0.712
## Year2002           -0.05762    0.03196   -1.80    0.071 .
## Year2003            0.02110    0.03040    0.69    0.488
## Year2004            0.04165    0.02999    1.39    0.165
## Year2005            0.01485    0.02917    0.51    0.611
## Year2006           -0.02462    0.02914   -0.84    0.398
## Year2007           -0.03870    0.02944   -1.31    0.189
## Year2008            0.00640    0.02882    0.22    0.824
## Year2009           -0.01173    0.02844   -0.41    0.680
## Year2010            0.00536    0.02803    0.19    0.848
## Year2011           -0.00633    0.02773   -0.23    0.819
## Year2012            0.01296    0.02792    0.46    0.643
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.61
## Multiple R-squared:  0.0583, Adjusted R-squared:  0.0574
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
##  observation 607 is an outlier with |weight| = 0 ( < 4.1e-06);
##  2049 weights are ~= 1. The remaining 22161 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0016  0.8620  0.9510  0.9040  0.9860  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          4.13e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.027  1           1.013
## LastAuthorFemale  1.024  1           1.012
## Year              1.013 16           1.000
```



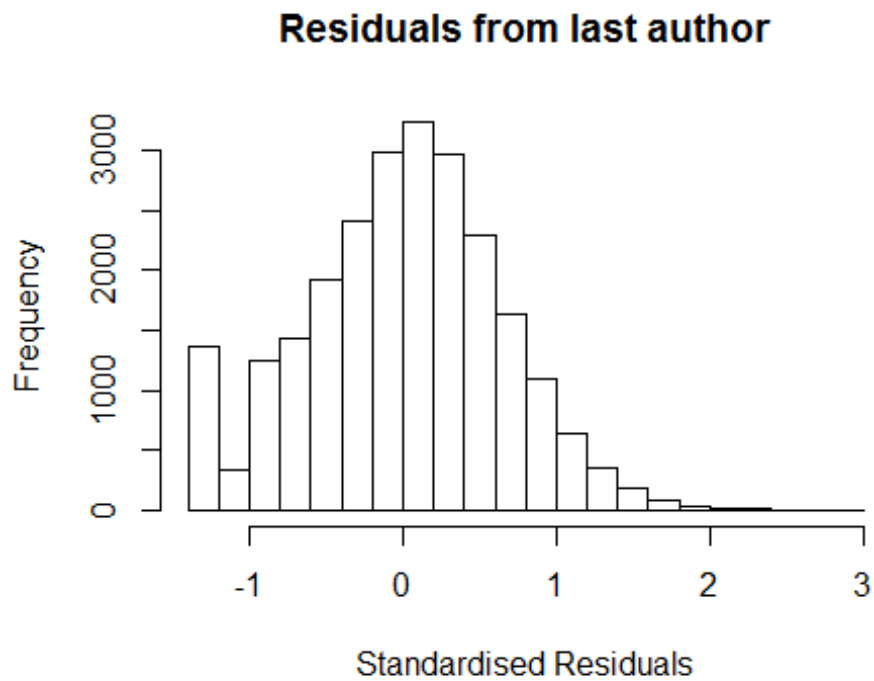

```
## [1] "List of  4 outliers with residuals above  2.5"
##          ScopusId  NLCS Year OneField Fields residuals
## 744    0030232761 3.847 1996     1502      6     2.680
## 2255   0029777844 4.113 1996     2210      3     2.946
## 47608 34249742469 3.790 2007     1600      2     2.534
## 64152 78650092372 4.275 2010     2200      3     2.950
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.3518 -0.4217  0.0263  0.4220  2.9500
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.16725    0.02575   45.33  < 2e-16 ***
## FirstAuthorFemale1  0.04501    0.01107    4.07  4.8e-05 ***
## LastAuthorFemale1   0.00208    0.01324    0.16  0.87491
## Year1997            0.03359    0.03381    0.99  0.32041
## Year1998            0.07032    0.03552    1.98  0.04777 *
## Year1999            0.08118    0.03416    2.38  0.01748 *
## Year2000           -0.00438    0.03365   -0.13  0.89653
## Year2001            0.03445    0.03615    0.95  0.34050
## Year2002            0.00269    0.03299    0.08  0.93492
```

```
## Year2003            0.08983    0.03152    2.85  0.00437 **
## Year2004            0.12365    0.03116    3.97  7.3e-05 ***
## Year2005            0.10011    0.03037    3.30  0.00098 ***
## Year2006            0.05172    0.03038    1.70  0.08872 .
## Year2007            0.04393    0.03078    1.43  0.15351
## Year2008            0.09375    0.03005    3.12  0.00181 **
## Year2009            0.08498    0.02971    2.86  0.00424 **
## Year2010            0.11273    0.02915    3.87  0.00011 ***
## Year2011            0.10592    0.02890    3.66  0.00025 ***
## Year2012            0.13741    0.02902    4.74  2.2e-06 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.625
## Multiple R-squared:  0.0051, Adjusted R-squared:  0.00436
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
##  2 observations c(607,18210) are outliers with |weight| = 0 ( < 4.1e-06);
##  2061 weights are ~= 1. The remaining 22148 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0266  0.8610  0.9510  0.9040  0.9860  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          4.13e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.008  1           1.004
## Year              1.008 16           1.000
```




```
## [1] "List of  4 outliers with residuals above  2.5"
##          ScopusId  NLCS Year OneField Fields residuals
## 744    0030232761 3.847 1996     1502      6     2.680
## 2255   0029777844 4.113 1996     2210      3     2.946
## 47608 34249742469 3.790 2007     1600      2     2.534
## 64152 78650092372 4.275 2010     2200      3     2.950
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.3502 -0.4217  0.0266  0.4218  2.9495
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.16738    0.02574   45.36  < 2e-16 ***
## FirstAuthorFemale1  0.04530    0.01097    4.13  3.7e-05 ***
## Year1997            0.03364    0.03380    1.00  0.31963
## Year1998            0.07032    0.03552    1.98  0.04777 *
## Year1999            0.08120    0.03416    2.38  0.01746 *
## Year2000           -0.00436    0.03364   -0.13  0.89701
## Year2001            0.03450    0.03614    0.95  0.33983
## Year2002            0.00276    0.03299    0.08  0.93334
## Year2003            0.08986    0.03152    2.85  0.00436 **
```

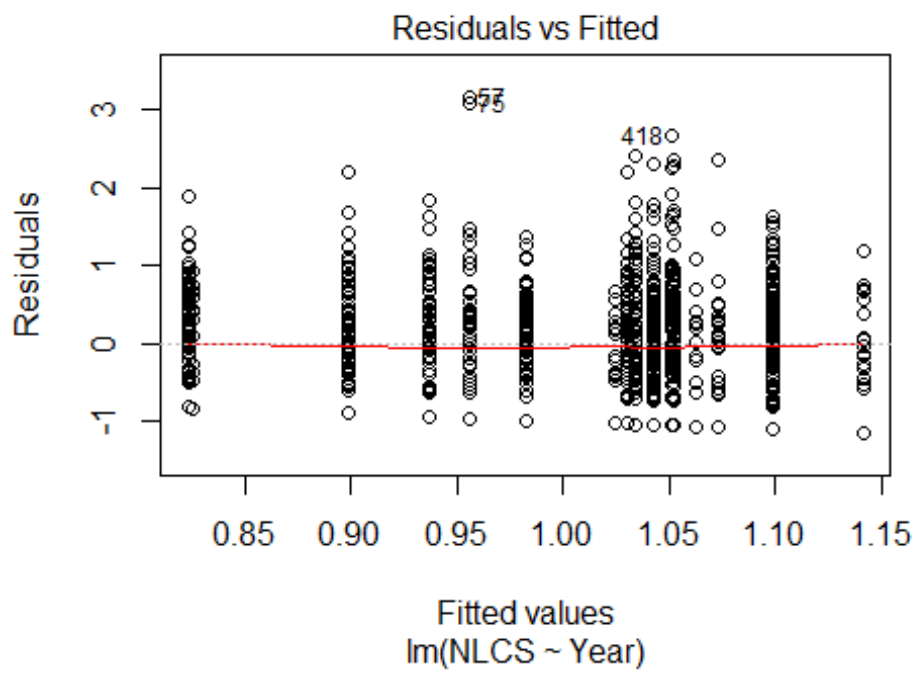
```
## Year2004            0.12372    0.03116    3.97  7.2e-05 ***
## Year2005            0.10018    0.03037    3.30  0.00097 ***
## Year2006            0.05179    0.03038    1.71  0.08819 .
## Year2007            0.04398    0.03077    1.43  0.15292
## Year2008            0.09383    0.03004    3.12  0.00179 **
## Year2009            0.08506    0.02971    2.86  0.00420 **
## Year2010            0.11282    0.02914    3.87  0.00011 ***
## Year2011            0.10602    0.02889    3.67  0.00024 ***
## Year2012            0.13755    0.02900    4.74  2.1e-06 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.625
## Multiple R-squared:  0.00509,    Adjusted R-squared:  0.00439
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
##  2 observations c(607,18210) are outliers with |weight| = 0 ( < 4.1e-06);
##  2058 weights are ~= 1. The remaining 22151 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0267  0.8610  0.9510  0.9040  0.9860  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          4.13e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                    psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##                   GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.005  1           1.003
## Year             1.005 16           1.000
```



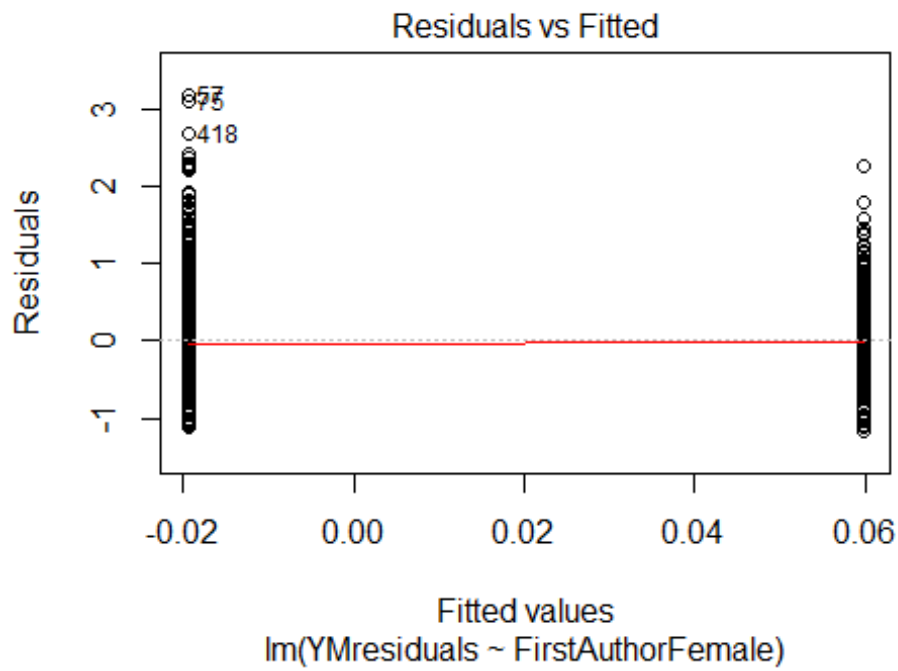

```
## [1] "List of  4 outliers with residuals above  2.5"
##          ScopusId  NLCS Year OneField Fields residuals
## 744    0030232761 3.847 1996     1502      6     2.680
## 2255   0029777844 4.113 1996     2210      3     2.946
## 47608 34249742469 3.790 2007     1600      2     2.534
## 64152 78650092372 4.275 2010     2200      3     2.950
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.3238 -0.4226  0.0262  0.4232  2.9874
##
## Coefficients:
##                   Estimate Std. Error t value Pr(>|t|)
## (Intercept)        1.17199    0.02572   45.56  < 2e-16 ***
## LastAuthorFemale1  0.01064    0.01311    0.81  0.41698
## Year1997           0.03340    0.03383    0.99  0.32353
## Year1998           0.07151    0.03551    2.01  0.04406 *
## Year1999           0.08199    0.03416    2.40  0.01638 *
## Year2000          -0.00408    0.03365   -0.12  0.90343
## Year2001           0.03573    0.03616    0.99  0.32320
## Year2002           0.00383    0.03298    0.12  0.90757
## Year2003           0.09223    0.03153    2.93  0.00344 **
```

```
## Year2004           0.12430    0.03116    3.99  6.6e-05 ***
## Year2005           0.10179    0.03036    3.35  0.00080 ***
## Year2006           0.05381    0.03036    1.77  0.07634 .
## Year2007           0.04562    0.03077    1.48  0.13820
## Year2008           0.09541    0.03005    3.18  0.00150 **
## Year2009           0.08721    0.02970    2.94  0.00333 **
## Year2010           0.11560    0.02913    3.97  7.3e-05 ***
## Year2011           0.10897    0.02889    3.77  0.00016 ***
## Year2012           0.14117    0.02899    4.87  1.1e-06 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.625
## Multiple R-squared:  0.00444,    Adjusted R-squared:  0.00375
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
##  2 observations c(607,18210) are outliers with |weight| = 0 ( < 4.1e-06);
##  2080 weights are ~= 1. The remaining 22129 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0278  0.8610  0.9500  0.9040  0.9860  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          4.13e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Sample size for the above analysis:  24211"
## [1] ""
## [1] ""
## [1] "##################################"
## [1] "Analysis of AJSC 2501"
## [1] "##################################"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  160  147  169  168  165  135  132  132  145  188  201  208  152   42   39
## 2011 2012
##   36   33
```

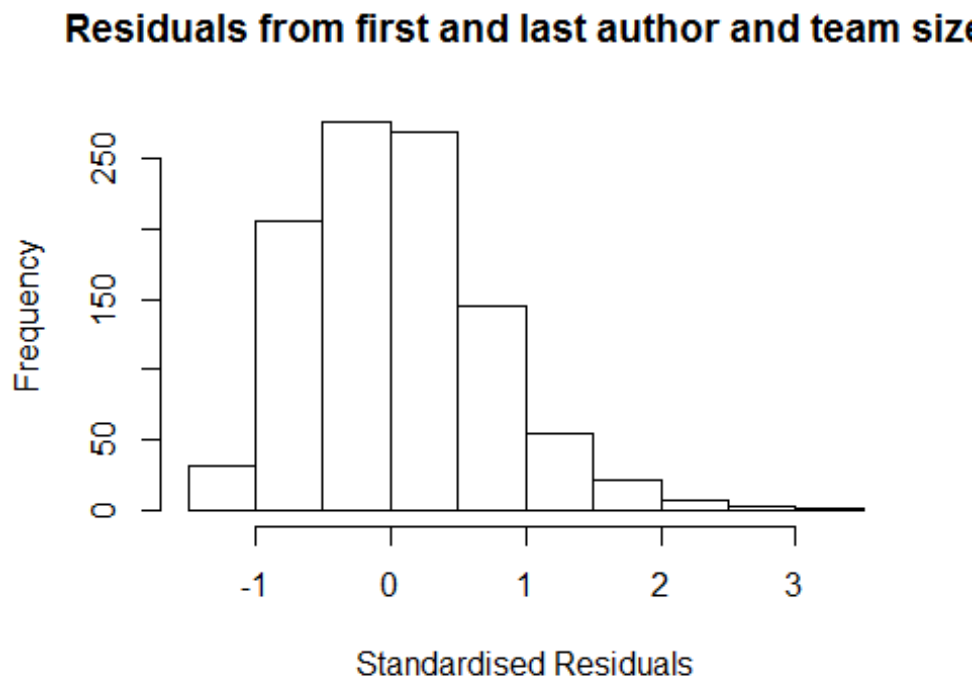
```
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##   76   83   79   73   73   31   82   88   91  116  119  130   84   17   18
## 2011 2012
##   22   17
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##   68   73   64   65   62   30   66   73   83  103   98  104   73   15   11
## 2011 2012
##   15   11
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
##  Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 40, df = 16, p-value = 7e-04
```




```
##
##  Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 2.6, df = 1, p-value = 0.1
```



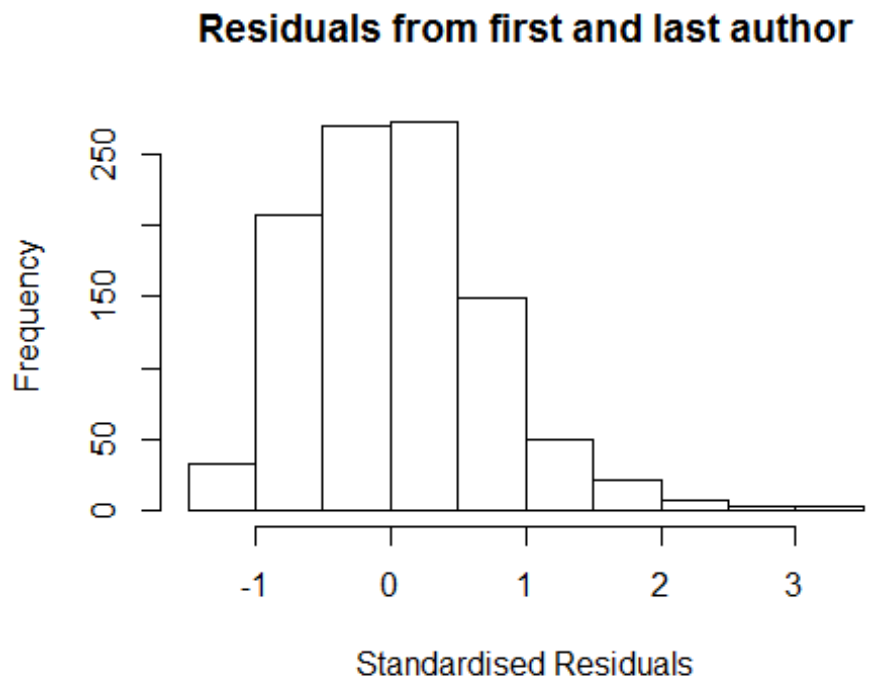

```
## [1] "Female first author team size 2018 geometric mean: 2.51258193288905"
## [1] "Male first author team size 2018 geometric mean: 3.35409819379607"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 1700, p-value = 0.02
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 2.60769675040748"
## [1] "Male last author team size 2018 geometric mean: 3.19948272192689"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 1800, p-value = 0.1
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.598  1           1.264
## LastAuthorFemale  1.541  1           1.241
## UniqueAuthors     1.328  4           1.036
## Year              1.514 16           1.013
```




```
## [1] "List of  5 outliers with residuals above  2.5"
##        ScopusId  NLCS Year OneField Fields residuals
## 57   0030284644 4.112 1996     2501      1     3.231
## 75   0030188676 4.047 1996     2501      1     3.414
## 179 24044543390 3.438 1997     2209      2     2.766
## 418  0032092047 3.725 1998     1501      2     2.867
## 905  0009735051 3.425 2001     2501      1     2.669
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
UniqueAuthors +
##     Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##     k.max = 1000))
##  \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.28126 -0.46592 -0.00793  0.44574  3.41412
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         0.63288    0.08202    7.72  2.9e-14 ***
## FirstAuthorFemale1  0.02823    0.06082    0.46    0.643
## LastAuthorFemale1   0.14183    0.06308    2.25    0.025 *
## UniqueAuthors2      0.24818    0.06034    4.11  4.2e-05 ***
## UniqueAuthors3      0.30549    0.06059    5.04  5.5e-07 ***
## UniqueAuthors4      0.42449    0.07802    5.44  6.7e-08 ***
```

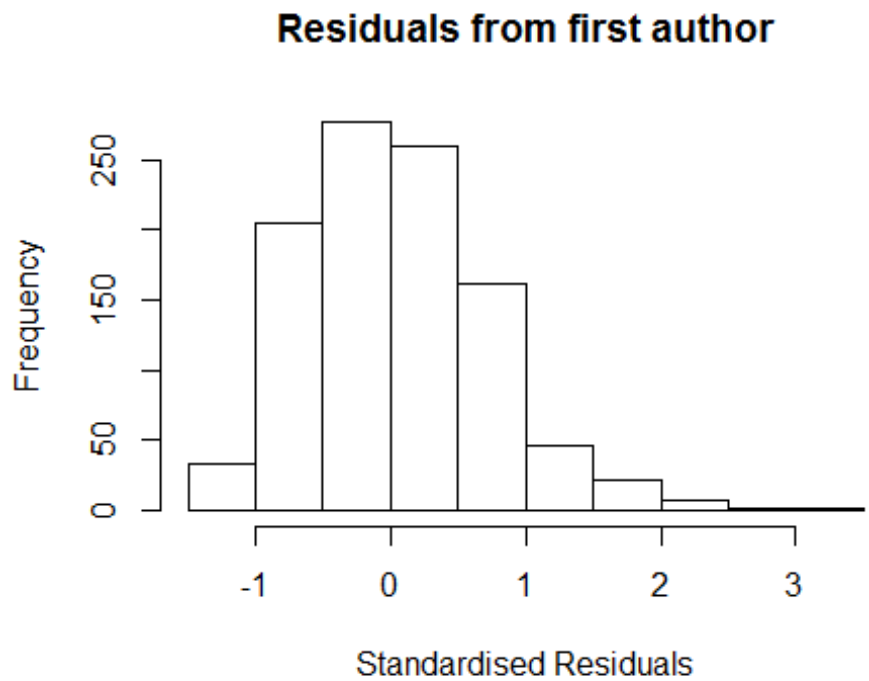
```
## UniqueAuthors5      0.40065    0.09046    4.43  1.1e-05 ***
## Year1997            0.03912    0.12479    0.31    0.754
## Year1998           -0.02287    0.12422   -0.18    0.854
## Year1999            0.19969    0.11564    1.73    0.085 .
## Year2000            0.09455    0.11154    0.85    0.397
## Year2001            0.12299    0.14797    0.83    0.406
## Year2002            0.03369    0.11468    0.29    0.769
## Year2003            0.00197    0.11851    0.02    0.987
## Year2004            0.04335    0.10778    0.40    0.688
## Year2005            0.17283    0.09947    1.74    0.083 .
## Year2006            0.05403    0.10218    0.53    0.597
## Year2007           -0.10677    0.09393   -1.14    0.256
## Year2008           -0.07125    0.10968   -0.65    0.516
## Year2009            0.00168    0.12374    0.01    0.989
## Year2010            0.08304    0.23237    0.36    0.721
## Year2011            0.07758    0.14586    0.53    0.595
## Year2012           -0.19420    0.15937   -1.22    0.223
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.633
## Multiple R-squared:  0.0804, Adjusted R-squared:  0.06
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
##  2 observations c(28,36) are outliers with |weight| = 0 ( < 9.9e-05);
##  81 weights are ~= 1. The remaining 931 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0042  0.8730  0.9430  0.8990  0.9860  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          9.86e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.592  1           1.262
```

```
## LastAuthorFemale  1.521  1           1.233
## Year              1.214 16           1.006
```



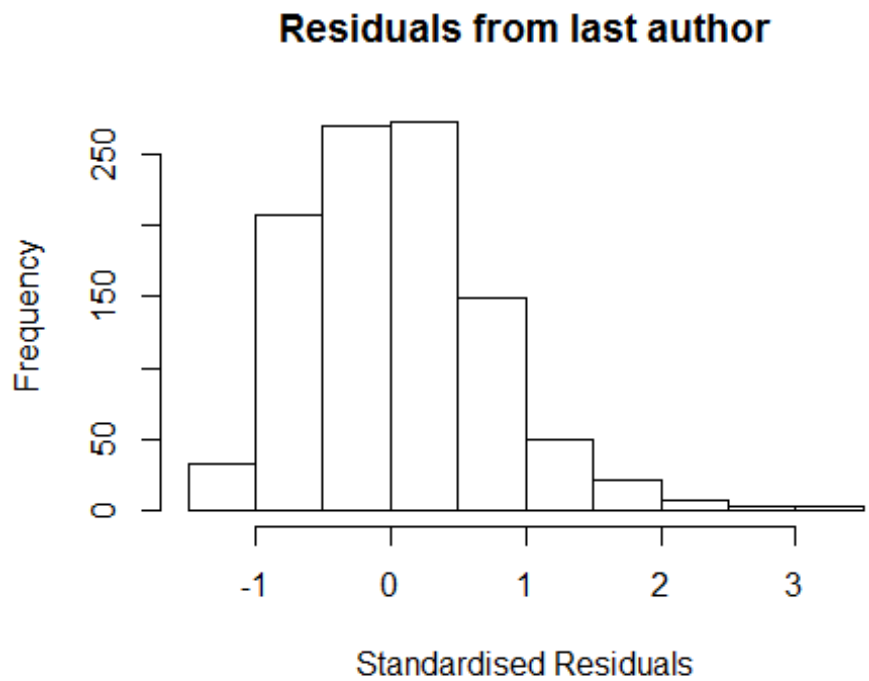

```
## [1] "List of  4 outliers with residuals above  2.5"
##        ScopusId  NLCS Year OneField Fields residuals
## 57   0030284644 4.112 1996     2501      1     3.303
## 75   0030188676 4.047 1996     2501      1     3.238
## 179 24044543390 3.438 1997     2209      2     2.584
## 418  0032092047 3.725 1998     1501      2     2.879
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.24089 -0.44836 -0.00847  0.45770  3.30306
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)           0.8089     0.0816    9.91   <2e-16 ***
## FirstAuthorFemale1    0.0573     0.0616    0.93    0.353
## LastAuthorFemale1     0.1291     0.0641    2.02    0.044 *
## Year1997              0.0455     0.1257    0.36    0.717
## Year1998              0.0374     0.1286    0.29    0.771
## Year1999              0.2455     0.1192    2.06    0.040 *
```

```
## Year2000             0.1398     0.1133    1.23    0.217
## Year2001             0.1774     0.1587    1.12    0.264
## Year2002             0.0993     0.1172    0.85    0.397
## Year2003             0.0292     0.1187    0.25    0.806
## Year2004             0.1157     0.1122    1.03    0.303
## Year2005             0.2338     0.1035    2.26    0.024 *
## Year2006             0.1099     0.1070    1.03    0.304
## Year2007            -0.0706     0.0993   -0.71    0.477
## Year2008            -0.0305     0.1163   -0.26    0.793
## Year2009             0.0852     0.1314    0.65    0.517
## Year2010             0.0753     0.2368    0.32    0.750
## Year2011             0.1325     0.1551    0.85    0.393
## Year2012            -0.1504     0.1899   -0.79    0.429
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.693
## Multiple R-squared:  0.031,  Adjusted R-squared:  0.0135
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  2 observations c(28,36) are outliers with |weight| <= 4.4e-05 ( < 9.9e-
05);
##  107 weights are ~= 1. The remaining 905 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0462  0.8710  0.9470  0.9080  0.9840  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          9.86e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##                     GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.155  1           1.075
## Year              1.155 16           1.005
```




```
## [1] "List of  4 outliers with residuals above  2.5"
##        ScopusId  NLCS Year OneField Fields residuals
## 57   0030284644 4.112 1996     2501      1     3.303
## 75   0030188676 4.047 1996     2501      1     3.238
## 179 24044543390 3.438 1997     2209      2     2.584
## 418  0032092047 3.725 1998     1501      2     2.879
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.18499 -0.45703 -0.00841  0.46478  3.28829
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)          0.8237     0.0824    9.99   <2e-16 ***
## FirstAuthorFemale1   0.1273     0.0527    2.41    0.016 *
## Year1997             0.0586     0.1267    0.46    0.644
## Year1998             0.0387     0.1296    0.30    0.765
## Year1999             0.2340     0.1197    1.96    0.051 .
## Year2000             0.1346     0.1146    1.17    0.241
## Year2001             0.1797     0.1589    1.13    0.258
## Year2002             0.1003     0.1169    0.86    0.391
## Year2003             0.0247     0.1199    0.21    0.837
```

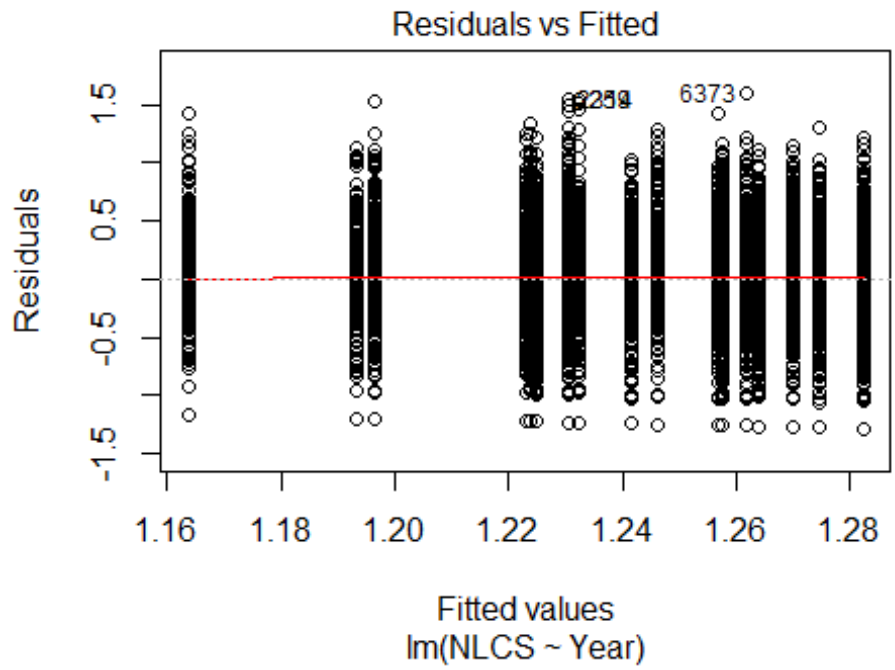
```
## Year2004             0.1153     0.1121    1.03    0.304
## Year2005             0.2337     0.1045    2.24    0.026 *
## Year2006             0.1081     0.1082    1.00    0.318
## Year2007            -0.0770     0.1004   -0.77    0.443
## Year2008            -0.0337     0.1168   -0.29    0.773
## Year2009             0.0880     0.1341    0.66    0.512
## Year2010             0.0937     0.2398    0.39    0.696
## Year2011             0.1517     0.1620    0.94    0.349
## Year2012            -0.1706     0.1940   -0.88    0.379
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.701
## Multiple R-squared:  0.0263, Adjusted R-squared:  0.00972
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  observation 28 is an outlier with |weight| = 0 ( < 9.9e-05);
##  99 weights are ~= 1. The remaining 914 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0013  0.8710  0.9490  0.9100  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          9.86e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##                   GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.117  1           1.057
## Year             1.117 16           1.003
```



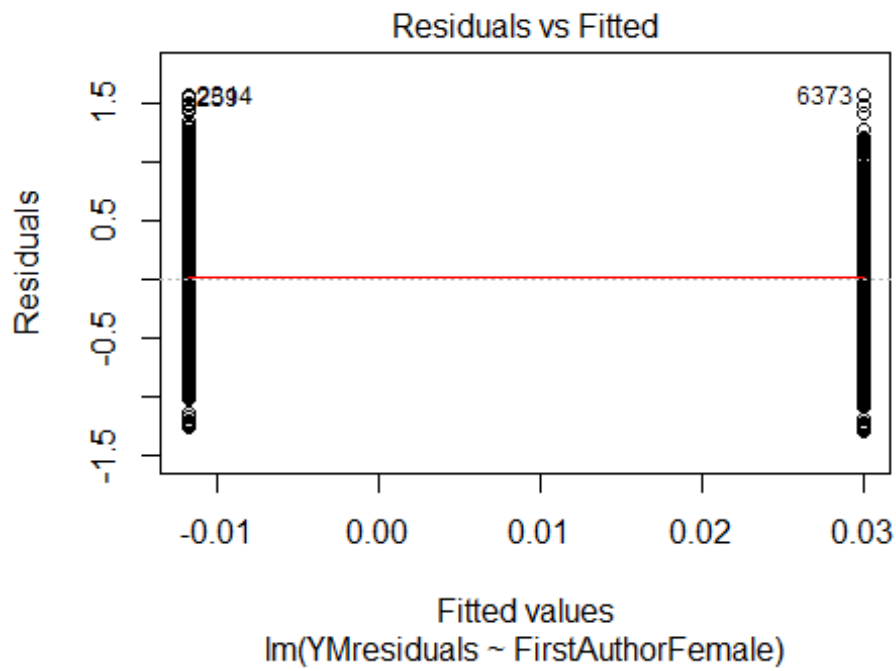

```
## [1] "List of  4 outliers with residuals above  2.5"
##        ScopusId  NLCS Year OneField Fields residuals
## 57   0030284644 4.112 1996     2501      1     3.303
## 75   0030188676 4.047 1996     2501      1     3.238
## 179 24044543390 3.438 1997     2209      2     2.584
## 418  0032092047 3.725 1998     1501      2     2.879
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.21719 -0.45145 -0.00368  0.45160  3.29559
##
## Coefficients:
##                   Estimate Std. Error t value Pr(>|t|)
## (Intercept)         0.8164     0.0813   10.04   <2e-16 ***
## LastAuthorFemale1   0.1615     0.0548    2.95   0.0033 **
## Year1997            0.0471     0.1258    0.37   0.7083
## Year1998            0.0334     0.1289    0.26   0.7953
## Year1999            0.2393     0.1192    2.01   0.0451 *
## Year2000            0.1352     0.1129    1.20   0.2313
## Year2001            0.1752     0.1589    1.10   0.2706
## Year2002            0.1013     0.1173    0.86   0.3879
## Year2003            0.0291     0.1184    0.25   0.8060
```

```
## Year2004            0.1100     0.1119    0.98   0.3258
## Year2005            0.2339     0.1032    2.27   0.0237 *
## Year2006            0.1112     0.1070    1.04   0.2986
## Year2007           -0.0715     0.0994   -0.72   0.4723
## Year2008           -0.0307     0.1167   -0.26   0.7924
## Year2009            0.0913     0.1305    0.70   0.4844
## Year2010            0.0921     0.2363    0.39   0.6969
## Year2011            0.1418     0.1509    0.94   0.3474
## Year2012           -0.1418     0.1919   -0.74   0.4599
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.693
## Multiple R-squared:  0.0303, Adjusted R-squared:  0.0137
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  2 observations c(28,36) are outliers with |weight| <= 7.4e-05 ( < 9.9e-
05);
##  101 weights are ~= 1. The remaining 911 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0461  0.8690  0.9490  0.9090  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          9.86e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Sample size for the above analysis:  1014"
## [1] ""
## [1] ""
## [1] "####################################"
## [1] "Analysis of AJSC 2502"
## [1] "####################################"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  514  495  514  520  688  499  566  576  699  816  906  969 1179 1471 1677
## 2011 2012
```

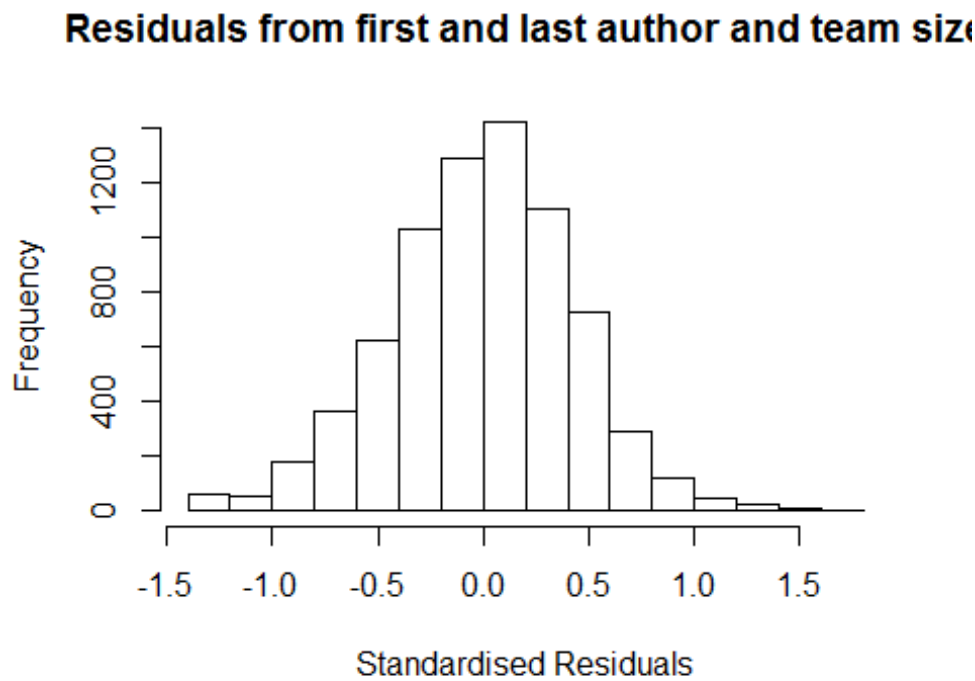
```
## 1638 1753
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  290  252  284  299  355  241  301  315  365  479  556  563  731  867 1024
## 2011 2012
##  980 1051
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  244  211  244  253  284  205  257  262  285  398  446  466  586  704  835
## 2011 2012
##  799  847
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
##  Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 170, df = 16, p-value <2e-16
```




```
##
##  Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 14, df = 1, p-value = 2e-04
```

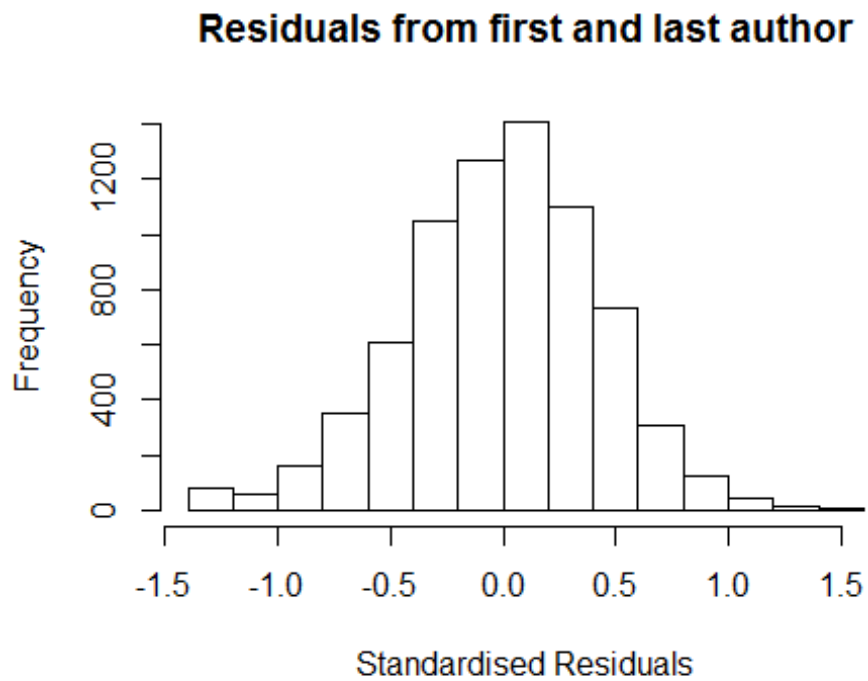



```
## [1] "Female first author team size 2018 geometric mean: 5.30976039226861"
## [1] "Male first author team size 2018 geometric mean: 4.73519110585622"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 67000, p-value = 0.04
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 5.14532186511055"
## [1] "Male last author team size 2018 geometric mean: 4.85459671279161"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 53000, p-value = 0.5
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.034  1           1.017
## LastAuthorFemale  1.024  1           1.012
## UniqueAuthors     1.141  4           1.017
## Year              1.149 16           1.004
```



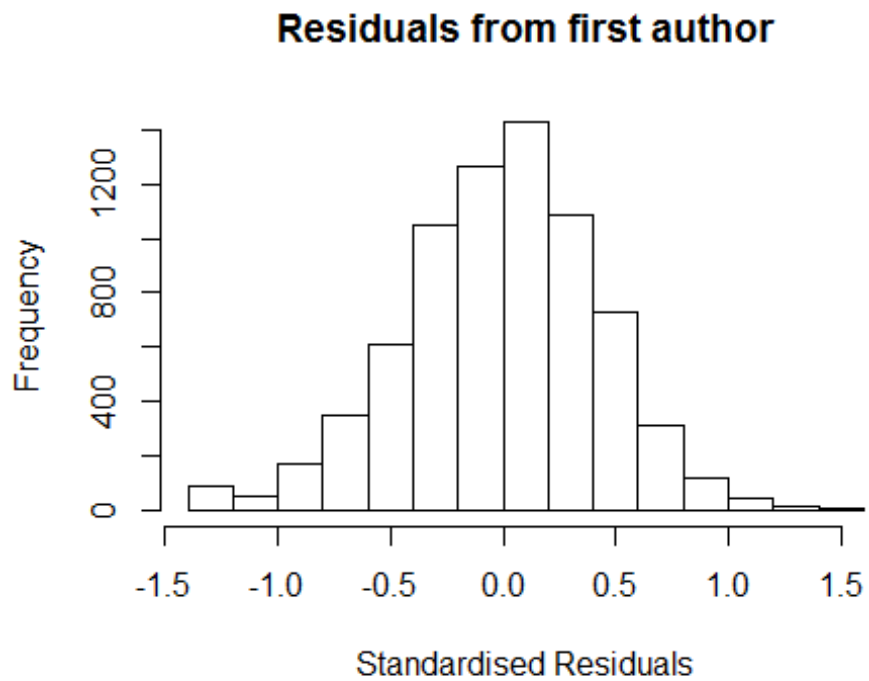

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
UniqueAuthors +
##     Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##     k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.3831 -0.2806  0.0102  0.2799  1.7622
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         0.96744    0.04017   24.08  < 2e-16 ***
## FirstAuthorFemale1  0.03248    0.01100    2.95   0.0032 **
## LastAuthorFemale1   0.04992    0.01255    3.98  7.0e-05 ***
## UniqueAuthors2      0.22296    0.03092    7.21  6.1e-13 ***
## UniqueAuthors3      0.26471    0.03073    8.61  < 2e-16 ***
## UniqueAuthors4      0.30511    0.03086    9.89  < 2e-16 ***
## UniqueAuthors5      0.32844    0.02981   11.02  < 2e-16 ***
## Year1997           -0.00568    0.04468   -0.13   0.8988
## Year1998           -0.02231    0.04179   -0.53   0.5935
## Year1999           -0.07251    0.04230   -1.71   0.0866 .
```

```
## Year2000            0.05031    0.04713    1.07   0.2857
## Year2001            0.02813    0.04661    0.60   0.5462
## Year2002           -0.04282    0.04248   -1.01   0.3135
## Year2003            0.02493    0.03921    0.64   0.5249
## Year2004            0.01007    0.03998    0.25   0.8012
## Year2005           -0.02239    0.03669   -0.61   0.5418
## Year2006            0.00379    0.03674    0.10   0.9179
## Year2007            0.02306    0.03548    0.65   0.5158
## Year2008           -0.02516    0.03478   -0.72   0.4695
## Year2009           -0.02478    0.03336   -0.74   0.4576
## Year2010            0.00651    0.03267    0.20   0.8421
## Year2011           -0.01061    0.03297   -0.32   0.7477
## Year2012           -0.05383    0.03296   -1.63   0.1025
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.413
## Multiple R-squared:  0.0383, Adjusted R-squared:  0.0354
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
##  670 weights are ~= 1. The remaining 6656 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0288  0.8680  0.9480  0.9000  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.37e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.021  1           1.011
## LastAuthorFemale  1.017  1           1.009
## Year              1.036 16           1.001
```



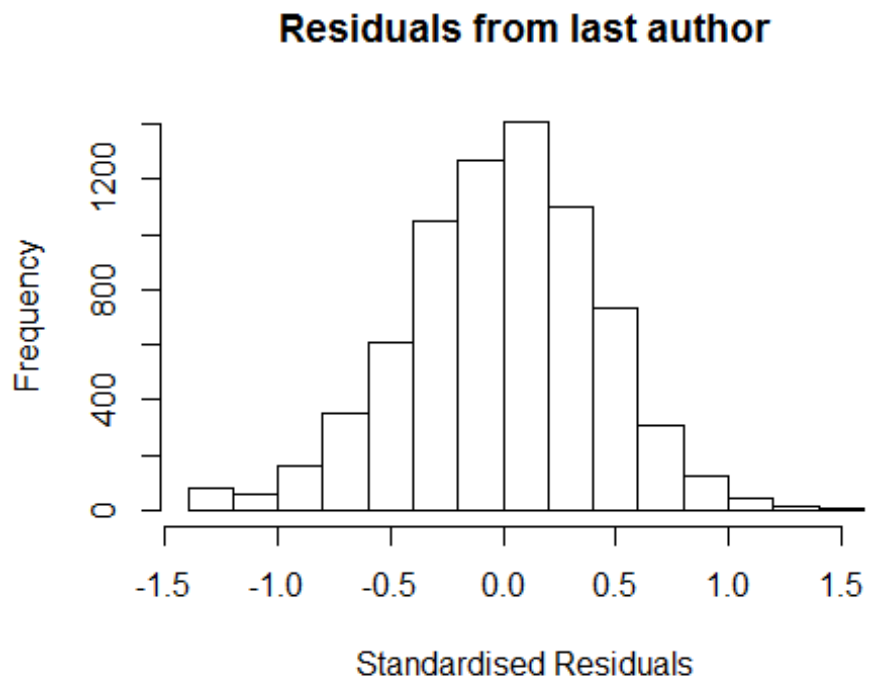

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##    Min     1Q Median     3Q    Max
## -1.344 -0.282  0.010  0.288  1.589
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)          1.21621    0.02966   41.01  < 2e-16 ***
## FirstAuthorFemale1   0.04294    0.01102    3.90  9.9e-05 ***
## LastAuthorFemale1    0.04992    0.01264    3.95  8.0e-05 ***
## Year1997            -0.00780    0.04435   -0.18     0.86
## Year1998            -0.02590    0.04168   -0.62     0.53
## Year1999            -0.06595    0.04184   -1.58     0.12
## Year2000             0.04881    0.04708    1.04     0.30
## Year2001             0.03446    0.04717    0.73     0.47
## Year2002            -0.02529    0.04295   -0.59     0.56
## Year2003             0.03600    0.03912    0.92     0.36
## Year2004             0.02339    0.04041    0.58     0.56
## Year2005             0.00775    0.03671    0.21     0.83
```

```
## Year2006            0.03313    0.03693    0.90     0.37
## Year2007            0.05248    0.03560    1.47     0.14
## Year2008            0.01046    0.03498    0.30     0.76
## Year2009            0.01196    0.03338    0.36     0.72
## Year2010            0.04478    0.03270    1.37     0.17
## Year2011            0.03035    0.03283    0.92     0.36
## Year2012           -0.01626    0.03292   -0.49     0.62
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.419
## Multiple R-squared:  0.00929,    Adjusted R-squared:  0.00685
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
##  660 weights are ~= 1. The remaining 6666 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.118   0.869   0.949   0.901   0.985   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.37e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.021  1           1.010
## Year              1.021 16           1.001
```




```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.31875 -0.28260  0.00892  0.28572  1.57495
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.22109    0.02980   40.97  < 2e-16 ***
## FirstAuthorFemale1  0.04765    0.01106    4.31  1.7e-05 ***
## Year1997           -0.00682    0.04434   -0.15     0.88
## Year1998           -0.02677    0.04176   -0.64     0.52
## Year1999           -0.06607    0.04204   -1.57     0.12
## Year2000            0.05002    0.04733    1.06     0.29
## Year2001            0.03866    0.04751    0.81     0.42
## Year2002           -0.02502    0.04315   -0.58     0.56
## Year2003            0.04054    0.03924    1.03     0.30
## Year2004            0.02615    0.04061    0.64     0.52
## Year2005            0.01232    0.03684    0.33     0.74
## Year2006            0.03557    0.03706    0.96     0.34
```

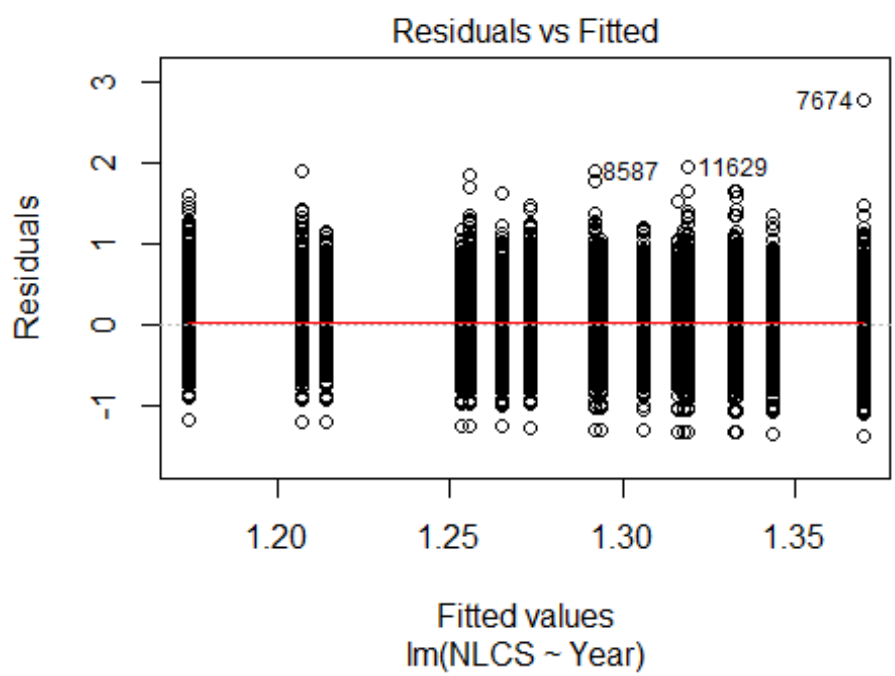
```
## Year2007            0.05365    0.03576    1.50     0.13
## Year2008            0.01358    0.03512    0.39     0.70
## Year2009            0.01618    0.03349    0.48     0.63
## Year2010            0.04815    0.03283    1.47     0.14
## Year2011            0.03584    0.03293    1.09     0.28
## Year2012           -0.01239    0.03304   -0.38     0.71
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.419
## Multiple R-squared:  0.00727,    Adjusted R-squared:  0.00496
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
##  664 weights are ~= 1. The remaining 6662 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.126   0.869   0.949   0.900   0.985   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.37e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##                   GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.017  1           1.008
## Year             1.017 16           1.001
```



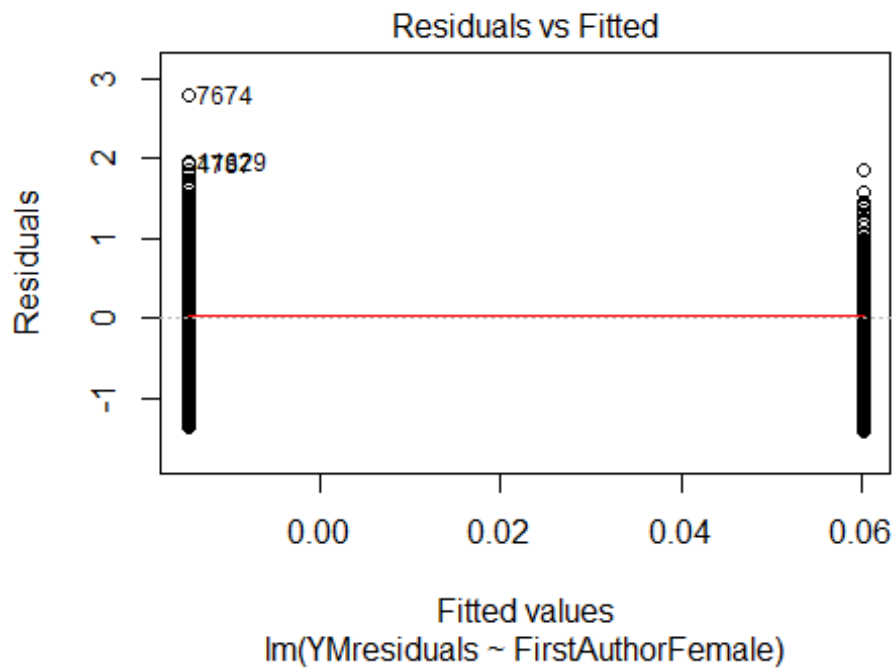

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.32894 -0.28493  0.00934  0.28784  1.61941
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.22244    0.02972   41.12  < 2e-16 ***
## LastAuthorFemale1   0.05544    0.01271    4.36  1.3e-05 ***
## Year1997           -0.00555    0.04439   -0.12     0.90
## Year1998           -0.02482    0.04178   -0.59     0.55
## Year1999           -0.06229    0.04198   -1.48     0.14
## Year2000            0.05192    0.04727    1.10     0.27
## Year2001            0.03536    0.04721    0.75     0.45
## Year2002           -0.02145    0.04288   -0.50     0.62
## Year2003            0.03833    0.03929    0.98     0.33
## Year2004            0.02887    0.04042    0.71     0.48
## Year2005            0.01415    0.03677    0.38     0.70
## Year2006            0.03698    0.03702    1.00     0.32
```

```
## Year2007           0.05853    0.03563    1.64     0.10
## Year2008           0.01688    0.03503    0.48     0.63
## Year2009           0.01972    0.03340    0.59     0.56
## Year2010           0.05106    0.03277    1.56     0.12
## Year2011           0.03735    0.03289    1.14     0.26
## Year2012          -0.01015    0.03296   -0.31     0.76
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.42
## Multiple R-squared:  0.00717,    Adjusted R-squared:  0.00486
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
##  650 weights are ~= 1. The remaining 6676 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.104   0.869   0.950   0.901   0.985   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.37e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Sample size for the above analysis:  7326"
## [1] ""
## [1] ""
## [1] "####################################"
## [1] "Analysis of AJSC 2503"
## [1] "####################################"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 1291 1249 1286 1153 1200 1056  952 1155 1092 1124 1083 1036 1131 1269 1444
## 2011 2012
## 1496 1338
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  534  550  522  499  456  363  443  532  487  514  495  487  547  619  732
## 2011 2012
```

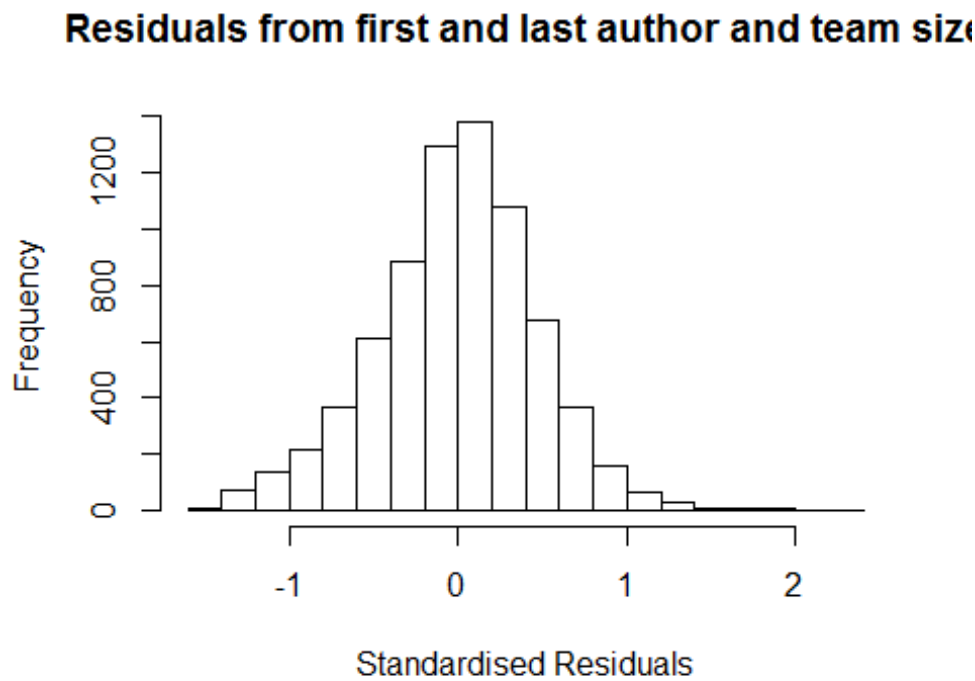
```
##  774  697
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  421  460  419  410  370  279  358  433  374  389  391  387  432  480  586
## 2011 2012
##  634  539
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
##  Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 160, df = 16, p-value <2e-16
```




```
##
##  Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 28, df = 1, p-value = 1e-07
```

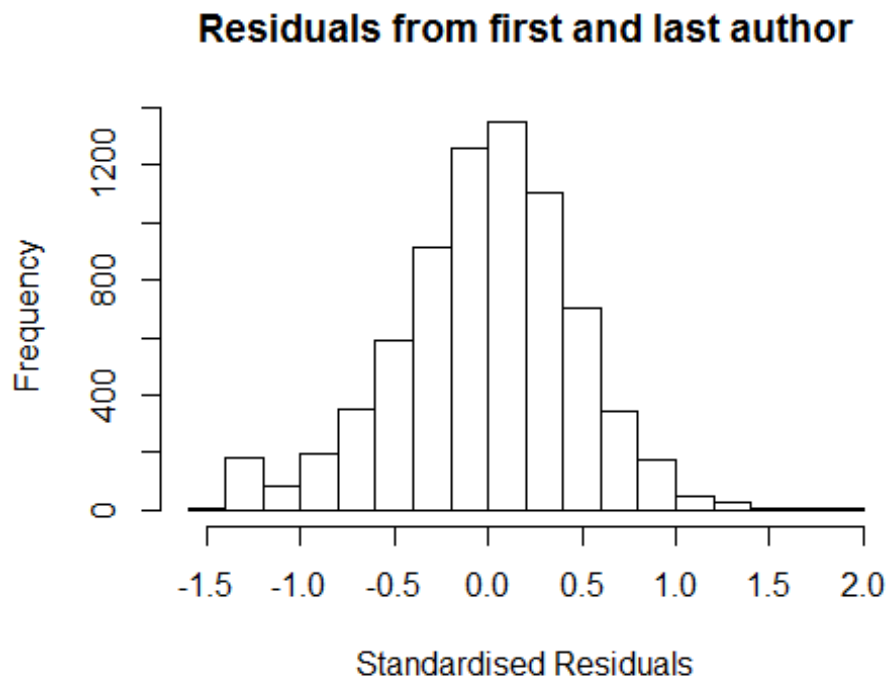



```
## [1] "Female first author team size 2018 geometric mean: 4.50253095844251"
## [1] "Male first author team size 2018 geometric mean: 3.85098014233113"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 29000, p-value = 0.01
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 4.69356841412715"
## [1] "Male last author team size 2018 geometric mean: 3.86538662888029"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 23000, p-value = 0.006
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.032  1           1.016
## LastAuthorFemale  1.031  1           1.015
## UniqueAuthors     1.136  4           1.016
## Year              1.139 16           1.004
```



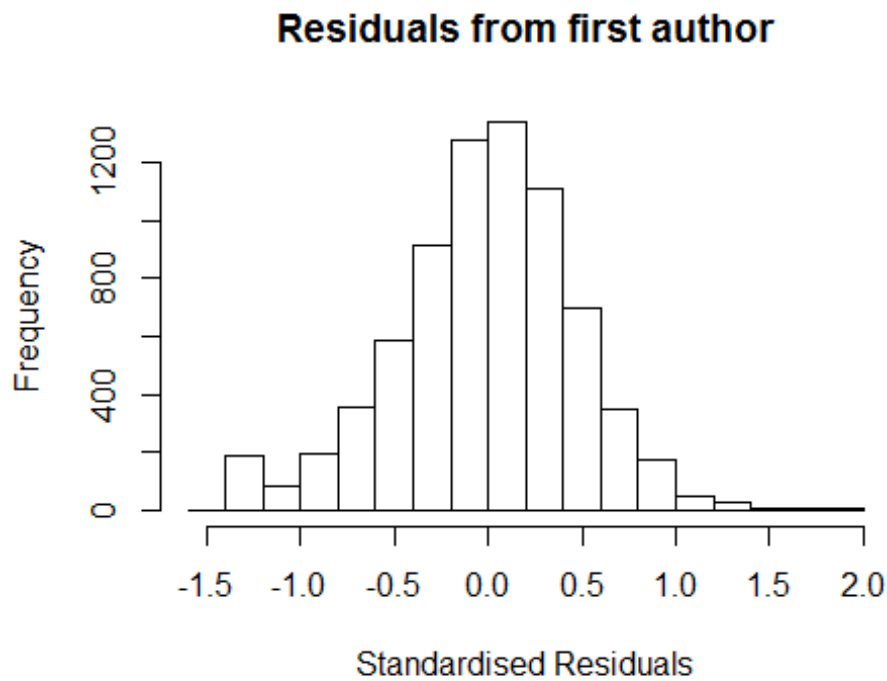

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
UniqueAuthors +
##     Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##     k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.4598 -0.2884  0.0134  0.2914  2.2555
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.043132   0.040041   26.05  < 2e-16 ***
## FirstAuthorFemale1  0.055508   0.012963    4.28  1.9e-05 ***
## LastAuthorFemale1  -0.000205   0.015783   -0.01   0.9896
## UniqueAuthors2      0.212404   0.033650    6.31  2.9e-10 ***
## UniqueAuthors3      0.267322   0.033597    7.96  2.0e-15 ***
## UniqueAuthors4      0.308213   0.033939    9.08  < 2e-16 ***
## UniqueAuthors5      0.351100   0.033921   10.35  < 2e-16 ***
## Year1997           -0.013502   0.036756   -0.37   0.7134
## Year1998            0.008641   0.037457    0.23   0.8176
## Year1999           -0.068473   0.037917   -1.81   0.0710 .
```

```
## Year2000           -0.118566   0.038680   -3.07   0.0022 **
## Year2001            0.064113   0.041910    1.53   0.1261
## Year2002            0.004557   0.035769    0.13   0.8986
## Year2003            0.034025   0.034418    0.99   0.3229
## Year2004            0.007273   0.035112    0.21   0.8359
## Year2005           -0.027657   0.034274   -0.81   0.4197
## Year2006            0.010080   0.035357    0.29   0.7756
## Year2007            0.023318   0.037198    0.63   0.5308
## Year2008            0.008163   0.033708    0.24   0.8086
## Year2009           -0.033851   0.033317   -1.02   0.3096
## Year2010           -0.005817   0.030801   -0.19   0.8502
## Year2011           -0.073492   0.031561   -2.33   0.0199 *
## Year2012           -0.108655   0.033501   -3.24   0.0012 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.436
## Multiple R-squared:  0.0531, Adjusted R-squared:  0.0503
## Convergence in 17 IRWLS iterations
##
## Robustness weights:
##  observation 3546 is an outlier with |weight| = 0 ( < 1.4e-05);
##  623 weights are ~= 1. The remaining 6738 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0123  0.8600  0.9520  0.8930  0.9860  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.36e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.014  1           1.007
## LastAuthorFemale  1.022  1           1.011
## Year              1.035 16           1.001
```



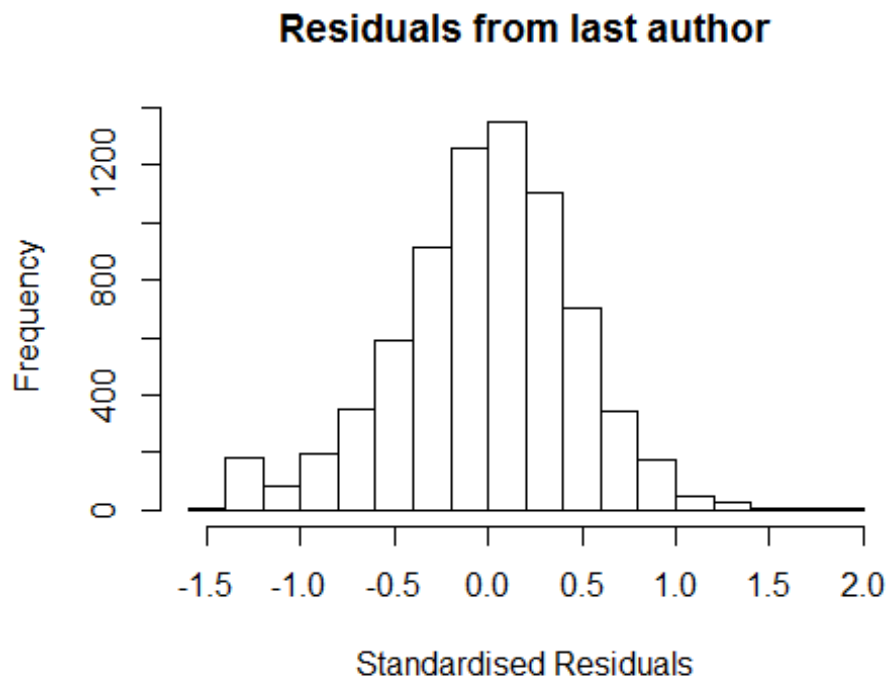

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.4200 -0.2936  0.0123  0.2939  1.9842
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)          1.27327    0.02701   47.14  < 2e-16 ***
## FirstAuthorFemale1   0.07612    0.01303    5.84  5.3e-09 ***
## LastAuthorFemale1    0.01331    0.01591    0.84   0.4029
## Year1997            -0.01103    0.03758   -0.29   0.7692
## Year1998             0.00264    0.03830    0.07   0.9450
## Year1999            -0.07096    0.03825   -1.86   0.0636 .
## Year2000            -0.11216    0.03943   -2.84   0.0045 **
## Year2001             0.08193    0.04254    1.93   0.0541 .
## Year2002             0.02636    0.03627    0.73   0.4673
## Year2003             0.05728    0.03443    1.66   0.0962 .
## Year2004             0.02673    0.03550    0.75   0.4516
## Year2005             0.01355    0.03469    0.39   0.6961
```

```
## Year2006            0.04353    0.03568    1.22   0.2226
## Year2007            0.04316    0.03689    1.17   0.2421
## Year2008            0.04782    0.03395    1.41   0.1591
## Year2009            0.01416    0.03355    0.42   0.6730
## Year2010            0.03789    0.03109    1.22   0.2229
## Year2011           -0.02187    0.03162   -0.69   0.4892
## Year2012           -0.05357    0.03396   -1.58   0.1147
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.441
## Multiple R-squared:  0.0146, Adjusted R-squared:  0.0121
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
##  596 weights are ~= 1. The remaining 6766 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0058  0.8650  0.9510  0.8920  0.9860  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.36e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.014  1           1.007
## Year              1.014 16           1.000
```




```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.4091 -0.2940  0.0119  0.2928  1.9828
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)          1.27418    0.02701   47.17  < 2e-16 ***
## FirstAuthorFemale1   0.07711    0.01307    5.90  3.8e-09 ***
## Year1997            -0.01062    0.03757   -0.28   0.7776
## Year1998             0.00295    0.03828    0.08   0.9387
## Year1999            -0.07103    0.03828   -1.86   0.0635 .
## Year2000            -0.11185    0.03945   -2.84   0.0046 **
## Year2001             0.08297    0.04253    1.95   0.0511 .
## Year2002             0.02679    0.03628    0.74   0.4603
## Year2003             0.05780    0.03445    1.68   0.0934 .
## Year2004             0.02715    0.03553    0.76   0.4449
## Year2005             0.01402    0.03470    0.40   0.6861
## Year2006             0.04407    0.03569    1.23   0.2169
```

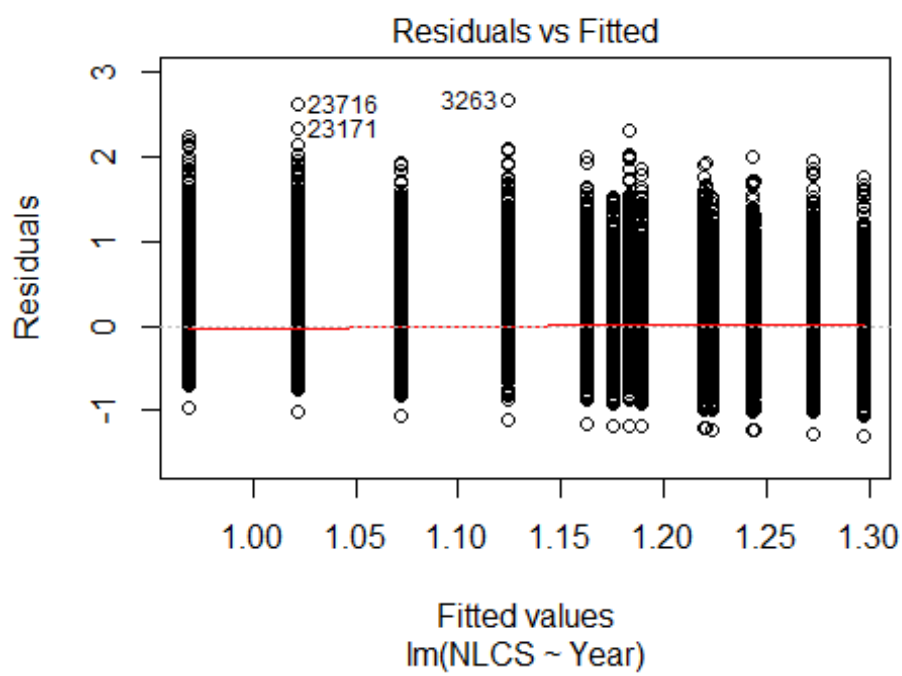
```
## Year2007            0.04398    0.03685    1.19   0.2327
## Year2008            0.04802    0.03396    1.41   0.1574
## Year2009            0.01498    0.03354    0.45   0.6553
## Year2010            0.03883    0.03104    1.25   0.2110
## Year2011           -0.02038    0.03154   -0.65   0.5180
## Year2012           -0.05234    0.03390   -1.54   0.1227
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.441
## Multiple R-squared:  0.0144, Adjusted R-squared:  0.0122
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
##  599 weights are ~= 1. The remaining 6763 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.006   0.864   0.950   0.892   0.986   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.36e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##                   GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.023  1           1.011
## Year             1.023 16           1.001
```



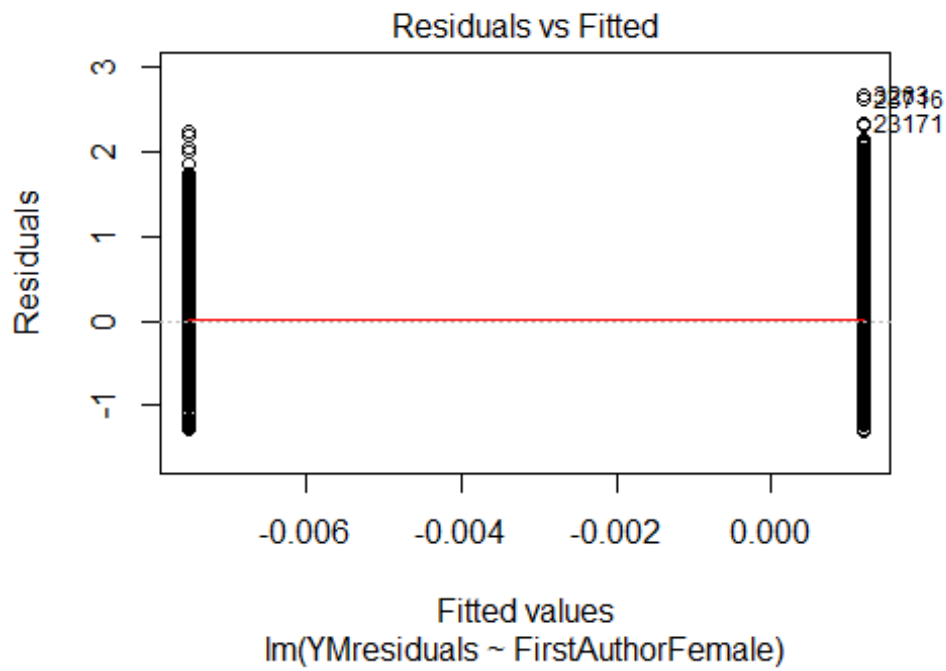

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.3895 -0.2987  0.0139  0.2956  1.9660
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.284456   0.027015   47.55   <2e-16 ***
## LastAuthorFemale1   0.021474   0.016069    1.34   0.1815
## Year1997           -0.013137   0.037759   -0.35   0.7279
## Year1998           -0.000463   0.038384   -0.01   0.9904
## Year1999           -0.071542   0.038285   -1.87   0.0617 .
## Year2000           -0.109900   0.039617   -2.77   0.0056 **
## Year2001            0.083571   0.042589    1.96   0.0498 *
## Year2002            0.026937   0.036309    0.74   0.4582
## Year2003            0.059614   0.034619    1.72   0.0851 .
## Year2004            0.031730   0.035531    0.89   0.3719
## Year2005            0.020537   0.034893    0.59   0.5562
## Year2006            0.046721   0.035879    1.30   0.1929
```

```
## Year2007           0.047511   0.036969    1.29   0.1988
## Year2008           0.052585   0.034084    1.54   0.1229
## Year2009           0.020165   0.033673    0.60   0.5493
## Year2010           0.044107   0.031206    1.41   0.1576
## Year2011          -0.014739   0.031749   -0.46   0.6425
## Year2012          -0.045981   0.034010   -1.35   0.1764
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.442
## Multiple R-squared:  0.0102, Adjusted R-squared:  0.00796
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
##  618 weights are ~= 1. The remaining 6744 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0094  0.8630  0.9510  0.8920  0.9860  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.36e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Sample size for the above analysis:  7362"
## [1] ""
## [1] ""
## [1] "####################################"
## [1] "Analysis of AJSC 2504"
## [1] "####################################"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 3742 2975 3242 2864 4733 5869 5408 3446 2911 3318 3605 4117 4251 4485 4706
## 2011 2012
## 4945 4833
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 1188 1009 1065  586 1723 2204 2252 1248 1079 1359 1420 1786 1812 2002 2033
## 2011 2012
```

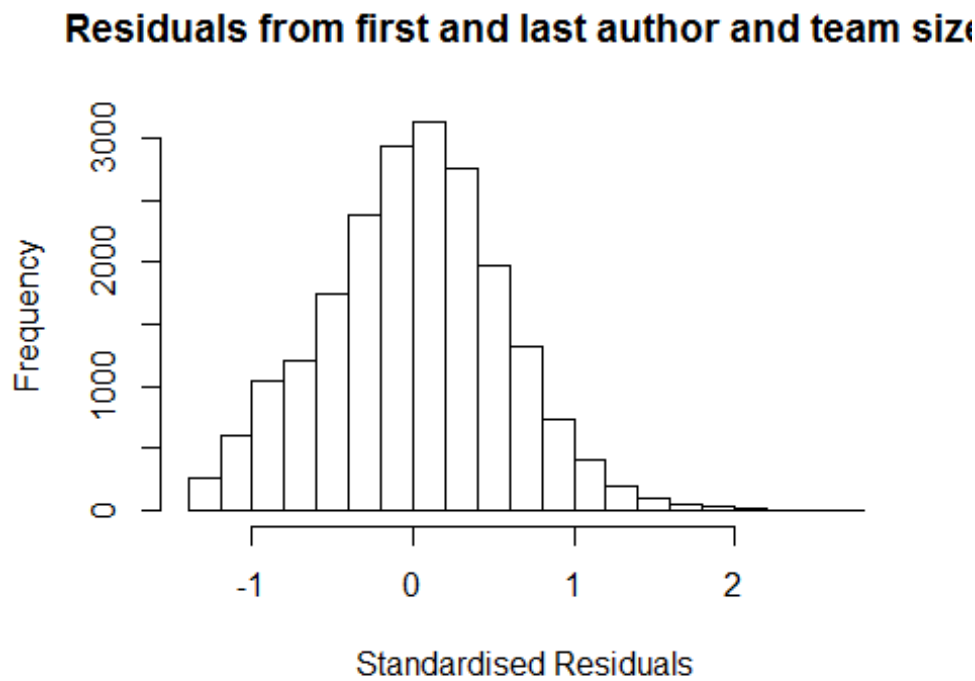
```
## 2165 2281
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  947  777  835  444 1372 1698 1743  915  831 1012 1055 1365 1391 1521 1590
## 2011 2012
## 1655 1724
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
##  Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 1000, df = 16, p-value <2e-16
```




```
##
##  Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 53, df = 1, p-value = 4e-13
```

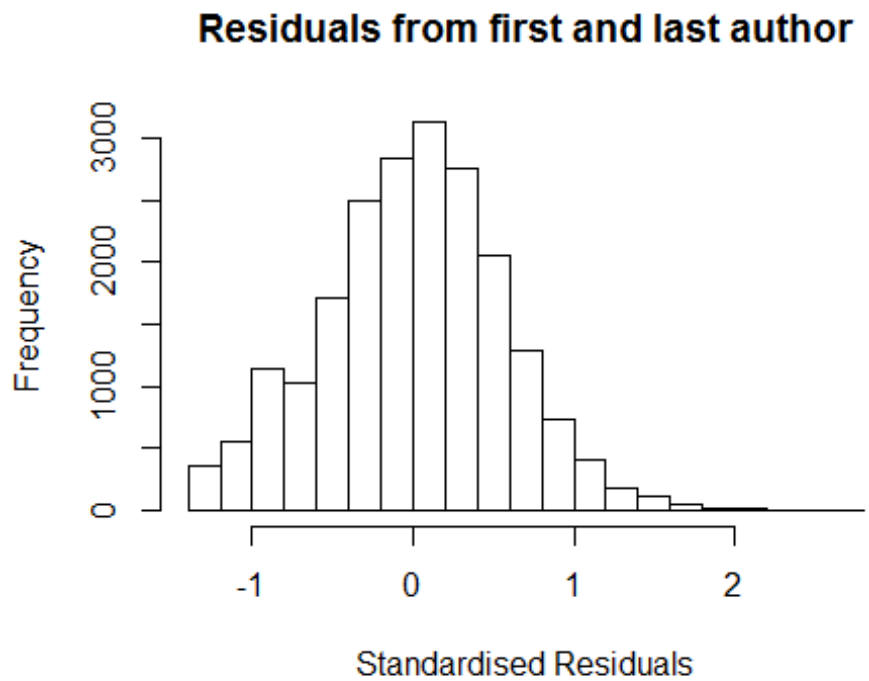



```
## [1] "Female first author team size 2018 geometric mean: 4.24726218919724"
## [1] "Male first author team size 2018 geometric mean: 3.60034434243734"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 110000, p-value = 7e-04
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 4.36528055743168"
## [1] "Male last author team size 2018 geometric mean: 3.59486344924415"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 99000, p-value = 1e-04
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.032  1           1.016
## LastAuthorFemale  1.023  1           1.012
## UniqueAuthors     1.062  4           1.008
## Year              1.062 16           1.002
```



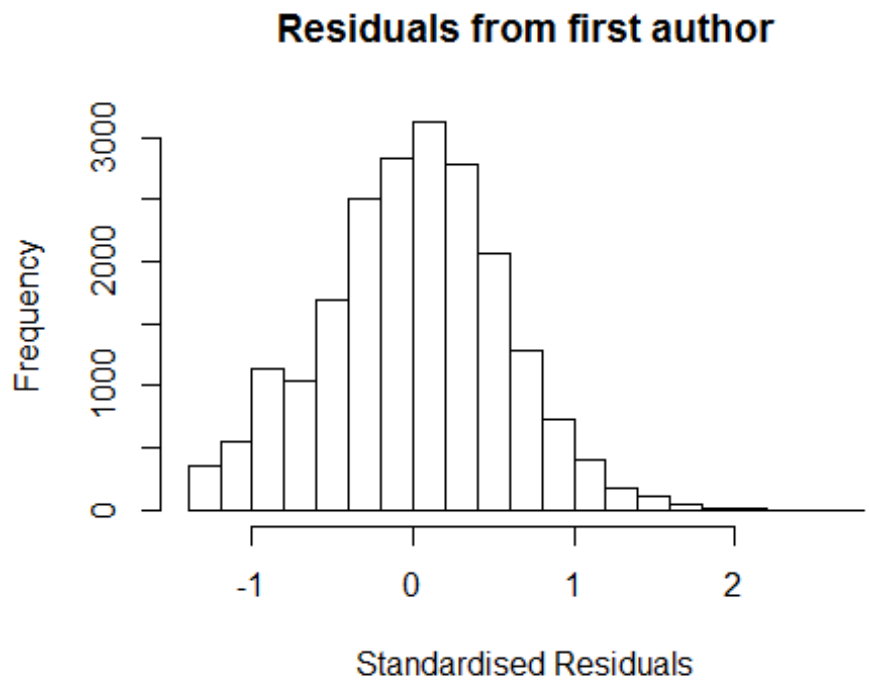

```
## [1] "List of  3 outliers with residuals above  2.5"
##         ScopusId  NLCS Year OneField Fields residuals
## 3263  2842565972 3.795 1996     2504      2     2.708
## 4658  0031249402 3.505 1997     2208      2     2.502
## 23716 0035278804 3.653 2001     2208      2     2.658
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
UniqueAuthors +
##     Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##     k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.3609 -0.3626  0.0161  0.3678  2.7076
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)          0.9361     0.0263   35.62  < 2e-16 ***
## FirstAuthorFemale1  -0.0142     0.0109   -1.30  0.19233
## LastAuthorFemale1    0.0133     0.0125    1.06  0.28752
## UniqueAuthors2       0.1514     0.0156    9.73  < 2e-16 ***
## UniqueAuthors3       0.2017     0.0159   12.71  < 2e-16 ***
## UniqueAuthors4       0.2122     0.0164   12.96  < 2e-16 ***
## UniqueAuthors5       0.2444     0.0156   15.68  < 2e-16 ***
## Year1997             0.0671     0.0346    1.94  0.05210 .
```

```
## Year1998             0.0769     0.0326    2.36  0.01835 *
## Year1999             0.1451     0.0349    4.15  3.3e-05 ***
## Year2000            -0.0444     0.0307   -1.45  0.14814
## Year2001            -0.1533     0.0304   -5.04  4.7e-07 ***
## Year2002            -0.1956     0.0297   -6.58  5.0e-11 ***
## Year2003             0.1212     0.0312    3.89  0.00010 ***
## Year2004             0.1016     0.0299    3.40  0.00067 ***
## Year2005             0.1671     0.0283    5.90  3.7e-09 ***
## Year2006             0.1004     0.0291    3.46  0.00055 ***
## Year2007             0.1626     0.0275    5.91  3.4e-09 ***
## Year2008             0.1281     0.0275    4.66  3.3e-06 ***
## Year2009             0.1003     0.0268    3.74  0.00018 ***
## Year2010             0.0872     0.0266    3.28  0.00105 **
## Year2011             0.0465     0.0266    1.74  0.08127 .
## Year2012             0.0555     0.0268    2.07  0.03857 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.545
## Multiple R-squared:  0.0526, Adjusted R-squared:  0.0516
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
##  2 observations c(747,5156) are outliers with |weight| = 0 ( < 4.8e-06);
##  1831 weights are ~= 1. The remaining 19042 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0016  0.8640  0.9500  0.9020  0.9860  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          4.79e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.022  1           1.011
## LastAuthorFemale  1.018  1           1.009
## Year              1.020 16           1.001
```



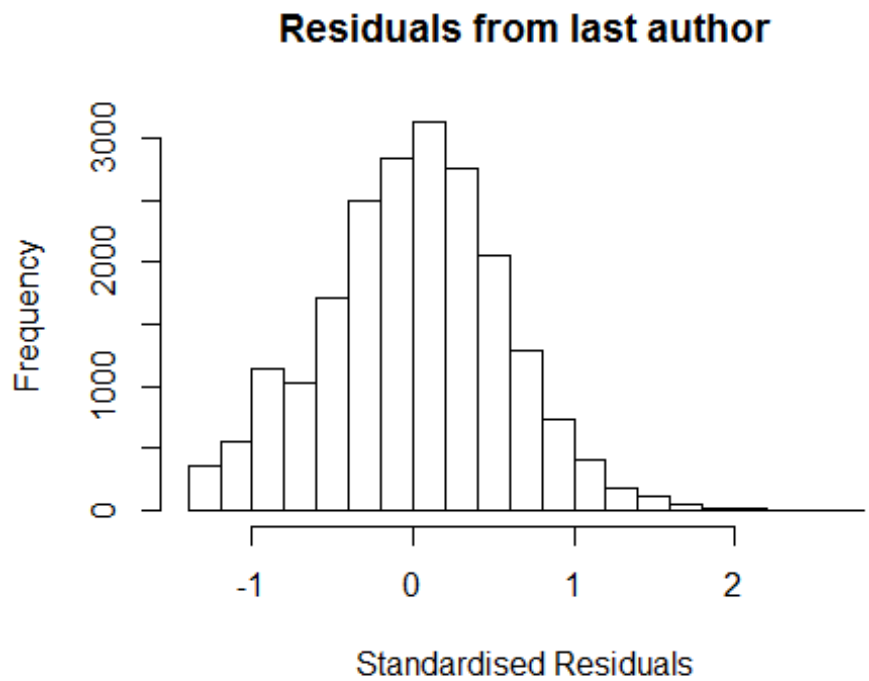

```
## [1] "List of  2 outliers with residuals above  2.5"
##         ScopusId  NLCS Year OneField Fields residuals
## 3263  2842565972 3.795 1996     2504      2     2.720
## 23716 0035278804 3.653 2001     2208      2     2.695
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.3067 -0.3737  0.0217  0.3712  2.7202
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)          1.07484    0.02394   44.89  < 2e-16 ***
## FirstAuthorFemale1   0.00382    0.01087    0.35  0.72506
## LastAuthorFemale1    0.02198    0.01261    1.74  0.08131 .
## Year1997             0.07359    0.03502    2.10  0.03564 *
## Year1998             0.08796    0.03313    2.65  0.00795 **
## Year1999             0.16105    0.03491    4.61  4.0e-06 ***
## Year2000            -0.02351    0.03092   -0.76  0.44707
## Year2001            -0.11660    0.03036   -3.84  0.00012 ***
## Year2002            -0.15598    0.02983   -5.23  1.7e-07 ***
## Year2003             0.14284    0.03151    4.53  5.9e-06 ***
## Year2004             0.13089    0.03023    4.33  1.5e-05 ***
```

```
## Year2005            0.20610    0.02856    7.22  5.5e-13 ***
## Year2006            0.13445    0.02933    4.58  4.6e-06 ***
## Year2007            0.20248    0.02765    7.32  2.5e-13 ***
## Year2008            0.16612    0.02770    6.00  2.1e-09 ***
## Year2009            0.13698    0.02700    5.07  3.9e-07 ***
## Year2010            0.13175    0.02671    4.93  8.2e-07 ***
## Year2011            0.09156    0.02672    3.43  0.00061 ***
## Year2012            0.10524    0.02689    3.91  9.1e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.549
## Multiple R-squared:  0.0363, Adjusted R-squared:  0.0354
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
##  2 observations c(747,5156) are outliers with |weight| = 0 ( < 4.8e-06);
##  1760 weights are ~= 1. The remaining 19113 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0163  0.8630  0.9510  0.9020  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          4.79e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.012  1           1.006
## Year              1.012 16           1.000
```




```
## [1] "List of  2 outliers with residuals above  2.5"
##         ScopusId  NLCS Year OneField Fields residuals
## 3263  2842565972 3.795 1996     2504      2     2.720
## 23716 0035278804 3.653 2001     2208      2     2.695
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.2892 -0.3746  0.0214  0.3702  2.7191
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)          1.07595    0.02395   44.93  < 2e-16 ***
## FirstAuthorFemale1   0.00636    0.01083    0.59  0.55723
## Year1997             0.07390    0.03503    2.11  0.03489 *
## Year1998             0.08782    0.03314    2.65  0.00805 **
## Year1999             0.16114    0.03492    4.61  4.0e-06 ***
## Year2000            -0.02276    0.03092   -0.74  0.46169
## Year2001            -0.11566    0.03037   -3.81  0.00014 ***
## Year2002            -0.15539    0.02984   -5.21  1.9e-07 ***
## Year2003             0.14336    0.03152    4.55  5.4e-06 ***
## Year2004             0.13084    0.03024    4.33  1.5e-05 ***
## Year2005             0.20685    0.02855    7.24  4.5e-13 ***
```

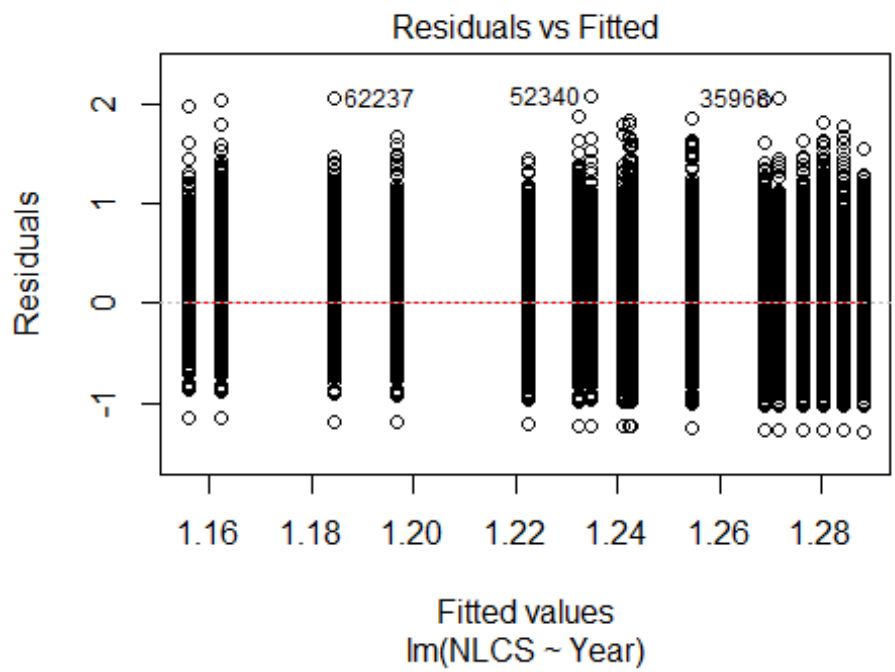
```
## Year2006            0.13500    0.02934    4.60  4.2e-06 ***
## Year2007            0.20332    0.02765    7.35  2.0e-13 ***
## Year2008            0.16732    0.02770    6.04  1.6e-09 ***
## Year2009            0.13787    0.02700    5.11  3.3e-07 ***
## Year2010            0.13283    0.02671    4.97  6.7e-07 ***
## Year2011            0.09257    0.02672    3.46  0.00053 ***
## Year2012            0.10631    0.02689    3.95  7.7e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.549
## Multiple R-squared:  0.0361, Adjusted R-squared:  0.0353
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
##  2 observations c(747,5156) are outliers with |weight| = 0 ( < 4.8e-06);
##  1778 weights are ~= 1. The remaining 19095 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0166  0.8630  0.9510  0.9020  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          4.79e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##                   GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.009  1           1.004
## Year             1.009 16           1.000
```



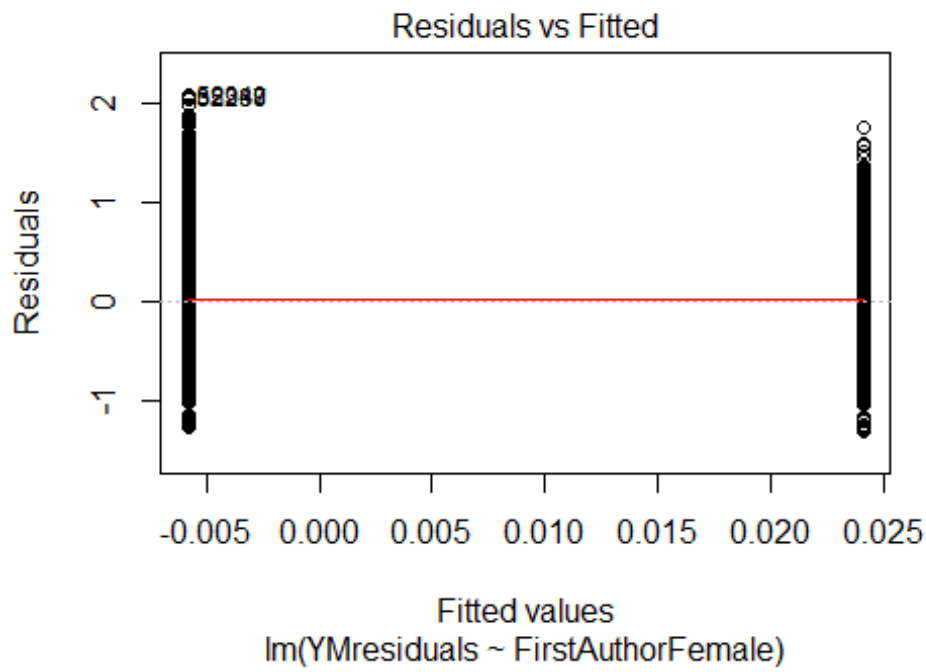

```
## [1] "List of  2 outliers with residuals above  2.5"
##         ScopusId  NLCS Year OneField Fields residuals
## 3263  2842565972 3.795 1996     2504      2     2.720
## 23716 0035278804 3.653 2001     2208      2     2.695
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##    Min     1Q Median     3Q    Max
## -1.304 -0.373  0.022  0.371  2.720
##
## Coefficients:
##                   Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.0751     0.0239   44.92  < 2e-16 ***
## LastAuthorFemale1   0.0226     0.0126    1.80  0.07229 .
## Year1997            0.0737     0.0350    2.10  0.03543 *
## Year1998            0.0880     0.0331    2.66  0.00791 **
## Year1999            0.1613     0.0349    4.62  3.8e-06 ***
## Year2000           -0.0234     0.0309   -0.76  0.44907
## Year2001           -0.1164     0.0304   -3.84  0.00013 ***
## Year2002           -0.1558     0.0298   -5.23  1.8e-07 ***
## Year2003            0.1430     0.0315    4.54  5.7e-06 ***
## Year2004            0.1310     0.0302    4.34  1.5e-05 ***
## Year2005            0.2064     0.0285    7.23  5.0e-13 ***
```

```
## Year2006            0.1346     0.0293    4.59  4.4e-06 ***
## Year2007            0.2028     0.0276    7.34  2.2e-13 ***
## Year2008            0.1665     0.0277    6.02  1.8e-09 ***
## Year2009            0.1373     0.0270    5.09  3.6e-07 ***
## Year2010            0.1321     0.0267    4.95  7.4e-07 ***
## Year2011            0.0919     0.0267    3.44  0.00058 ***
## Year2012            0.1056     0.0269    3.93  8.4e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.549
## Multiple R-squared:  0.0363, Adjusted R-squared:  0.0355
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
##  2 observations c(747,5156) are outliers with |weight| = 0 ( < 4.8e-06);
##  1777 weights are ~= 1. The remaining 19096 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0164  0.8630  0.9510  0.9020  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          4.79e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                  psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Sample size for the above analysis:  20875"
## [1] ""
## [1] ""
## [1] "##################################"
## [1] "Analysis of AJSC 2505"
## [1] "##################################"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 4464 5239 4591 4287 4260 3940 3578 3532 3641 3857 4022 3551 3710 3749 3983
## 2011 2012
## 3915 3827
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
```

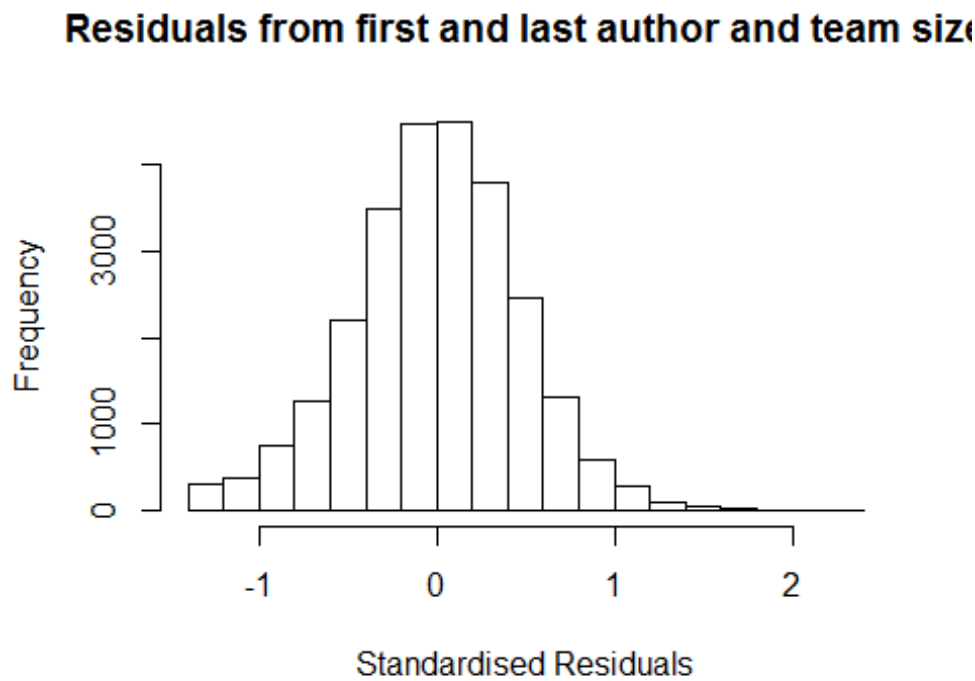
```
## 1928 2243 2024 1827 1792 1487 1707 1724 1741 1888 1963 1800 1871 1960 2011
## 2011 2012
## 2028 2009
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 1583 1825 1662 1548 1481 1218 1402 1385 1420 1515 1565 1452 1504 1578 1642
## 2011 2012
## 1595 1596
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
##  Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 140, df = 16, p-value <2e-16
```




```
##
##  Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 27, df = 1, p-value = 3e-07
```

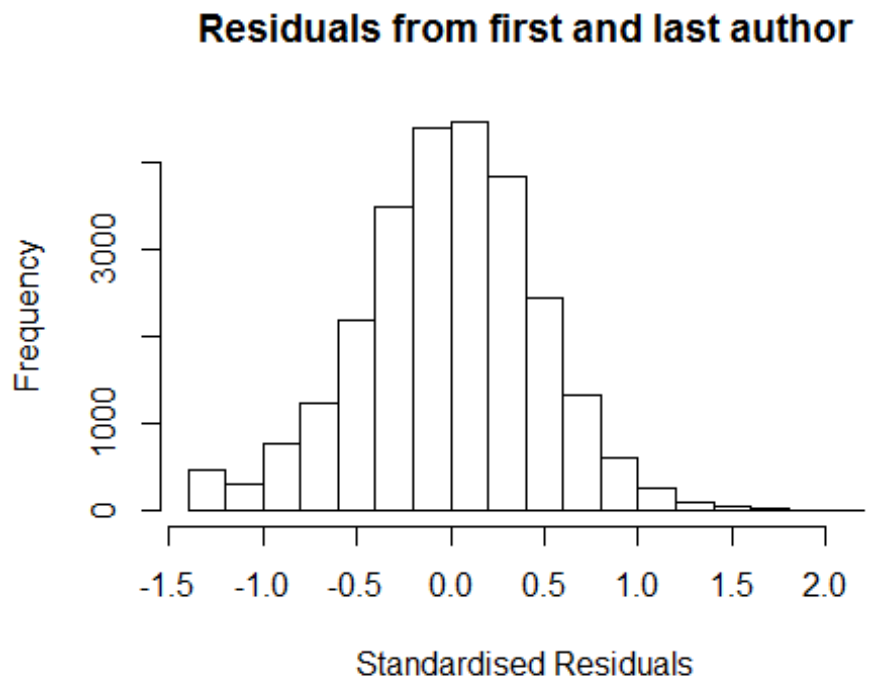



```
## [1] "Female first author team size 2018 geometric mean: 3.99644854782466"
## [1] "Male first author team size 2018 geometric mean: 3.6104663313621"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 52000, p-value = 0.2
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 3.90197775024769"
## [1] "Male last author team size 2018 geometric mean: 3.66146722288053"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 41000, p-value = 0.3
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.019  1           1.009
## LastAuthorFemale  1.012  1           1.006
## UniqueAuthors     1.072  4           1.009
## Year              1.070 16           1.002
```



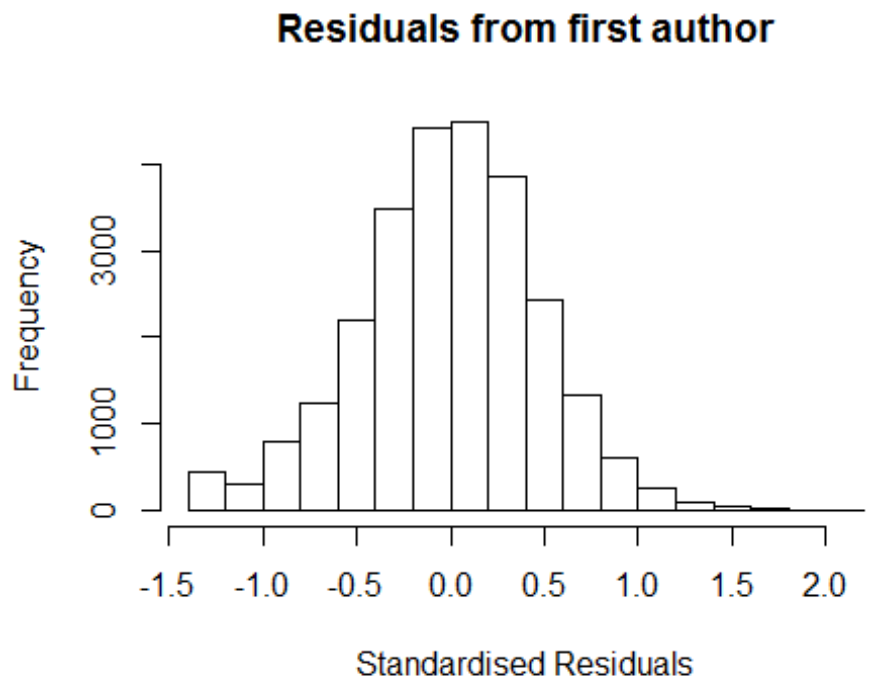

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
UniqueAuthors +
##     Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##     k.max = 1000))
##  \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.38705 -0.30049  0.00555  0.30323  2.28764
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.05959    0.01834   57.76  < 2e-16 ***
## FirstAuthorFemale1  0.01398    0.00714    1.96   0.0504 .
## LastAuthorFemale1   0.00989    0.00886    1.12   0.2642
## UniqueAuthors2      0.18664    0.01473   12.67  < 2e-16 ***
## UniqueAuthors3      0.20894    0.01485   14.07  < 2e-16 ***
## UniqueAuthors4      0.25012    0.01526   16.39  < 2e-16 ***
## UniqueAuthors5      0.29258    0.01509   19.39  < 2e-16 ***
## Year1997            0.01067    0.01793    0.60   0.5517
## Year1998            0.03489    0.01849    1.89   0.0592 .
## Year1999           -0.01787    0.01858   -0.96   0.3363
```

```
## Year2000           -0.03077    0.01854   -1.66   0.0970 .
## Year2001            0.01430    0.01943    0.74   0.4617
## Year2002            0.00132    0.01854    0.07   0.9432
## Year2003            0.01230    0.01793    0.69   0.4927
## Year2004           -0.01541    0.01778   -0.87   0.3859
## Year2005            0.01135    0.01754    0.65   0.5175
## Year2006           -0.03931    0.01787   -2.20   0.0278 *
## Year2007           -0.03228    0.01826   -1.77   0.0771 .
## Year2008           -0.05070    0.01802   -2.81   0.0049 **
## Year2009           -0.08859    0.01806   -4.91  9.3e-07 ***
## Year2010           -0.08814    0.01729   -5.10  3.5e-07 ***
## Year2011           -0.13122    0.01777   -7.38  1.6e-13 ***
## Year2012           -0.13328    0.01795   -7.42  1.2e-13 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.449
## Multiple R-squared:  0.0341, Adjusted R-squared:  0.0333
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
##  3 observations c(12620,24105,25341)
##   are outliers with |weight| = 0 ( < 3.9e-06);
##  2232 weights are ~= 1. The remaining 23736 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0005  0.8670  0.9500  0.8980  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          3.85e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.010  1           1.005
## LastAuthorFemale  1.010  1           1.005
## Year              1.013 16           1.000
```



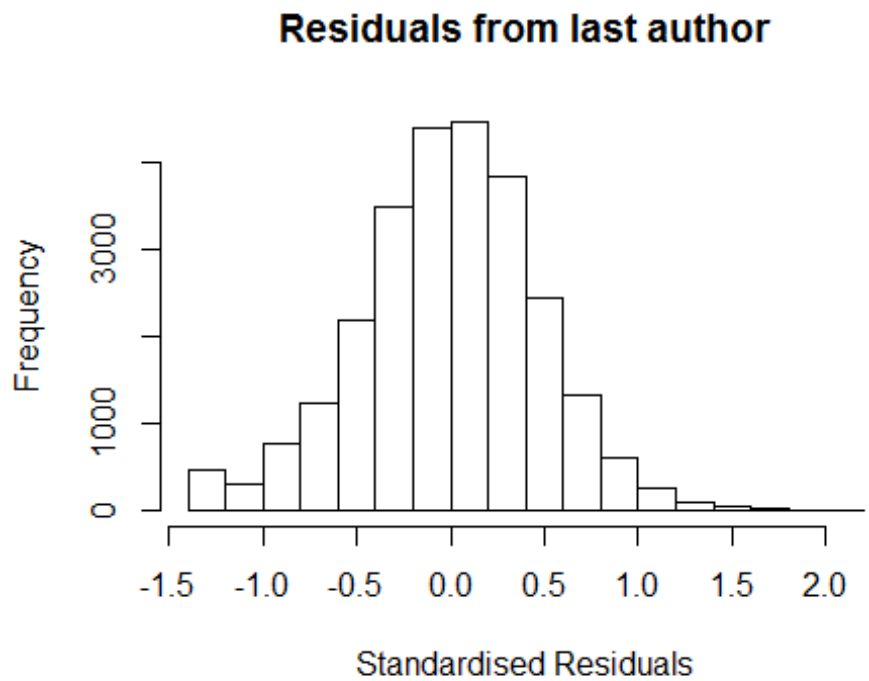

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.33913 -0.30036  0.00678  0.30398  2.07285
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)          1.24649    0.01351   92.26  < 2e-16 ***
## FirstAuthorFemale1   0.03040    0.00716    4.25  2.2e-05 ***
## LastAuthorFemale1    0.01725    0.00890    1.94  0.05260 .
## Year1997             0.00863    0.01802    0.48  0.63194
## Year1998             0.04499    0.01850    2.43  0.01502 *
## Year1999            -0.00864    0.01860   -0.46  0.64210
## Year2000            -0.01740    0.01859   -0.94  0.34925
## Year2001             0.02851    0.01946    1.47  0.14293
## Year2002             0.02327    0.01858    1.25  0.21054
## Year2003             0.03451    0.01795    1.92  0.05456 .
## Year2004             0.01173    0.01784    0.66  0.51082
## Year2005             0.04041    0.01755    2.30  0.02132 *
```

```
## Year2006           -0.00856    0.01796   -0.48  0.63384
## Year2007           -0.00467    0.01822   -0.26  0.79762
## Year2008           -0.01844    0.01803   -1.02  0.30627
## Year2009           -0.06134    0.01811   -3.39  0.00071 ***
## Year2010           -0.05047    0.01721   -2.93  0.00337 **
## Year2011           -0.09081    0.01770   -5.13  2.9e-07 ***
## Year2012           -0.09196    0.01800   -5.11  3.3e-07 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.451
## Multiple R-squared:  0.00853,    Adjusted R-squared:  0.00785
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  2212 weights are ~= 1. The remaining 23759 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0015  0.8670  0.9500  0.8970  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          3.85e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.007  1           1.003
## Year              1.007 16           1.000
```




```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.32470 -0.30006  0.00677  0.30427  2.07094
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.24794    0.01349   92.52  < 2e-16 ***
## FirstAuthorFemale1  0.03167    0.00716    4.43  9.6e-06 ***
## Year1997            0.00859    0.01803    0.48  0.63354
## Year1998            0.04509    0.01850    2.44  0.01482 *
## Year1999           -0.00871    0.01861   -0.47  0.63976
## Year2000           -0.01719    0.01859   -0.92  0.35512
## Year2001            0.02874    0.01947    1.48  0.13984
## Year2002            0.02294    0.01859    1.23  0.21712
## Year2003            0.03492    0.01796    1.94  0.05184 .
## Year2004            0.01211    0.01785    0.68  0.49741
## Year2005            0.04083    0.01755    2.33  0.01999 *
## Year2006           -0.00832    0.01797   -0.46  0.64336
```

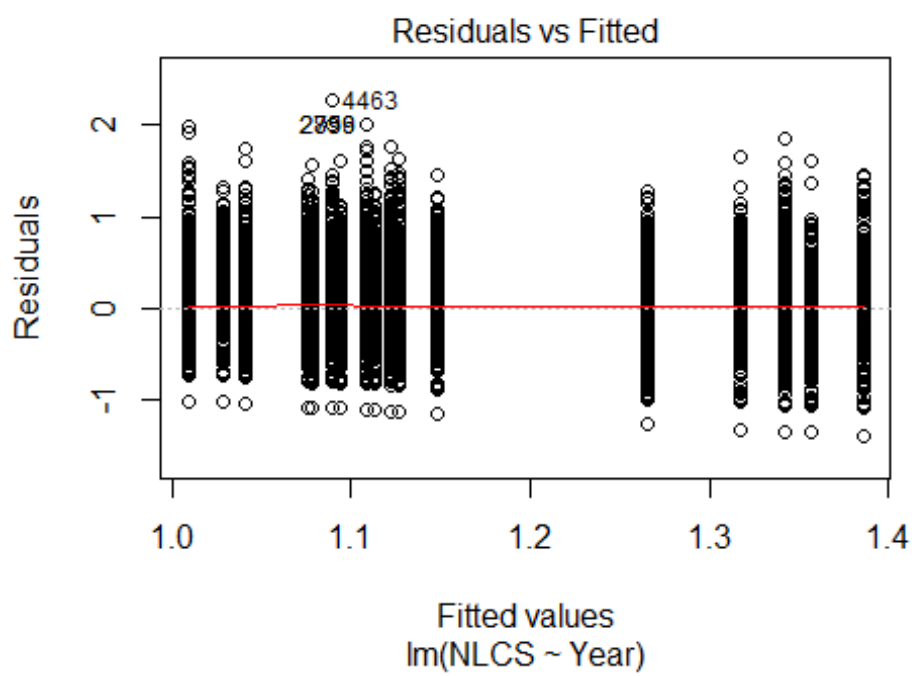
```
## Year2007           -0.00407    0.01821   -0.22  0.82296
## Year2008           -0.01759    0.01802   -0.98  0.32882
## Year2009           -0.06089    0.01811   -3.36  0.00078 ***
## Year2010           -0.04980    0.01722   -2.89  0.00383 **
## Year2011           -0.08993    0.01770   -5.08  3.8e-07 ***
## Year2012           -0.09095    0.01800   -5.05  4.4e-07 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.451
## Multiple R-squared:  0.00839,    Adjusted R-squared:  0.00774
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  2203 weights are ~= 1. The remaining 23768 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0016  0.8670  0.9510  0.8970  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          3.85e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##                   GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.007  1           1.004
## Year             1.007 16           1.000
```



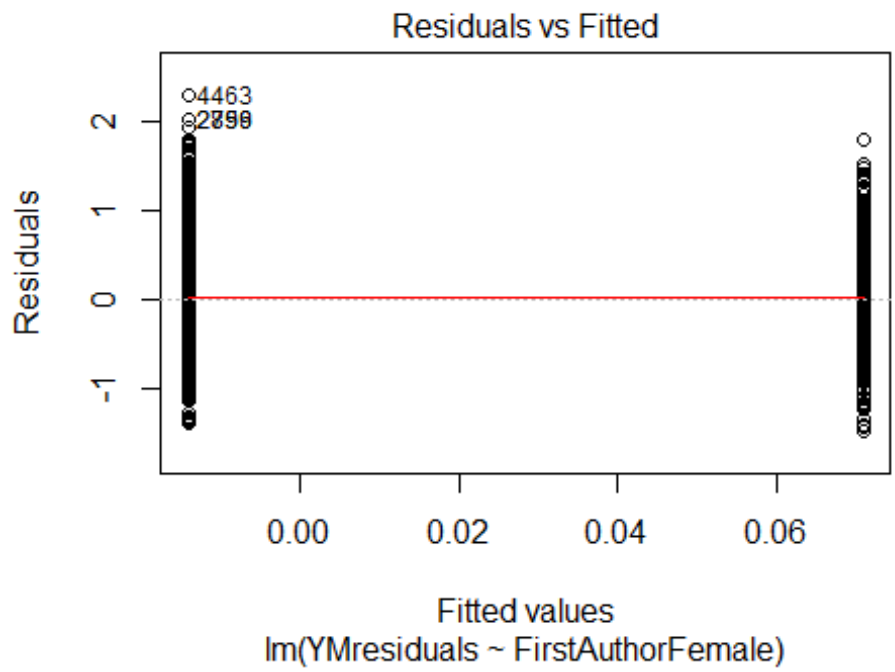

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.31702 -0.30130  0.00672  0.30530  2.06630
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.25172    0.01346   93.02  < 2e-16 ***
## LastAuthorFemale1  0.02070    0.00890    2.33  0.01999 *
## Year1997            0.00831    0.01803    0.46  0.64479
## Year1998            0.04460    0.01851    2.41  0.01598 *
## Year1999           -0.00887    0.01861   -0.48  0.63373
## Year2000           -0.01713    0.01860   -0.92  0.35689
## Year2001            0.02888    0.01948    1.48  0.13814
## Year2002            0.02340    0.01858    1.26  0.20782
## Year2003            0.03555    0.01796    1.98  0.04771 *
## Year2004            0.01212    0.01784    0.68  0.49691
## Year2005            0.04108    0.01756    2.34  0.01932 *
## Year2006           -0.00776    0.01797   -0.43  0.66589
```

```
## Year2007          -0.00387    0.01823   -0.21  0.83175
## Year2008          -0.01761    0.01804   -0.98  0.32901
## Year2009          -0.06003    0.01813   -3.31  0.00093 ***
## Year2010          -0.04882    0.01721   -2.84  0.00455 **
## Year2011          -0.08970    0.01772   -5.06  4.2e-07 ***
## Year2012          -0.09044    0.01801   -5.02  5.2e-07 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.451
## Multiple R-squared:  0.00784,    Adjusted R-squared:  0.00719
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  2196 weights are ~= 1. The remaining 23775 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.002   0.867   0.950   0.897   0.985   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          3.85e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Sample size for the above analysis:  25971"
## [1] ""
## [1] ""
## [1] "####################################"
## [1] "Analysis of AJSC 2506"
## [1] "####################################"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 1250 1425 1362 1161 1113 1049  940  993 1015  941 1056 1186 1069 1246 1351
## 2011 2012
## 1245 1183
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  444  521  482  382  368  240  325  378  384  350  430  508  450  600  645
## 2011 2012
```

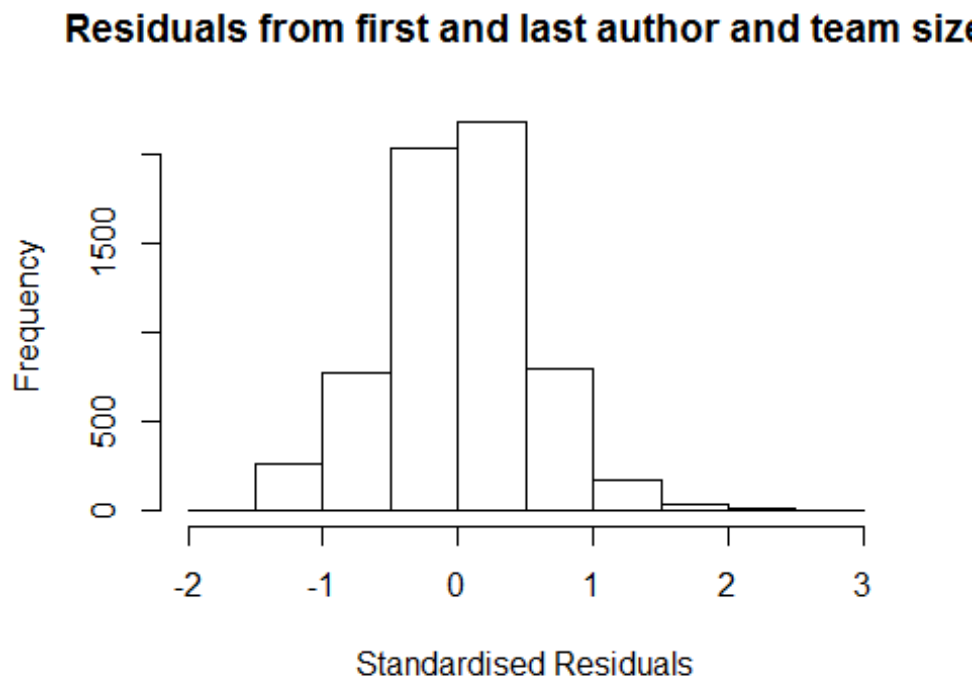
```
##  584  597
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  374  436  418  313  299  195  267  315  289  270  321  419  352  495  538
## 2011 2012
##  477  475
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
##  Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 110, df = 16, p-value = 1e-15
```




```
##
##  Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 25, df = 1, p-value = 5e-07
```

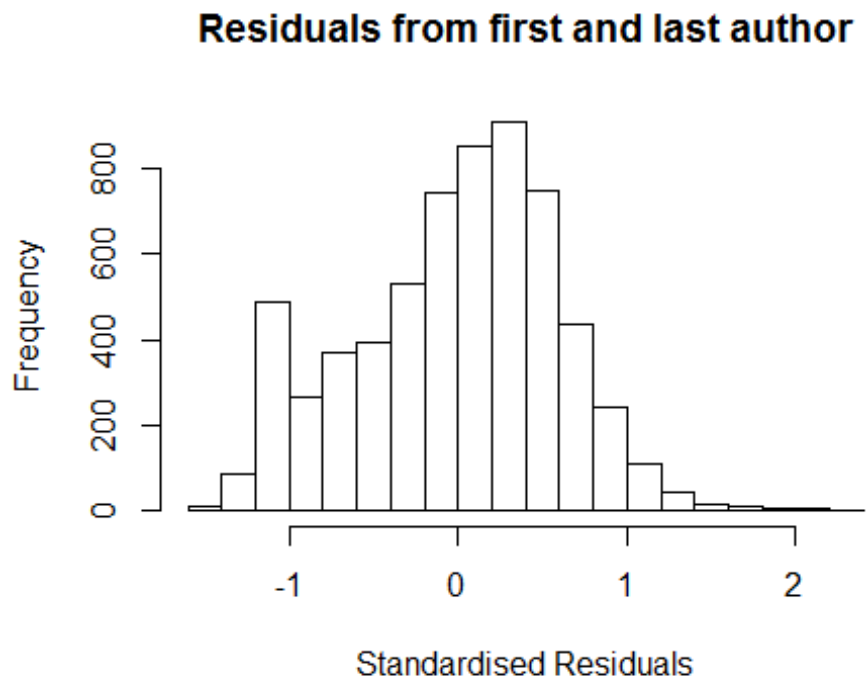



```
## [1] "Female first author team size 2018 geometric mean: 3.77665276903957"
## [1] "Male first author team size 2018 geometric mean: 3.12918137987467"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 26000, p-value = 0.006
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 3.80602577520531"
## [1] "Male last author team size 2018 geometric mean: 3.15004268797717"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 23000, p-value = 0.006
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.047  1           1.023
## LastAuthorFemale  1.047  1           1.023
## UniqueAuthors     1.187  4           1.022
## Year              1.211 16           1.006
```



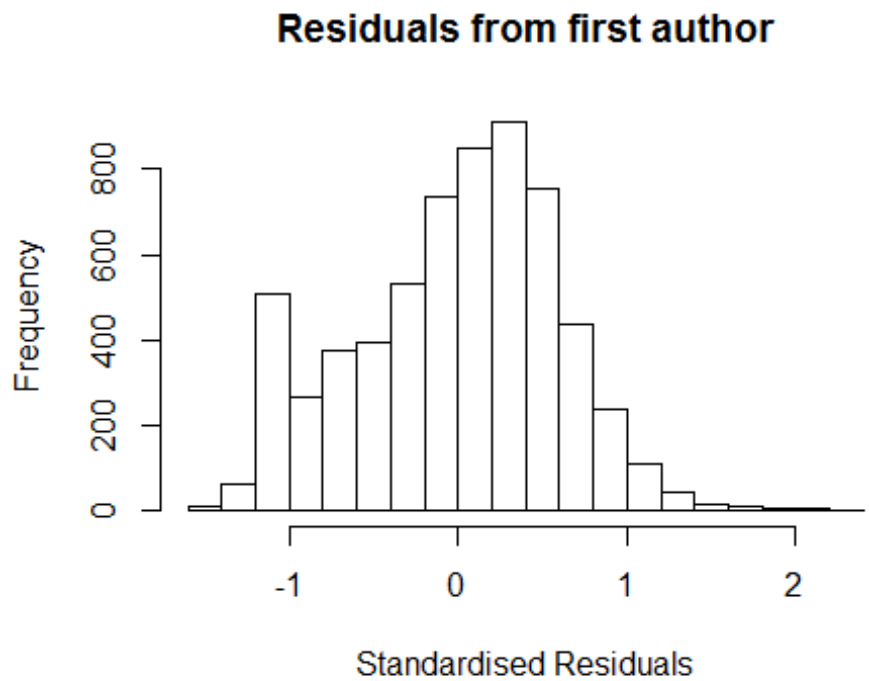

```
## [1] "List of  2 outliers with residuals above  2.5"
##        ScopusId  NLCS Year OneField Fields residuals
## 2796 0030671739 3.118 1997     2505      5     2.549
## 2859 0030719030 3.116 1997     2210      5     2.547
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
UniqueAuthors +
##     Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##     k.max = 1000))
##  \--> method = "MM"
## Residuals:
##    Min     1Q Median     3Q    Max
## -1.594 -0.365  0.013  0.338  2.549
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)         0.559501   0.038782   14.43  < 2e-16 ***
## FirstAuthorFemale1  0.024062   0.016786    1.43  0.15179
## LastAuthorFemale1   0.000545   0.019814    0.03  0.97805
## UniqueAuthors2      0.648734   0.025557   25.38  < 2e-16 ***
## UniqueAuthors3      0.750185   0.025412   29.52  < 2e-16 ***
## UniqueAuthors4      0.810516   0.024425   33.18  < 2e-16 ***
## UniqueAuthors5      0.849294   0.023123   36.73  < 2e-16 ***
## Year1997            0.009147   0.045590    0.20  0.84100
## Year1998           -0.043257   0.047176   -0.92  0.35922
```

```
## Year1999           -0.103371   0.051403   -2.01  0.04437 *
## Year2000            0.009304   0.048238    0.19  0.84707
## Year2001            0.184963   0.058105    3.18  0.00146 **
## Year2002            0.088824   0.053092    1.67  0.09437 .
## Year2003            0.083111   0.045459    1.83  0.06756 .
## Year2004            0.029192   0.046209    0.63  0.52758
## Year2005            0.035205   0.043984    0.80  0.42351
## Year2006           -0.070334   0.046350   -1.52  0.12920
## Year2007           -0.043886   0.043519   -1.01  0.31329
## Year2008           -0.096405   0.043454   -2.22  0.02655 *
## Year2009           -0.168285   0.040856   -4.12  3.9e-05 ***
## Year2010           -0.131492   0.039745   -3.31  0.00094 ***
## Year2011           -0.170961   0.041375   -4.13  3.6e-05 ***
## Year2012           -0.219280   0.042365   -5.18  2.3e-07 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.471
## Multiple R-squared:  0.297,  Adjusted R-squared:  0.295
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
##  5 observations c(749,770,1234,2215,6142)
##   are outliers with |weight| = 0 ( < 1.6e-05);
##  467 weights are ~= 1. The remaining 5781 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0104  0.8610  0.9400  0.8850  0.9840  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.60e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.034  1           1.017
## LastAuthorFemale  1.047  1           1.023
## Year              1.069 16           1.002
```



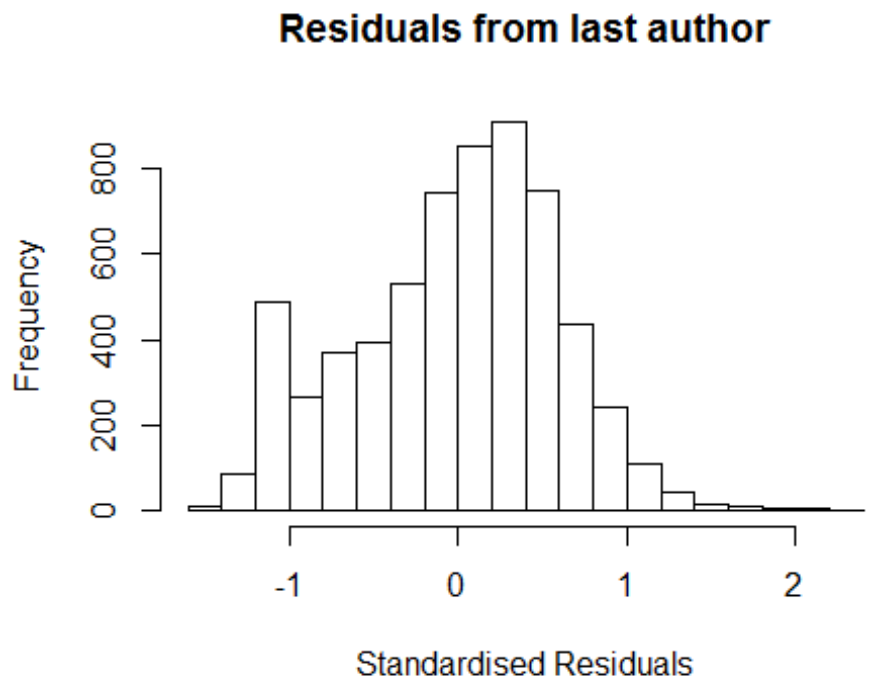

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.4682 -0.4173  0.0649  0.4102  2.3006
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)          1.07498    0.03915   27.46  < 2e-16 ***
## FirstAuthorFemale1   0.11622    0.01908    6.09  1.2e-09 ***
## LastAuthorFemale1    0.02692    0.02426    1.11   0.2673
## Year1997            -0.00909    0.05079   -0.18   0.8580
## Year1998            -0.01657    0.05216   -0.32   0.7507
## Year1999            -0.11918    0.05731   -2.08   0.0376 *
## Year2000            -0.00472    0.05550   -0.09   0.9322
## Year2001             0.25006    0.06207    4.03  5.7e-05 ***
## Year2002             0.22040    0.05608    3.93  8.6e-05 ***
## Year2003             0.20743    0.04860    4.27  2.0e-05 ***
## Year2004             0.14869    0.05064    2.94   0.0033 **
## Year2005             0.23193    0.04791    4.84  1.3e-06 ***
```

```
## Year2006           -0.00315    0.05471   -0.06   0.9541
## Year2007           -0.02940    0.05227   -0.56   0.5738
## Year2008           -0.01246    0.05250   -0.24   0.8124
## Year2009           -0.09380    0.04844   -1.94   0.0529 .
## Year2010           -0.02519    0.04679   -0.54   0.5904
## Year2011           -0.01165    0.04628   -0.25   0.8013
## Year2012           -0.09296    0.04871   -1.91   0.0564 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.574
## Multiple R-squared:  0.0364, Adjusted R-squared:  0.0336
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
##  465 weights are ~= 1. The remaining 5788 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0723  0.8520  0.9460  0.8990  0.9840  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.60e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.029  1           1.015
## Year              1.029 16           1.001
```




```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.4458 -0.4182  0.0651  0.4117  2.2987
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.07649    0.03913   27.51  < 2e-16 ***
## FirstAuthorFemale1  0.11959    0.01954    6.12  1.0e-09 ***
## Year1997           -0.00934    0.05081   -0.18   0.8542
## Year1998           -0.01622    0.05216   -0.31   0.7559
## Year1999           -0.11987    0.05729   -2.09   0.0365 *
## Year2000           -0.00405    0.05552   -0.07   0.9419
## Year2001            0.24974    0.06203    4.03  5.7e-05 ***
## Year2002            0.22135    0.05611    3.94  8.1e-05 ***
## Year2003            0.20881    0.04861    4.30  1.8e-05 ***
## Year2004            0.14949    0.05070    2.95   0.0032 **
## Year2005            0.23252    0.04787    4.86  1.2e-06 ***
## Year2006           -0.00285    0.05470   -0.05   0.9584
```

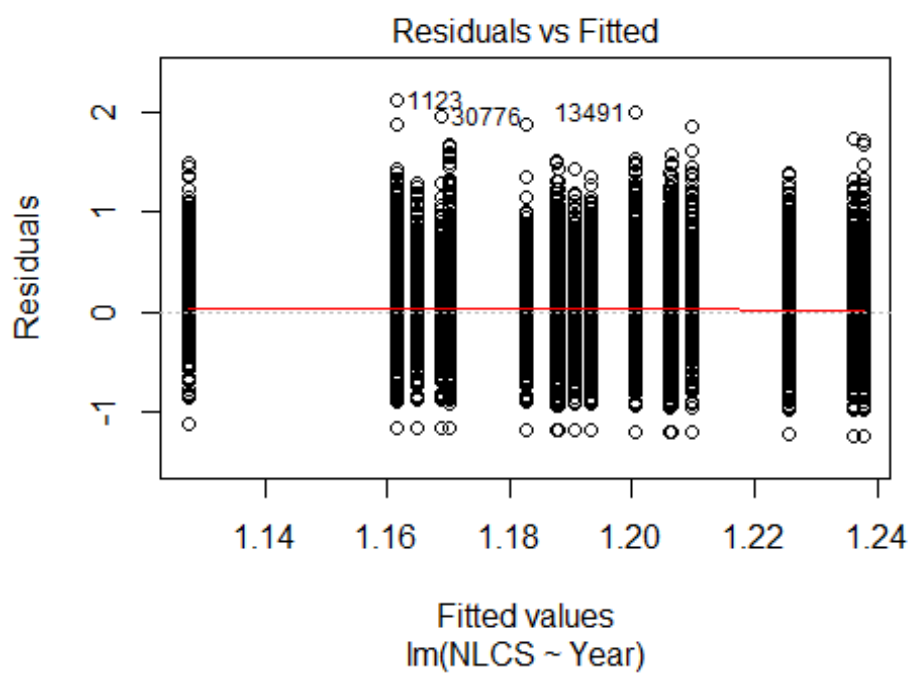
```
## Year2007           -0.02904    0.05230   -0.56   0.5788
## Year2008           -0.01156    0.05250   -0.22   0.8258
## Year2009           -0.09149    0.04841   -1.89   0.0588 .
## Year2010           -0.02373    0.04679   -0.51   0.6120
## Year2011           -0.00929    0.04621   -0.20   0.8407
## Year2012           -0.09074    0.04868   -1.86   0.0624 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.574
## Multiple R-squared:  0.0361, Adjusted R-squared:  0.0334
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
##  467 weights are ~= 1. The remaining 5786 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0728  0.8520  0.9460  0.8990  0.9840  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.60e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##                  GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.04  1           1.020
## Year             1.04 16           1.001
```



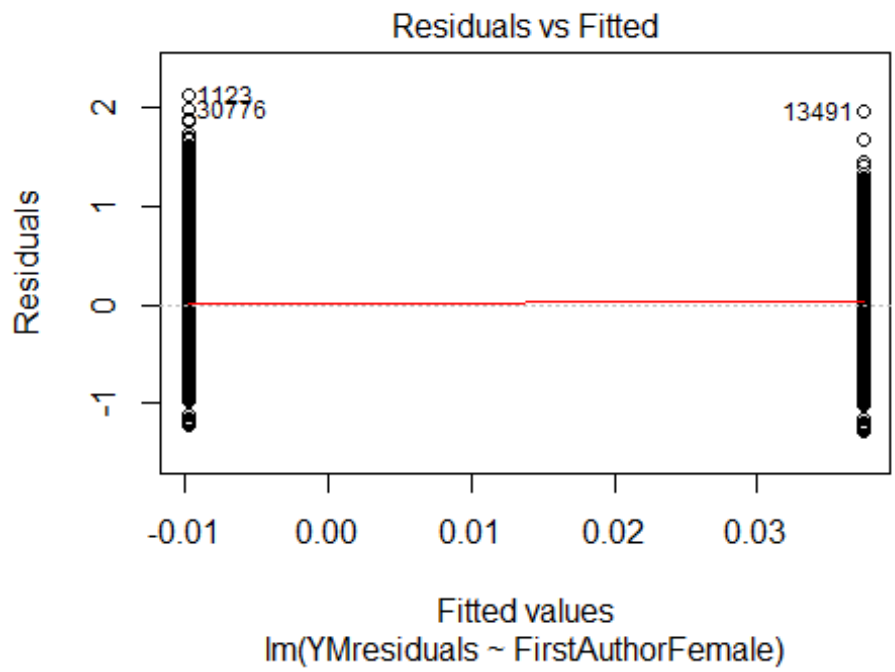

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.3870 -0.4201  0.0645  0.4088  2.2906
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.08912    0.03887   28.02  < 2e-16 ***
## LastAuthorFemale1   0.04839    0.02439    1.98   0.0473 *
## Year1997           -0.01137    0.05072   -0.22   0.8226
## Year1998           -0.02070    0.05201   -0.40   0.6907
## Year1999           -0.12219    0.05755   -2.12   0.0338 *
## Year2000           -0.00413    0.05545   -0.07   0.9406
## Year2001            0.24945    0.06187    4.03  5.6e-05 ***
## Year2002            0.22712    0.05564    4.08  4.5e-05 ***
## Year2003            0.21337    0.04837    4.41  1.0e-05 ***
## Year2004            0.15881    0.05031    3.16   0.0016 **
## Year2005            0.23877    0.04777    5.00  5.9e-07 ***
## Year2006            0.00106    0.05453    0.02   0.9845
```

```
## Year2007          -0.02510    0.05235   -0.48   0.6316
## Year2008          -0.00338    0.05241   -0.06   0.9486
## Year2009          -0.09004    0.04842   -1.86   0.0630 .
## Year2010          -0.02048    0.04675   -0.44   0.6614
## Year2011          -0.00225    0.04621   -0.05   0.9611
## Year2012          -0.08534    0.04872   -1.75   0.0799 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.575
## Multiple R-squared:  0.0314, Adjusted R-squared:  0.0288
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
##  490 weights are ~= 1. The remaining 5763 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0766  0.8500  0.9450  0.8990  0.9840  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.60e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Sample size for the above analysis:  6253"
## [1] ""
## [1] ""
## [1] "####################################"
## [1] "Analysis of AJSC 2507"
## [1] "####################################"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 2040 2140 1948 1760 1634 1557 1463 1408 1604 1495 1619 1528 1567 1487 1637
## 2011 2012
## 1464 1635
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  888  967  820  715  635  508  659  670  752  764  814  801  844  784  844
## 2011 2012
```

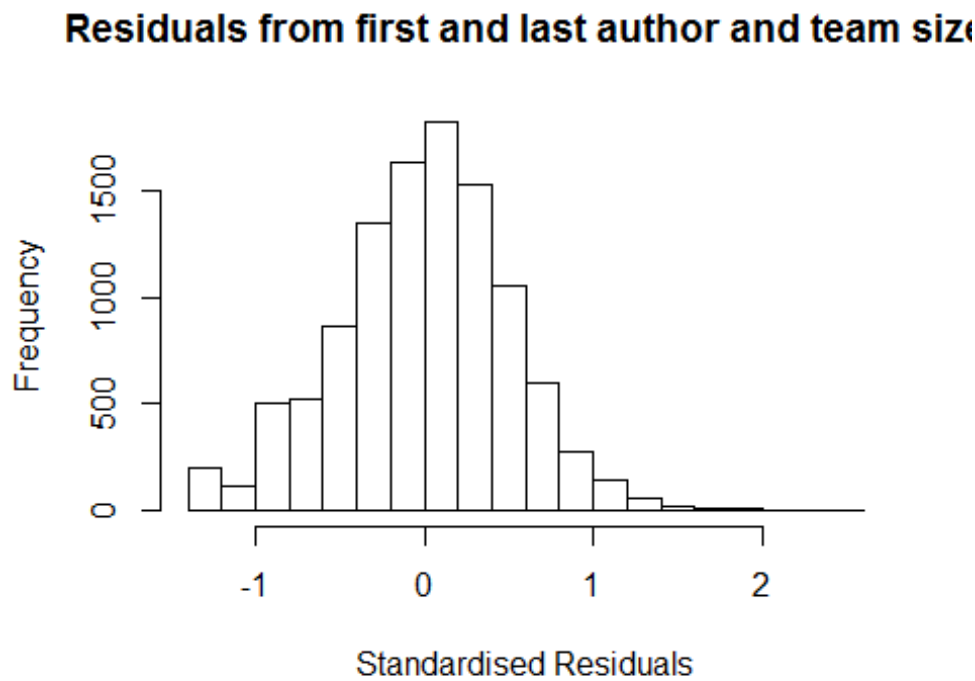
```
##  755  837
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  749  815  682  623  527  417  552  552  621  624  635  648  670  642  687
## 2011 2012
##  566  671
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
##  Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 100, df = 16, p-value = 6e-15
```




```
##
##  Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 13, df = 1, p-value = 3e-04
```

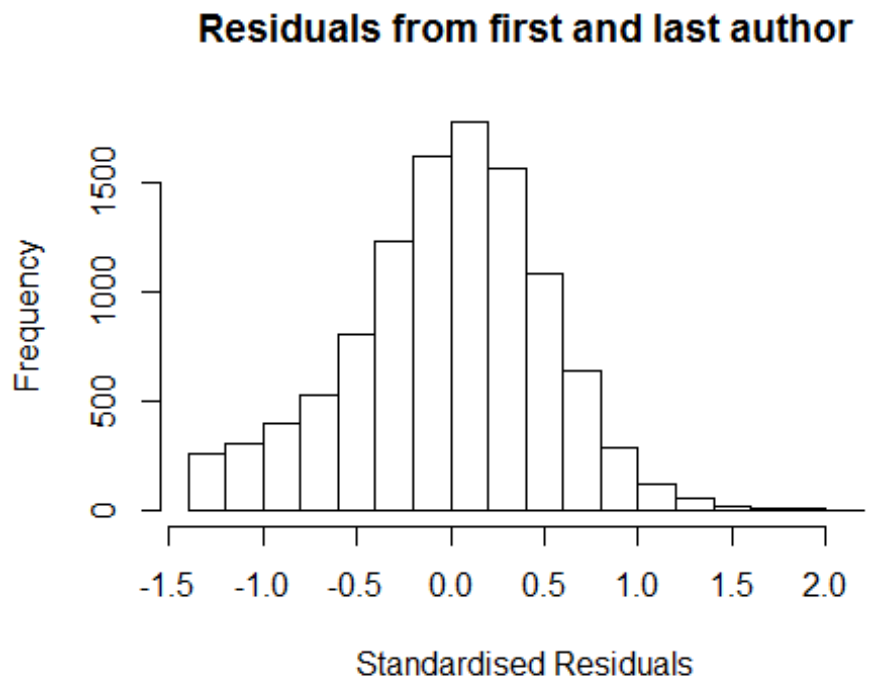



```
## [1] "Female first author team size 2018 geometric mean: 3.92607542216874"
## [1] "Male first author team size 2018 geometric mean: 3.59558064580367"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 47000, p-value = 0.08
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 3.7849360396679"
## [1] "Male last author team size 2018 geometric mean: 3.64027423817039"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 38000, p-value = 0.5
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.029  1           1.015
## LastAuthorFemale  1.015  1           1.007
## UniqueAuthors     1.126  4           1.015
## Year              1.128 16           1.004
```



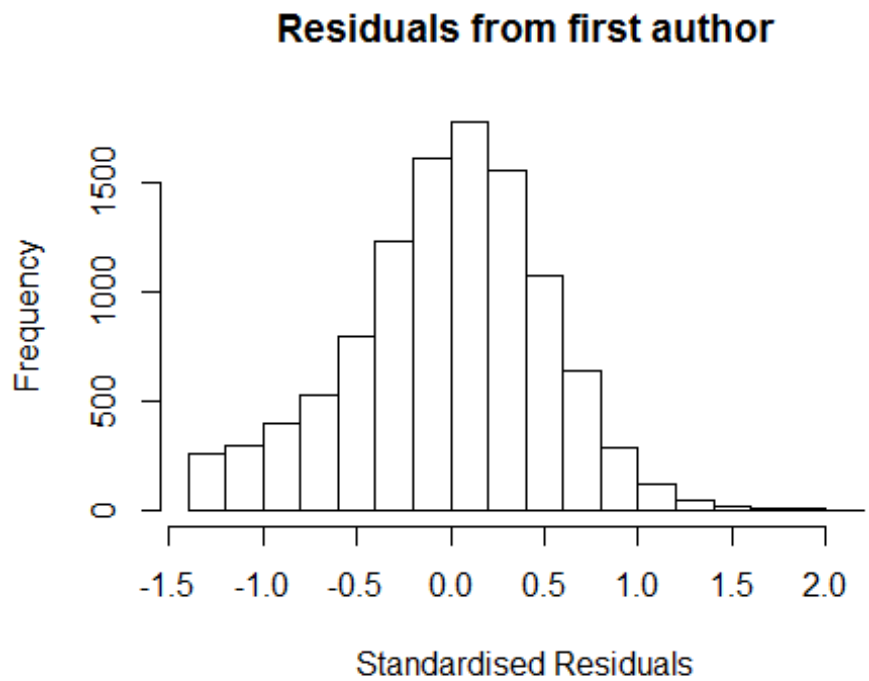

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
UniqueAuthors +
##     Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##     k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.3664 -0.3265  0.0186  0.3241  2.4321
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)         0.846868   0.026229   32.29  < 2e-16 ***
## FirstAuthorFemale1  0.021311   0.011680    1.82   0.0681 .
## LastAuthorFemale1  -0.014139   0.014832   -0.95   0.3405
## UniqueAuthors2      0.377051   0.019983   18.87  < 2e-16 ***
## UniqueAuthors3      0.441111   0.020602   21.41  < 2e-16 ***
## UniqueAuthors4      0.475960   0.021291   22.36  < 2e-16 ***
## UniqueAuthors5      0.491331   0.020896   23.51  < 2e-16 ***
## Year1997            0.002152   0.028950    0.07   0.9408
## Year1998            0.023497   0.030236    0.78   0.4371
## Year1999            0.000968   0.030269    0.03   0.9745
```

```
## Year2000           -0.016633   0.032236   -0.52   0.6059
## Year2001           -0.018985   0.034414   -0.55   0.5812
## Year2002           -0.023146   0.031515   -0.73   0.4627
## Year2003            0.028231   0.031095    0.91   0.3640
## Year2004           -0.008430   0.029507   -0.29   0.7751
## Year2005           -0.027251   0.028047   -0.97   0.3313
## Year2006           -0.025261   0.029155   -0.87   0.3863
## Year2007           -0.031632   0.029376   -1.08   0.2816
## Year2008           -0.034600   0.028272   -1.22   0.2210
## Year2009           -0.059994   0.029040   -2.07   0.0389 *
## Year2010           -0.111453   0.028531   -3.91  9.4e-05 ***
## Year2011           -0.058166   0.029024   -2.00   0.0451 *
## Year2012           -0.086883   0.027811   -3.12   0.0018 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.481
## Multiple R-squared:  0.094,  Adjusted R-squared:  0.0922
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
##  2 observations c(384,10412) are outliers with |weight| = 0 ( < 9.4e-06);
##  910 weights are ~= 1. The remaining 9769 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0001  0.8640  0.9500  0.8990  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          9.36e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.018  1           1.009
## LastAuthorFemale  1.009  1           1.005
## Year              1.025 16           1.001
```



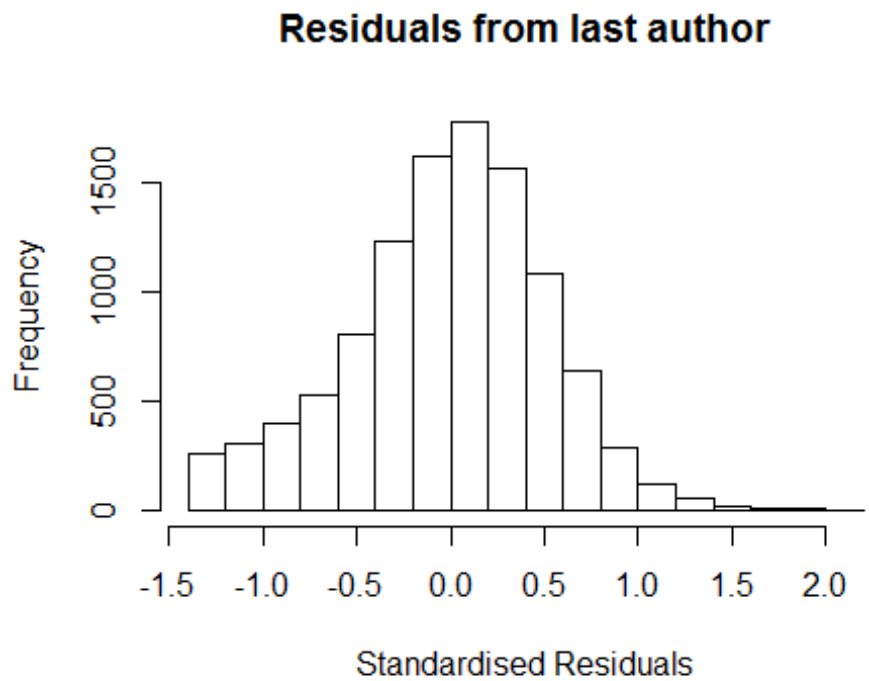

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.3153 -0.3286  0.0233  0.3339  2.1149
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)          1.16413    0.02276   51.16   <2e-16 ***
## FirstAuthorFemale1   0.05846    0.01189    4.91    9e-07 ***
## LastAuthorFemale1   -0.00483    0.01525   -0.32   0.7515
## Year1997            -0.00608    0.03029   -0.20   0.8409
## Year1998             0.05075    0.03191    1.59   0.1117
## Year1999             0.02372    0.03083    0.77   0.4417
## Year2000             0.01673    0.03333    0.50   0.6157
## Year2001             0.03432    0.03542    0.97   0.3325
## Year2002             0.04312    0.03290    1.31   0.1900
## Year2003             0.09275    0.03245    2.86   0.0043 **
## Year2004             0.06174    0.03092    2.00   0.0459 *
## Year2005             0.06319    0.02912    2.17   0.0300 *
```

```
## Year2006            0.04341    0.03069    1.41   0.1572
## Year2007            0.03057    0.03056    1.00   0.3173
## Year2008            0.04122    0.02944    1.40   0.1615
## Year2009            0.01742    0.03039    0.57   0.5665
## Year2010           -0.02580    0.03049   -0.85   0.3976
## Year2011            0.03890    0.02989    1.30   0.1932
## Year2012            0.01646    0.02914    0.56   0.5721
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.49
## Multiple R-squared:  0.00501,    Adjusted R-squared:  0.00333
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
##  910 weights are ~= 1. The remaining 9771 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0225  0.8590  0.9490  0.8940  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          9.36e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.017  1           1.008
## Year              1.017 16           1.001
```




```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.3143 -0.3285  0.0236  0.3341  2.1154
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)          1.16361    0.02267   51.33  < 2e-16 ***
## FirstAuthorFemale1   0.05787    0.01199    4.83  1.4e-06 ***
## Year1997            -0.00605    0.03029   -0.20   0.8418
## Year1998             0.05076    0.03190    1.59   0.1117
## Year1999             0.02382    0.03083    0.77   0.4398
## Year2000             0.01692    0.03333    0.51   0.6118
## Year2001             0.03459    0.03540    0.98   0.3285
## Year2002             0.04334    0.03288    1.32   0.1875
## Year2003             0.09286    0.03245    2.86   0.0042 **
## Year2004             0.06191    0.03092    2.00   0.0453 *
## Year2005             0.06315    0.02912    2.17   0.0302 *
## Year2006             0.04341    0.03069    1.41   0.1573
```

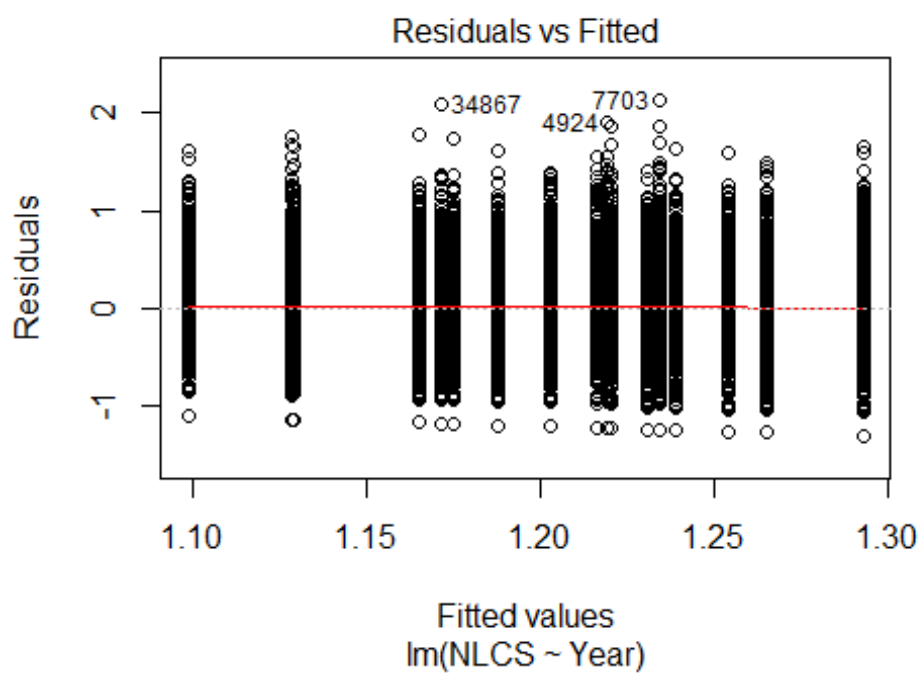
```
## Year2007            0.03056    0.03057    1.00   0.3175
## Year2008            0.04115    0.02944    1.40   0.1623
## Year2009            0.01742    0.03039    0.57   0.5665
## Year2010           -0.02597    0.03048   -0.85   0.3942
## Year2011            0.03868    0.02989    1.29   0.1957
## Year2012            0.01638    0.02913    0.56   0.5739
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.49
## Multiple R-squared:  0.00502,    Adjusted R-squared:  0.00343
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
##  911 weights are ~= 1. The remaining 9770 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0223  0.8590  0.9490  0.8940  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          9.36e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##                   GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.009  1           1.004
## Year             1.009 16           1.000
```



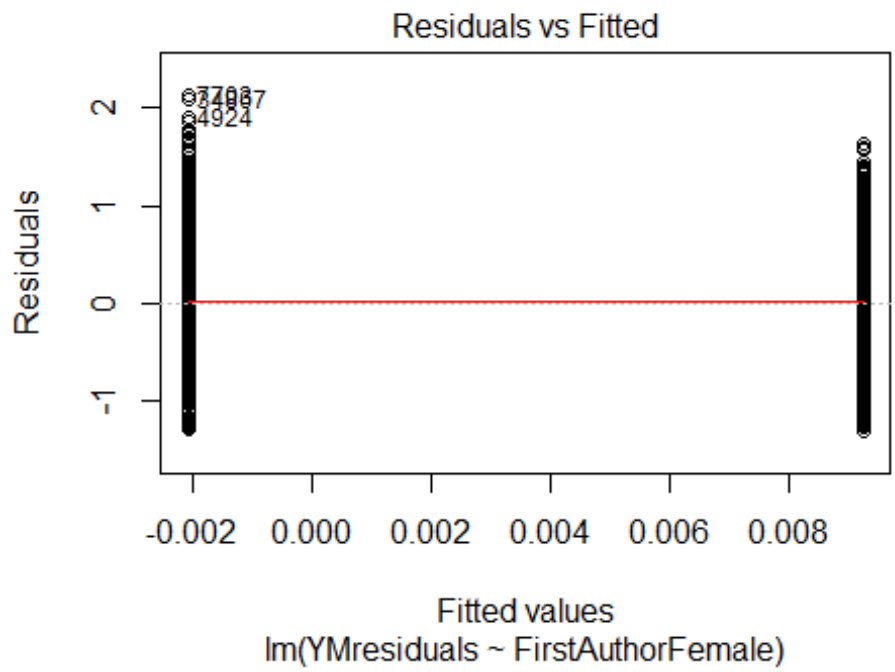

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.2730 -0.3282  0.0245  0.3371  2.1051
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.17389    0.02266   51.81   <2e-16 ***
## LastAuthorFemale1   0.00596    0.01526    0.39   0.6960
## Year1997           -0.00636    0.03030   -0.21   0.8336
## Year1998            0.04998    0.03196    1.56   0.1179
## Year1999            0.02457    0.03090    0.80   0.4265
## Year2000            0.01725    0.03333    0.52   0.6048
## Year2001            0.03484    0.03540    0.98   0.3250
## Year2002            0.04453    0.03287    1.35   0.1755
## Year2003            0.09317    0.03248    2.87   0.0041 **
## Year2004            0.06175    0.03096    1.99   0.0462 *
## Year2005            0.06432    0.02914    2.21   0.0273 *
## Year2006            0.04558    0.03077    1.48   0.1386
```

```
## Year2007           0.03328    0.03061    1.09   0.2769
## Year2008           0.04503    0.02945    1.53   0.1263
## Year2009           0.02276    0.03034    0.75   0.4531
## Year2010          -0.02243    0.03050   -0.74   0.4622
## Year2011           0.04343    0.02994    1.45   0.1469
## Year2012           0.01925    0.02916    0.66   0.5091
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.491
## Multiple R-squared:  0.00288,    Adjusted R-squared:  0.00129
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
##  882 weights are ~= 1. The remaining 9799 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.026   0.859   0.950   0.895   0.986   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          9.36e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Sample size for the above analysis:  10681"
## [1] ""
## [1] ""
## [1] "####################################"
## [1] "Analysis of AJSC 2508"
## [1] "####################################"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 2104 2609 2471 2275 2217 2242 1866 1967 2088 2136 2394 2608 2650 2707 2758
## 2011 2012
## 2722 2681
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  803 1057 1036  833  926  855  859  896  922  982 1081 1200 1287 1319 1398
## 2011 2012
```

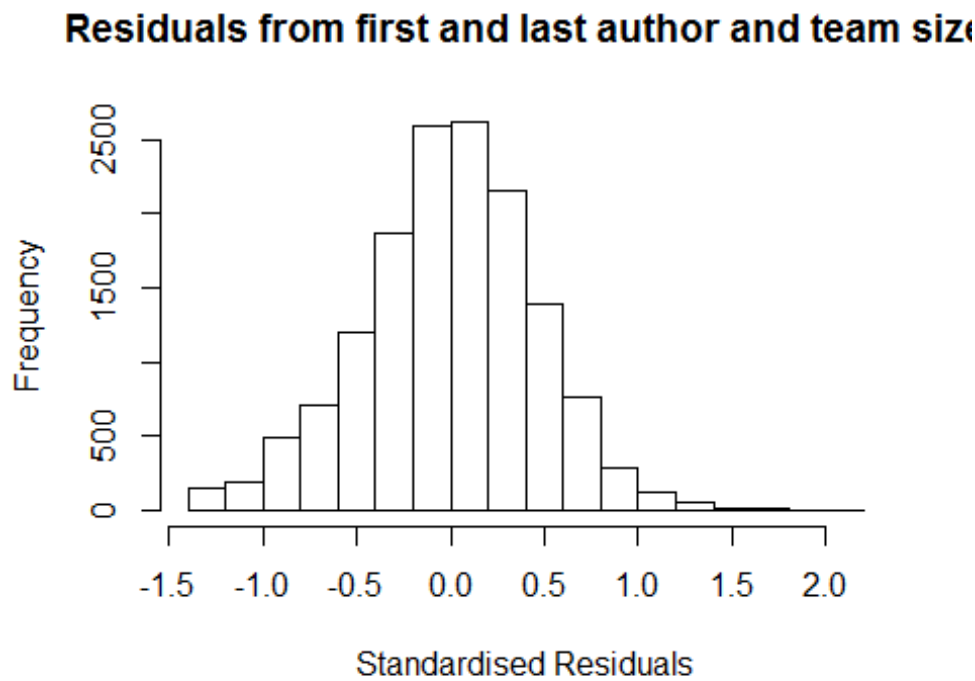
```
## 1429 1447
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  657  843  848  690  767  702  700  701  739  763  833  942 1016 1029 1130
## 2011 2012
## 1124 1139
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
##  Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 120, df = 16, p-value <2e-16
```




```
##
##  Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 19, df = 1, p-value = 1e-05
```

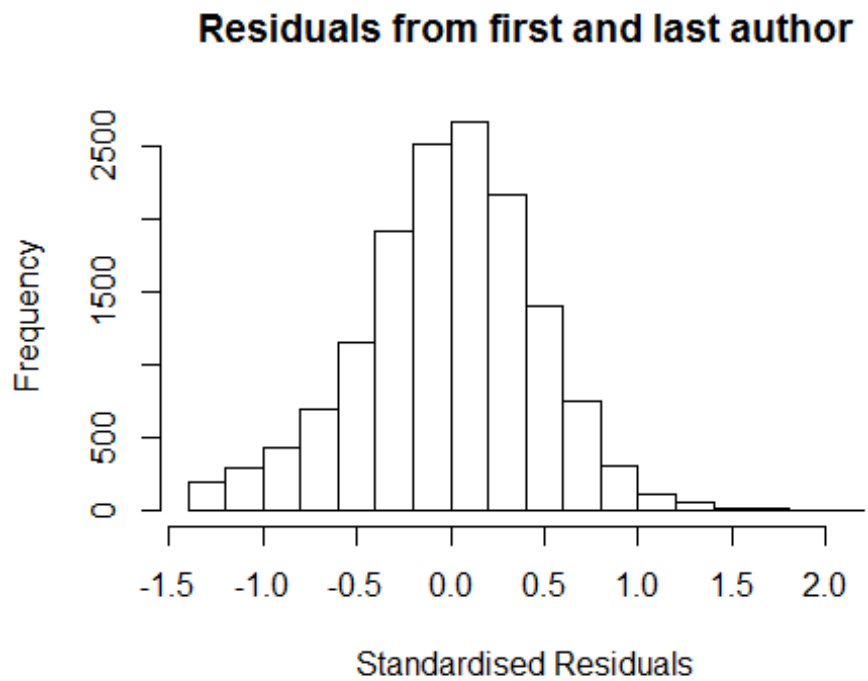



```
## [1] "Female first author team size 2018 geometric mean: 3.81087793058056"
## [1] "Male first author team size 2018 geometric mean: 3.66282815844253"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 99000, p-value = 0.6
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 3.7571877115154"
## [1] "Male last author team size 2018 geometric mean: 3.68428130459692"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 79000, p-value = 0.5
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.025  1           1.012
## LastAuthorFemale  1.017  1           1.008
## UniqueAuthors     1.086  4           1.010
## Year              1.082 16           1.002
```



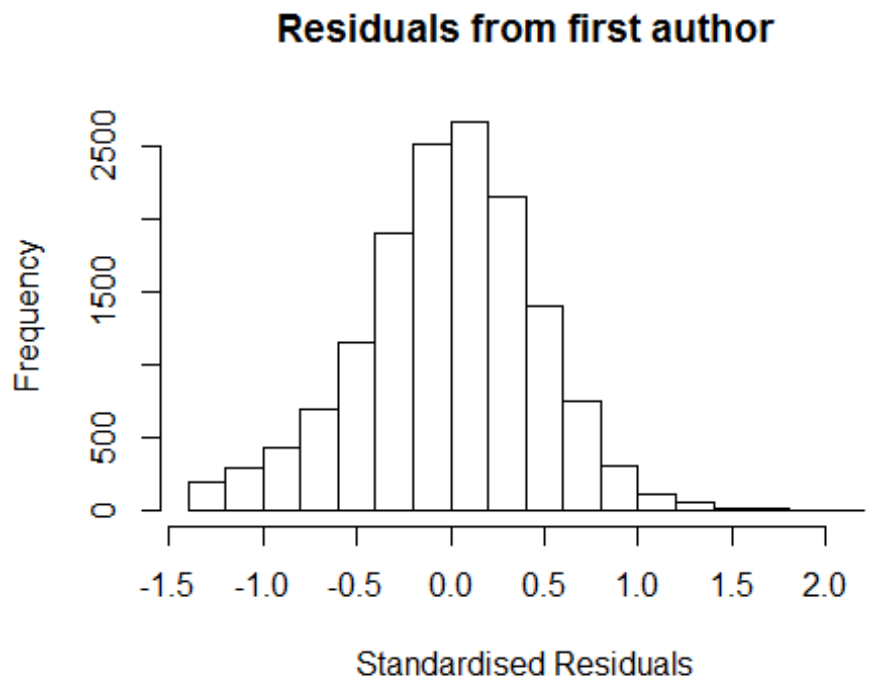

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
UniqueAuthors +
##     Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##     k.max = 1000))
##  \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.39970 -0.29074  0.00766  0.30012  2.12292
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         0.97984    0.02556   38.34  < 2e-16 ***
## FirstAuthorFemale1 -0.00575    0.00946   -0.61  0.54343
## LastAuthorFemale1  -0.00595    0.01170   -0.51  0.61115
## UniqueAuthors2      0.24424    0.01885   12.96  < 2e-16 ***
## UniqueAuthors3      0.28502    0.01903   14.97  < 2e-16 ***
## UniqueAuthors4      0.32240    0.01961   16.44  < 2e-16 ***
## UniqueAuthors5      0.35773    0.01941   18.43  < 2e-16 ***
## Year1997            0.01524    0.02697    0.57  0.57195
## Year1998            0.02777    0.02748    1.01  0.31220
## Year1999           -0.00182    0.02704   -0.07  0.94623
```

```
## Year2000           -0.03379    0.02758   -1.22  0.22060
## Year2001            0.06213    0.02774    2.24  0.02512 *
## Year2002            0.02847    0.02701    1.05  0.29191
## Year2003           -0.00727    0.02625   -0.28  0.78173
## Year2004           -0.01570    0.02603   -0.60  0.54649
## Year2005           -0.01187    0.02504   -0.47  0.63537
## Year2006           -0.09715    0.02579   -3.77  0.00017 ***
## Year2007           -0.05388    0.02473   -2.18  0.02939 *
## Year2008           -0.07762    0.02500   -3.10  0.00191 **
## Year2009           -0.07810    0.02459   -3.18  0.00150 **
## Year2010           -0.12875    0.02408   -5.35  9.1e-08 ***
## Year2011           -0.17569    0.02435   -7.22  5.6e-13 ***
## Year2012           -0.12794    0.02483   -5.15  2.6e-07 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.439
## Multiple R-squared:  0.0555, Adjusted R-squared:  0.0541
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
##  3 observations c(1439,2318,10819)
##   are outliers with |weight| = 0 ( < 6.8e-06);
##  1274 weights are ~= 1. The remaining 13346 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0024  0.8650  0.9500  0.8980  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          6.84e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.010  1           1.005
## LastAuthorFemale  1.011  1           1.005
## Year              1.020 16           1.001
```



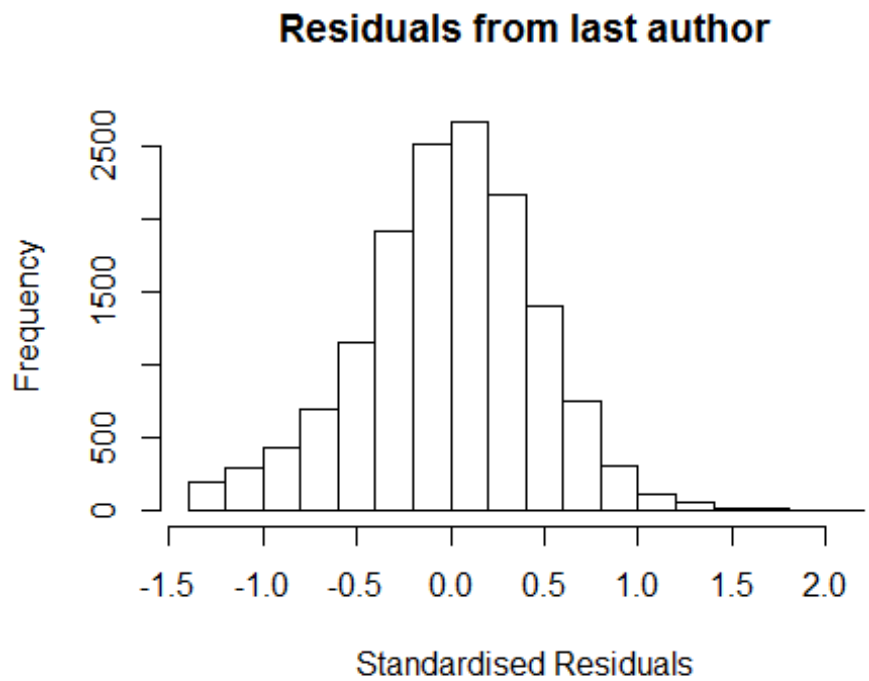

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.3098 -0.2948  0.0102  0.3005  2.1035
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)          1.21250    0.02014   60.20  < 2e-16 ***
## FirstAuthorFemale1   0.01592    0.00955    1.67  0.09567 .
## LastAuthorFemale1    0.00238    0.01194    0.20  0.84220
## Year1997             0.01651    0.02714    0.61  0.54299
## Year1998             0.04297    0.02737    1.57  0.11648
## Year1999             0.01913    0.02723    0.70  0.48238
## Year2000            -0.01597    0.02769   -0.58  0.56407
## Year2001             0.07898    0.02771    2.85  0.00437 **
## Year2002             0.05537    0.02725    2.03  0.04221 *
## Year2003             0.02309    0.02619    0.88  0.37795
## Year2004             0.01909    0.02600    0.73  0.46286
## Year2005             0.02812    0.02502    1.12  0.26108
```

```
## Year2006           -0.05486    0.02574   -2.13  0.03308 *
## Year2007           -0.01131    0.02475   -0.46  0.64753
## Year2008           -0.03493    0.02498   -1.40  0.16204
## Year2009           -0.03846    0.02444   -1.57  0.11555
## Year2010           -0.07999    0.02391   -3.35  0.00082 ***
## Year2011           -0.12097    0.02414   -5.01  5.5e-07 ***
## Year2012           -0.07187    0.02478   -2.90  0.00374 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.443
## Multiple R-squared:  0.0132, Adjusted R-squared:  0.012
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  2 observations c(2318,10819) are outliers with |weight| = 0 ( < 6.8e-06);
##  1228 weights are ~= 1. The remaining 13393 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0294  0.8640  0.9510  0.8960  0.9860  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          6.84e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##                   GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.01  1           1.005
## Year              1.01 16           1.000
```




```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.3078 -0.2946  0.0101  0.3009  2.1034
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)          1.21267    0.02012   60.26  < 2e-16 ***
## FirstAuthorFemale1   0.01612    0.00959    1.68  0.09274 .
## Year1997             0.01656    0.02713    0.61  0.54157
## Year1998             0.04295    0.02737    1.57  0.11668
## Year1999             0.01913    0.02723    0.70  0.48236
## Year2000            -0.01590    0.02769   -0.57  0.56582
## Year2001             0.07903    0.02771    2.85  0.00435 **
## Year2002             0.05540    0.02726    2.03  0.04214 *
## Year2003             0.02314    0.02619    0.88  0.37683
## Year2004             0.01917    0.02600    0.74  0.46087
## Year2005             0.02818    0.02502    1.13  0.26005
## Year2006            -0.05482    0.02574   -2.13  0.03320 *
```

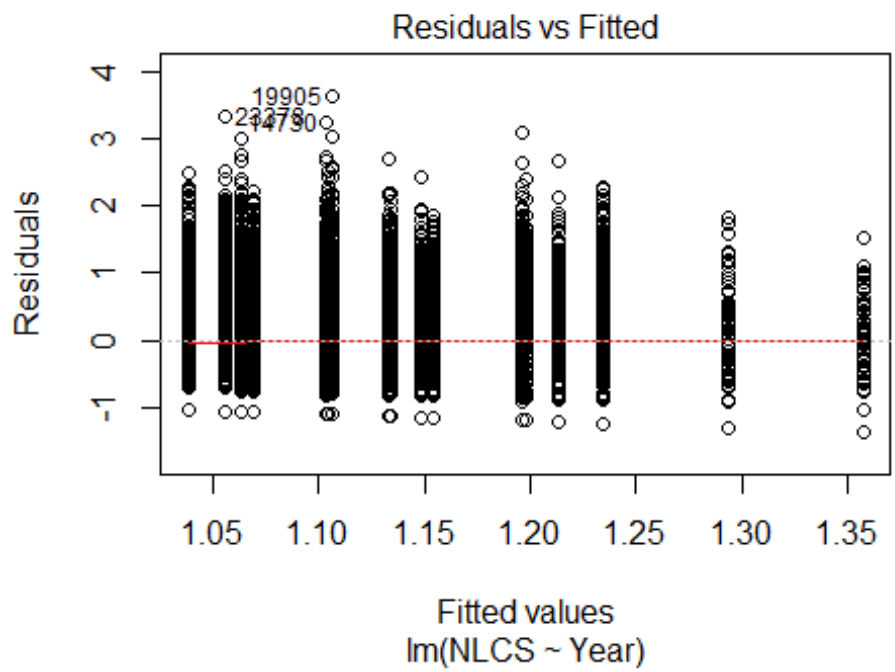
```
## Year2007           -0.01125    0.02474   -0.45  0.64936
## Year2008           -0.03481    0.02497   -1.39  0.16325
## Year2009           -0.03836    0.02443   -1.57  0.11634
## Year2010           -0.07989    0.02391   -3.34  0.00084 ***
## Year2011           -0.12088    0.02414   -5.01  5.6e-07 ***
## Year2012           -0.07172    0.02476   -2.90  0.00378 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.443
## Multiple R-squared:  0.0132, Adjusted R-squared:  0.012
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  2 observations c(2318,10819) are outliers with |weight| = 0 ( < 6.8e-06);
##  1238 weights are ~= 1. The remaining 13383 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0295  0.8640  0.9510  0.8960  0.9860  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          6.84e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##                   GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.011  1           1.005
## Year             1.011 16           1.000
```



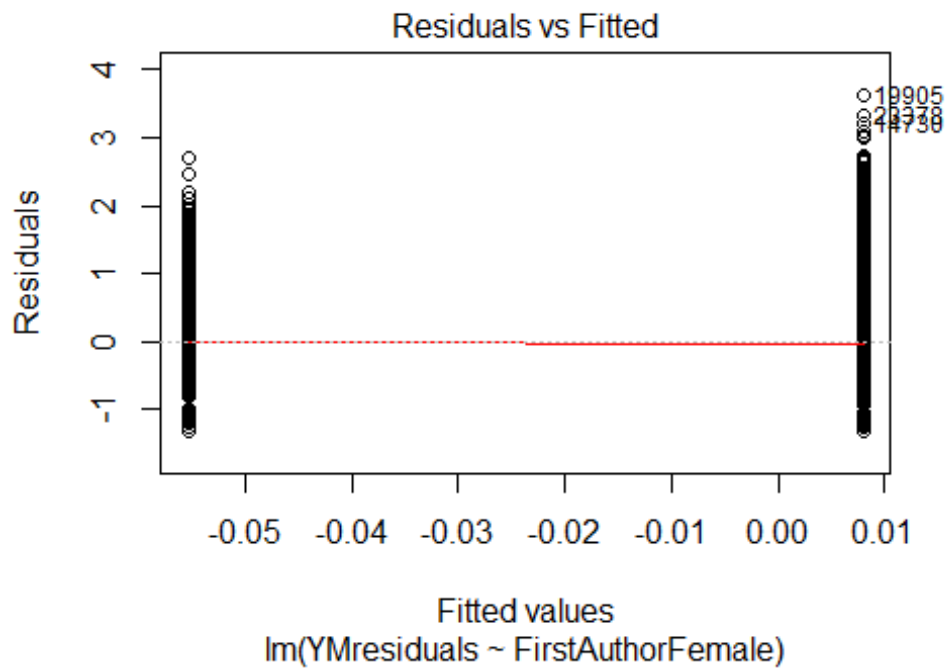

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.29834 -0.29548  0.00979  0.30131  2.10126
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.21453    0.02009   60.46  < 2e-16 ***
## LastAuthorFemale1   0.00439    0.01198    0.37  0.71397
## Year1997            0.01711    0.02714    0.63  0.52846
## Year1998            0.04321    0.02738    1.58  0.11455
## Year1999            0.01951    0.02723    0.72  0.47377
## Year2000           -0.01564    0.02770   -0.56  0.57226
## Year2001            0.07942    0.02773    2.86  0.00419 **
## Year2002            0.05579    0.02725    2.05  0.04068 *
## Year2003            0.02366    0.02618    0.90  0.36612
## Year2004            0.02009    0.02599    0.77  0.43954
## Year2005            0.02891    0.02502    1.16  0.24789
## Year2006           -0.05364    0.02572   -2.09  0.03705 *
```

```
## Year2007          -0.01019    0.02474   -0.41  0.68055
## Year2008          -0.03361    0.02496   -1.35  0.17816
## Year2009          -0.03732    0.02442   -1.53  0.12649
## Year2010          -0.07886    0.02390   -3.30  0.00097 ***
## Year2011          -0.12002    0.02415   -4.97  6.7e-07 ***
## Year2012          -0.07072    0.02477   -2.85  0.00431 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.443
## Multiple R-squared:  0.013,  Adjusted R-squared:  0.0118
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  2 observations c(2318,10819) are outliers with |weight| = 0 ( < 6.8e-06);
##  1239 weights are ~= 1. The remaining 13382 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0302  0.8640  0.9500  0.8960  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          6.84e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Sample size for the above analysis:  14623"
## [1] ""
## [1] ""
## [1] "####################################"
## [1] "Analysis of AJSC 2600"
## [1] "####################################"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 1800  912  163 1803  114 1245 1196 1493 1353 1351 1711 1989 2020 1986 2066
## 2011 2012
## 2052 2108
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 1148  570   93 1173   70  821  792  998  878  874 1137 1341 1366 1367 1437
```

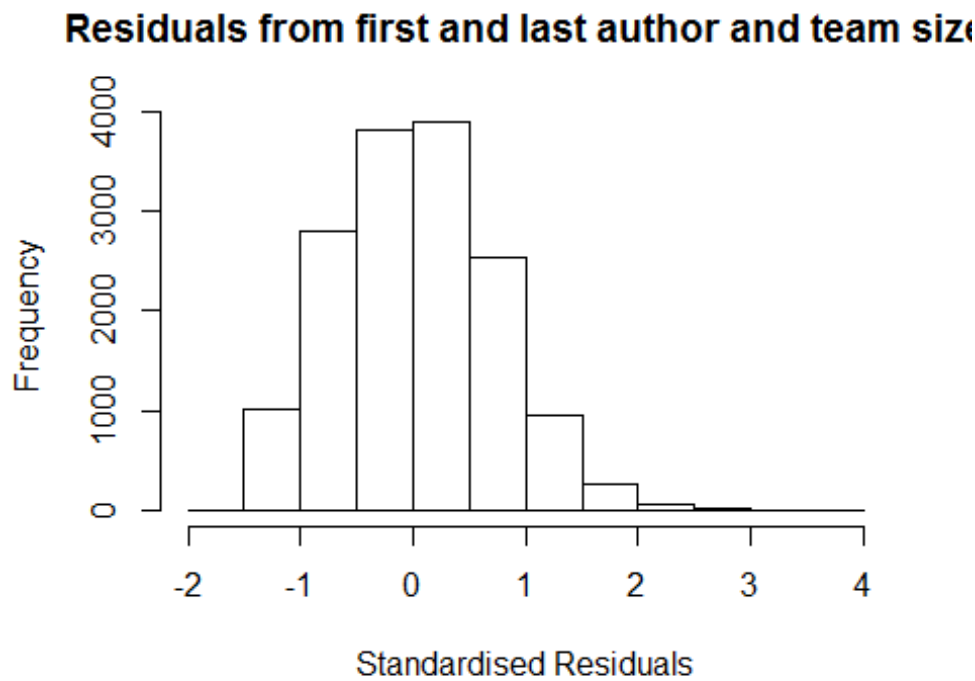
```
## 2011 2012
## 1451 1465
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 1029  520   87 1058   65  755  715  907  804  806 1034 1210 1210 1234 1268
## 2011 2012
## 1299 1312
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
##  Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 160, df = 16, p-value <2e-16
```




```
##
##  Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 5.9, df = 1, p-value = 0.02
```



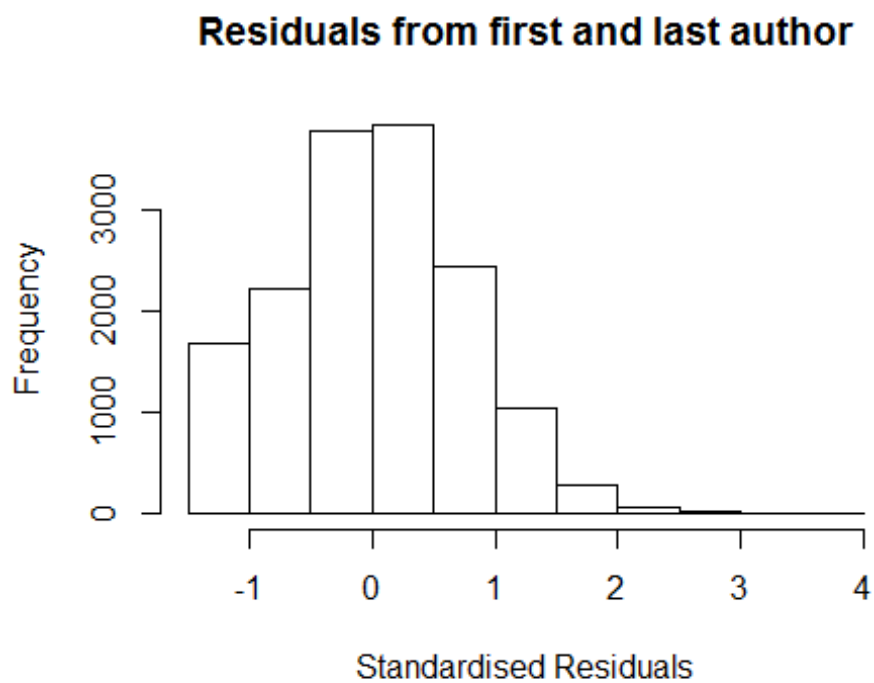

```
## [1] "Female first author team size 2018 geometric mean: 1.94727062825705"
## [1] "Male first author team size 2018 geometric mean: 1.46100408645266"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 89000, p-value = 2e-08
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 1.86676993042369"
## [1] "Male last author team size 2018 geometric mean: 1.47045642597383"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 88000, p-value = 6e-07
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.618  1           1.272
## LastAuthorFemale  1.612  1           1.270
## UniqueAuthors     1.033  4           1.004
## Year              1.034 16           1.001
```




```
## [1] "List of  16 outliers with residuals above  2.5"
##          ScopusId  NLCS Year OneField Fields residuals
## 2901   0031161793 3.605 1997     2600      1     2.503
## 4858   0033239662 4.280 1999     2600      2     3.192
## 7466   0036020892 3.891 2002     2600      2     2.797
## 9606   0037623983 3.818 2003     1700      2     2.562
## 14730 33745604236 4.326 2006     2600      2     2.981
## 15035 33646365077 3.822 2006     2600      2     2.842
## 15258 29144523061 3.760 2006     1700      7     2.540
## 18799 57349174008 4.128 2008     2600      3     2.802
## 19905 42649140570 4.730 2008     2600      1     3.769
## 22536 65549134300 3.585 2009     2200      3     2.678
## 22621 70049091784 3.562 2009     2600      2     2.655
## 23378 84969334819 4.392 2009     2600      2     3.245
## 24195 77957933737 3.515 2010     2600      1     2.629
## 26821 80053974840 3.721 2011     1700      2     2.562
## 27988 79551550744 4.060 2011     2600      2     2.776
## 29921 84861223341 3.537 2012     2600      1     2.589
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
UniqueAuthors +
##     Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##     k.max = 1000))
##  \--> method = "MM"
## Residuals:
```

```
##     Min      1Q  Median      3Q     Max
## -1.5975 -0.4913  0.0044  0.4932  3.7694
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)          1.1362     0.0255   44.59  < 2e-16 ***
## FirstAuthorFemale1  -0.0528     0.0218   -2.43  0.01525 *
## LastAuthorFemale1   -0.0555     0.0218   -2.54  0.01100 *
## UniqueAuthors2       0.2398     0.0136   17.64  < 2e-16 ***
## UniqueAuthors3       0.3652     0.0220   16.64  < 2e-16 ***
## UniqueAuthors4       0.4087     0.0457    8.93  < 2e-16 ***
## UniqueAuthors5       0.4547     0.0667    6.82  9.5e-12 ***
## Year1997            -0.0338     0.0409   -0.83  0.40897
## Year1998             0.0526     0.0871    0.60  0.54562
## Year1999            -0.0483     0.0341   -1.42  0.15687
## Year2000             0.1255     0.0729    1.72  0.08519 .
## Year2001            -0.1154     0.0355   -3.25  0.00114 **
## Year2002            -0.0422     0.0354   -1.19  0.23354
## Year2003            -0.1199     0.0335   -3.58  0.00035 ***
## Year2004            -0.1040     0.0336   -3.10  0.00196 **
## Year2005            -0.0995     0.0341   -2.92  0.00355 **
## Year2006            -0.1563     0.0329   -4.74  2.1e-06 ***
## Year2007            -0.1970     0.0318   -6.20  5.7e-10 ***
## Year2008            -0.1755     0.0317   -5.53  3.2e-08 ***
## Year2009            -0.2288     0.0324   -7.05  1.8e-12 ***
## Year2010            -0.2501     0.0328   -7.62  2.8e-14 ***
## Year2011            -0.2170     0.0338   -6.42  1.4e-10 ***
## Year2012            -0.1884     0.0346   -5.44  5.3e-08 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.734
## Multiple R-squared:  0.047,  Adjusted R-squared:  0.0456
## Convergence in 10 IRWLS iterations
##
## Robustness weights:
##  observation 9699 is an outlier with |weight| = 0 ( < 6.5e-06);
##  1263 weights are ~= 1. The remaining 14049 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0123  0.8660  0.9520  0.9150  0.9860  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          6.53e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
```

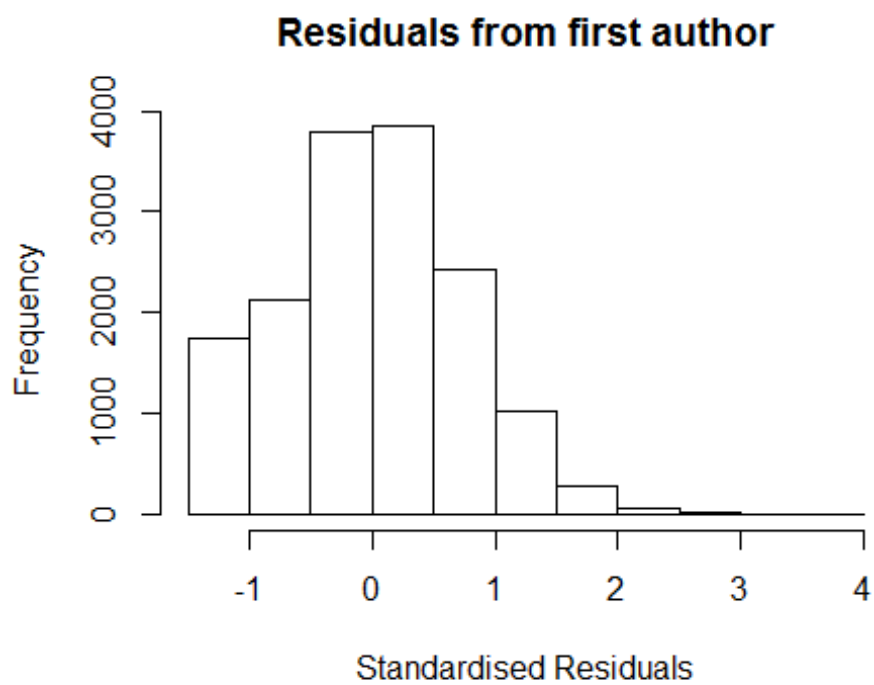
```
##           0        1000           0
##                   psi           subsampling                    cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.584  1           1.259
## LastAuthorFemale  1.582  1           1.258
## Year              1.013 16           1.000
```




```
## [1] "List of  18 outliers with residuals above  2.5"
##          ScopusId  NLCS Year OneField Fields residuals
## 4858   0033239662 4.280 1999     2600      2     3.111
## 5080   0001259111 3.825 1999     1804      3     2.690
## 7466   0036020892 3.891 2002     2600      2     2.695
## 9606   0037623983 3.818 2003     1700      2     2.701
## 14730 33745604236 4.326 2006     2600      2     3.249
## 15035 33646365077 3.822 2006     2600      2     2.745
## 15258 29144523061 3.760 2006     1700      7     2.683
## 18786 55649115527 3.676 2008     2600      3     2.607
## 18799 57349174008 4.128 2008     2600      3     3.059
## 18854 52149117439 3.641 2008     1700      2     2.572
## 19905 42649140570 4.730 2008     2600      1     3.661
## 22536 65549134300 3.585 2009     2200      3     2.566
```

```
## 22621 70049091784 3.562 2009     2600      2     2.543
## 23378 84969334819 4.392 2009     2600      2     3.373
## 24195 77957933737 3.515 2010     2600      1     2.518
## 26821 80053974840 3.721 2011     1700      2     2.688
## 27988 79551550744 4.060 2011     2600      2     3.027
## 30656 84856822033 3.605 2012     2600      1     2.575
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.26885 -0.50289 -0.00125  0.49661  3.66134
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)          1.2139     0.0257   47.15  < 2e-16 ***
## FirstAuthorFemale1  -0.0337     0.0221   -1.52   0.1274
## LastAuthorFemale1   -0.0387     0.0221   -1.75   0.0800 .
## Year1997            -0.0387     0.0414   -0.94   0.3498
## Year1998             0.0499     0.0862    0.58   0.5628
## Year1999            -0.0450     0.0345   -1.30   0.1924
## Year2000             0.1274     0.0740    1.72   0.0851 .
## Year2001            -0.1058     0.0361   -2.93   0.0034 **
## Year2002            -0.0179     0.0363   -0.49   0.6226
## Year2003            -0.0973     0.0342   -2.84   0.0045 **
## Year2004            -0.0735     0.0345   -2.13   0.0331 *
## Year2005            -0.0714     0.0349   -2.04   0.0409 *
## Year2006            -0.1370     0.0335   -4.08  4.5e-05 ***
## Year2007            -0.1696     0.0325   -5.21  1.9e-07 ***
## Year2008            -0.1452     0.0323   -4.49  7.0e-06 ***
## Year2009            -0.1950     0.0333   -5.85  5.0e-09 ***
## Year2010            -0.2170     0.0337   -6.45  1.2e-10 ***
## Year2011            -0.1807     0.0344   -5.26  1.5e-07 ***
## Year2012            -0.1449     0.0354   -4.10  4.2e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.752
## Multiple R-squared:  0.00978,    Adjusted R-squared:  0.00861
## Convergence in 10 IRWLS iterations
##
## Robustness weights:
##  observation 9699 is an outlier with |weight| = 0 ( < 6.5e-06);
##  1235 weights are ~= 1. The remaining 14077 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.007   0.858   0.952   0.916   0.985   0.999
## Algorithmic parameters:
```

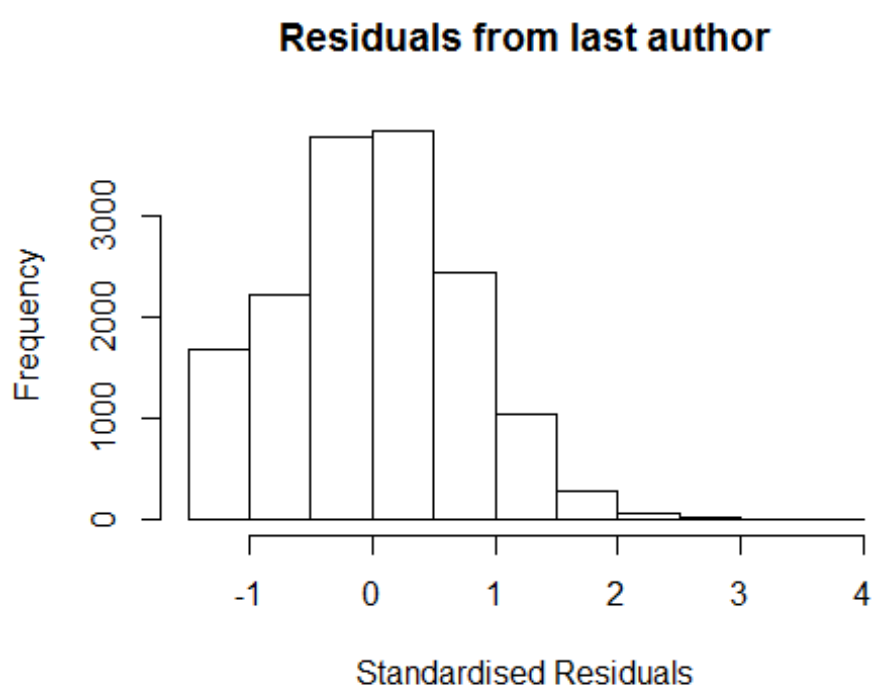
```
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          6.53e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##                   GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.01  1           1.005
## Year              1.01 16           1.000
```




```
## [1] "List of  18 outliers with residuals above  2.5"
##          ScopusId  NLCS Year OneField Fields residuals
## 4858   0033239662 4.280 1999     2600      2     3.111
## 5080   0001259111 3.825 1999     1804      3     2.690
## 7466   0036020892 3.891 2002     2600      2     2.695
## 9606   0037623983 3.818 2003     1700      2     2.701
## 14730 33745604236 4.326 2006     2600      2     3.249
```

```
## 15035 33646365077 3.822 2006     2600      2     2.745
## 15258 29144523061 3.760 2006     1700      7     2.683
## 18786 55649115527 3.676 2008     2600      3     2.607
## 18799 57349174008 4.128 2008     2600      3     3.059
## 18854 52149117439 3.641 2008     1700      2     2.572
## 19905 42649140570 4.730 2008     2600      1     3.661
## 22536 65549134300 3.585 2009     2200      3     2.566
## 22621 70049091784 3.562 2009     2600      2     2.543
## 23378 84969334819 4.392 2009     2600      2     3.373
## 24195 77957933737 3.515 2010     2600      1     2.518
## 26821 80053974840 3.721 2011     1700      2     2.688
## 27988 79551550744 4.060 2011     2600      2     3.027
## 30656 84856822033 3.605 2012     2600      1     2.575
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##       Min        1Q    Median        3Q       Max
## -1.283460 -0.501139  0.000542  0.495542  3.663439
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)          1.2116     0.0257   47.15  < 2e-16 ***
## FirstAuthorFemale1  -0.0572     0.0177   -3.23   0.0012 **
## Year1997            -0.0387     0.0414   -0.94   0.3498
## Year1998             0.0496     0.0863    0.58   0.5652
## Year1999            -0.0445     0.0345   -1.29   0.1976
## Year2000             0.1290     0.0739    1.75   0.0809 .
## Year2001            -0.1052     0.0361   -2.91   0.0036 **
## Year2002            -0.0173     0.0363   -0.48   0.6345
## Year2003            -0.0972     0.0342   -2.84   0.0045 **
## Year2004            -0.0729     0.0345   -2.11   0.0345 *
## Year2005            -0.0709     0.0349   -2.03   0.0423 *
## Year2006            -0.1364     0.0335   -4.07  4.8e-05 ***
## Year2007            -0.1697     0.0325   -5.21  1.9e-07 ***
## Year2008            -0.1451     0.0323   -4.49  7.2e-06 ***
## Year2009            -0.1949     0.0333   -5.85  5.1e-09 ***
## Year2010            -0.2165     0.0337   -6.43  1.3e-10 ***
## Year2011            -0.1811     0.0344   -5.27  1.4e-07 ***
## Year2012            -0.1446     0.0354   -4.09  4.4e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.752
## Multiple R-squared:  0.00958,    Adjusted R-squared:  0.00848
## Convergence in 10 IRWLS iterations
##
```

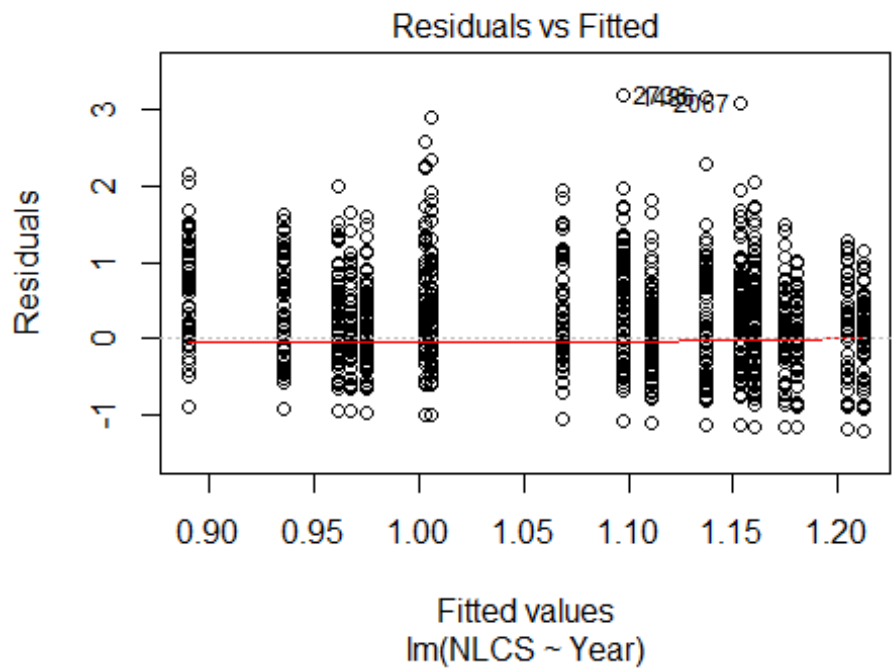
```
## Robustness weights:
##  observation 9699 is an outlier with |weight| = 0 ( < 6.5e-06);
##  1271 weights are ~= 1. The remaining 14041 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0068  0.8570  0.9510  0.9160  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          6.53e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##                   GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.008  1           1.004
## Year             1.008 16           1.000
```



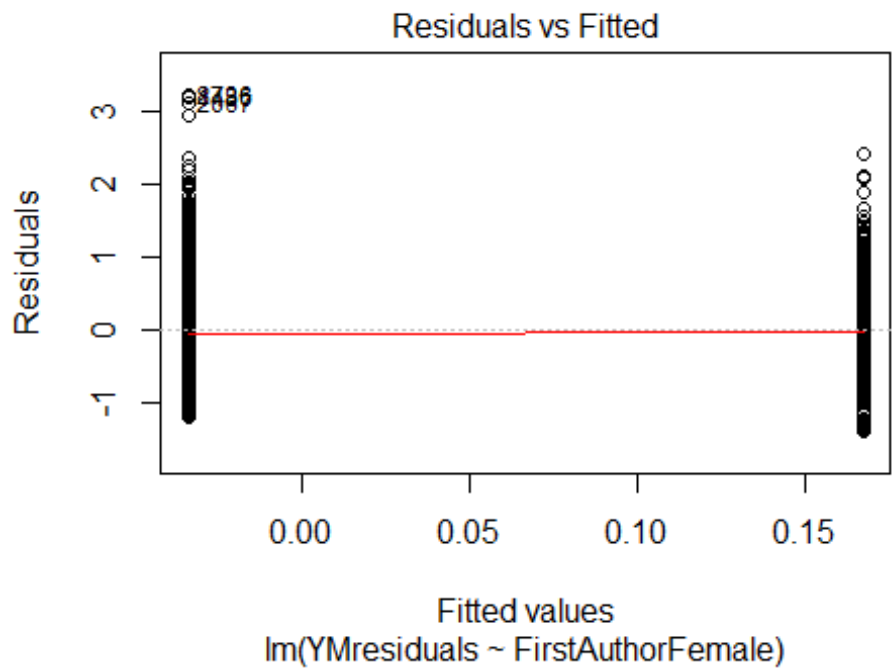
```
## [1] "List of  18 outliers with residuals above  2.5"
##          ScopusId  NLCS Year OneField Fields residuals
## 4858   0033239662 4.280 1999     2600      2     3.111
## 5080   0001259111 3.825 1999     1804      3     2.690
## 7466   0036020892 3.891 2002     2600      2     2.695
## 9606   0037623983 3.818 2003     1700      2     2.701
## 14730 33745604236 4.326 2006     2600      2     3.249
## 15035 33646365077 3.822 2006     2600      2     2.745
## 15258 29144523061 3.760 2006     1700      7     2.683
## 18786 55649115527 3.676 2008     2600      3     2.607
## 18799 57349174008 4.128 2008     2600      3     3.059
## 18854 52149117439 3.641 2008     1700      2     2.572
## 19905 42649140570 4.730 2008     2600      1     3.661
## 22536 65549134300 3.585 2009     2200      3     2.566
## 22621 70049091784 3.562 2009     2600      2     2.543
## 23378 84969334819 4.392 2009     2600      2     3.373
## 24195 77957933737 3.515 2010     2600      1     2.518
## 26821 80053974840 3.721 2011     1700      2     2.688
## 27988 79551550744 4.060 2011     2600      2     3.027
## 30656 84856822033 3.605 2012     2600      1     2.575
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##       Min        1Q    Median        3Q       Max
## -1.280612 -0.500878  0.000172  0.495049  3.663030
##
## Coefficients:
##                   Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.2123     0.0257   47.13  < 2e-16 ***
## LastAuthorFemale1  -0.0592     0.0177   -3.34  0.00085 ***
## Year1997           -0.0380     0.0414   -0.92  0.35928
## Year1998            0.0509     0.0861    0.59  0.55396
## Year1999           -0.0446     0.0346   -1.29  0.19723
## Year2000            0.1275     0.0740    1.72  0.08496 .
## Year2001           -0.1059     0.0361   -2.93  0.00340 **
## Year2002           -0.0177     0.0363   -0.49  0.62678
## Year2003           -0.0974     0.0342   -2.84  0.00447 **
## Year2004           -0.0736     0.0345   -2.13  0.03299 *
## Year2005           -0.0717     0.0349   -2.05  0.04024 *
## Year2006           -0.1372     0.0336   -4.09  4.3e-05 ***
## Year2007           -0.1695     0.0326   -5.21  1.9e-07 ***
## Year2008           -0.1453     0.0323   -4.50  6.9e-06 ***
## Year2009           -0.1957     0.0333   -5.87  4.4e-09 ***
## Year2010           -0.2174     0.0337   -6.46  1.1e-10 ***
## Year2011           -0.1814     0.0344   -5.28  1.3e-07 ***
## Year2012           -0.1456     0.0354   -4.12  3.8e-05 ***
```

```
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.752
## Multiple R-squared:  0.00963,    Adjusted R-squared:  0.00853
## Convergence in 10 IRWLS iterations
##
## Robustness weights:
##  observation 9699 is an outlier with |weight| = 0 ( < 6.5e-06);
##  1320 weights are ~= 1. The remaining 13992 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0068  0.8560  0.9510  0.9160  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          6.53e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Sample size for the above analysis:  15313"
## [1] ""
## [1] ""
## [1] "##################################"
## [1] "Analysis of AJSC 2601"
## [1] "##################################"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  132   94  100   88   85  105  133  139  118  149  143  167  213  168  184
## 2011 2012
##  177  187
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##   87   66   70   52   57   73   87   89   77  103   94  103  146  116  135
## 2011 2012
##  126  131
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##   75   62   64   49   55   64   77   80   69   93   86   95  129  106  128
## 2011 2012
```

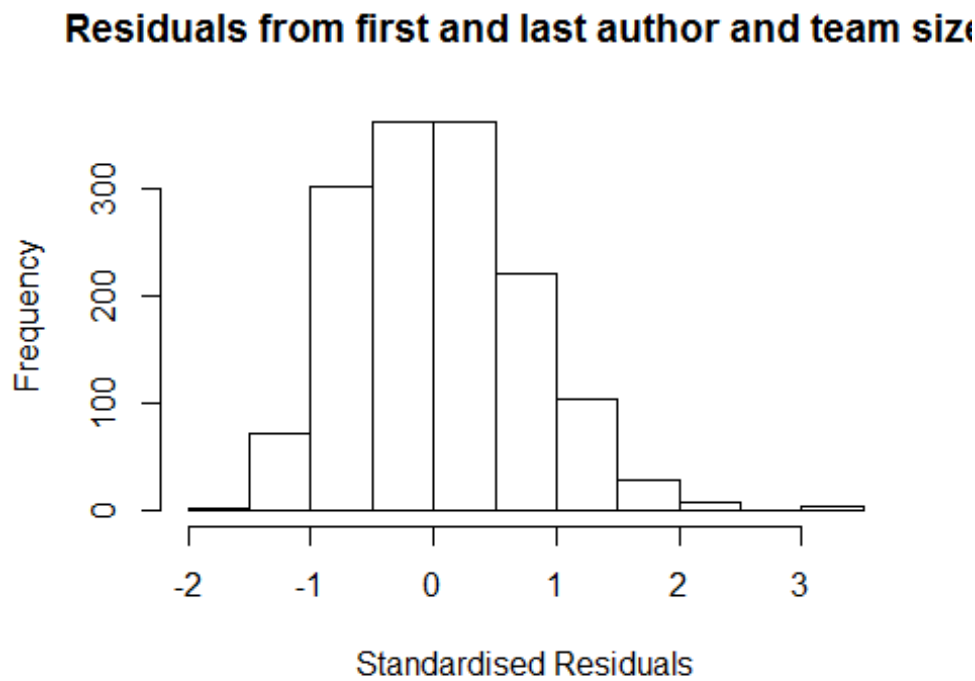
```
##  116  114
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
##  Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 50, df = 16, p-value = 3e-05
```




```
##
##  Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 4.4, df = 1, p-value = 0.04
```

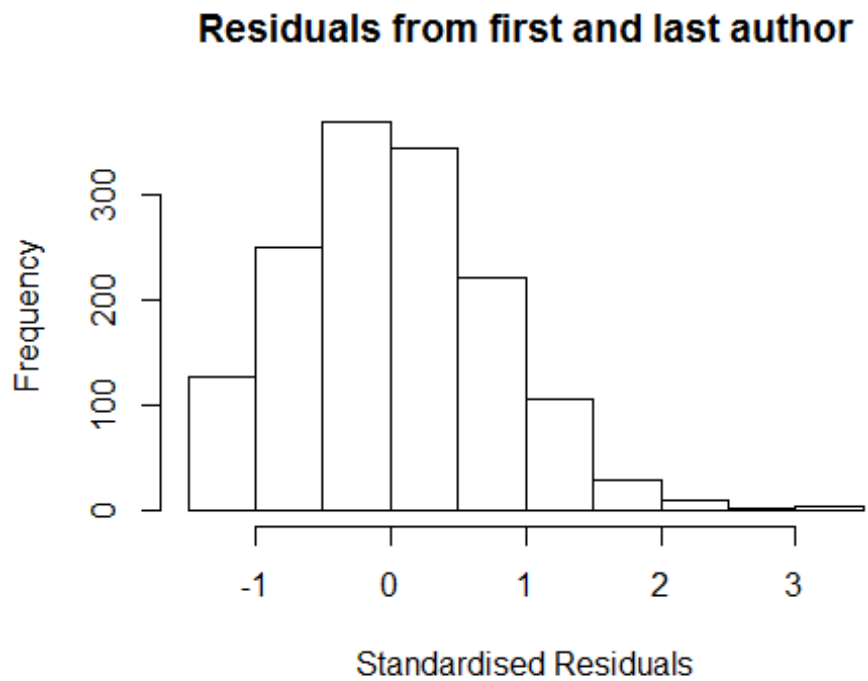



```
## [1] "Female first author team size 2018 geometric mean: 1.70939735716892"
## [1] "Male first author team size 2018 geometric mean: 1.43543107101746"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 1400, p-value = 0.2
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 1.63827604774522"
## [1] "Male last author team size 2018 geometric mean: 1.45426972332782"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 1300, p-value = 0.3
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.519  1           1.233
## LastAuthorFemale  1.504  1           1.227
## UniqueAuthors     1.201  4           1.023
## Year              1.240 16           1.007
```



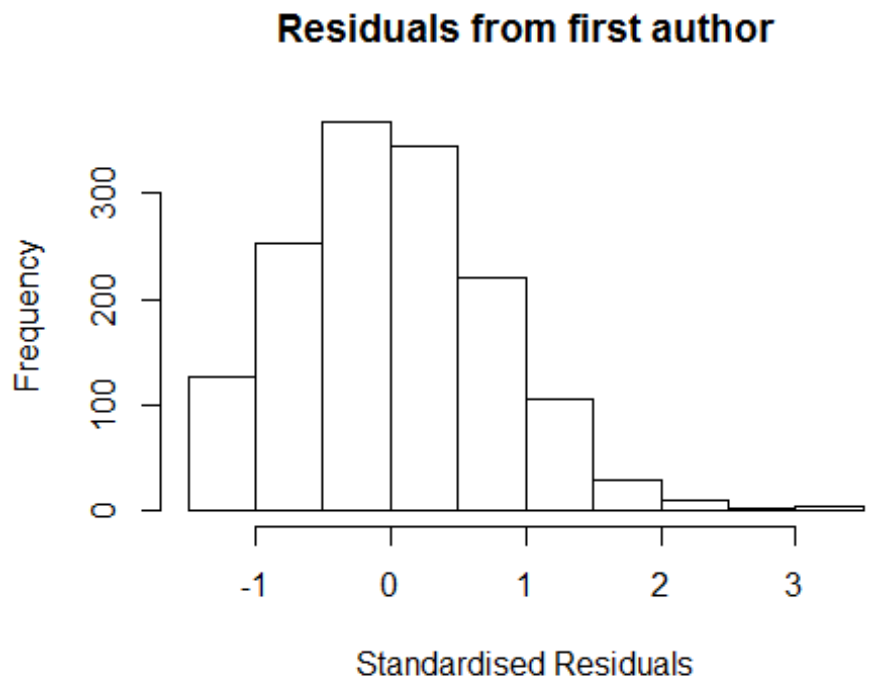

```
## [1] "List of  4 outliers with residuals above  2.5"
##         ScopusId  NLCS Year OneField Fields residuals
## 1486 33646521094 4.309 2006     2601      1     3.423
## 2067 68149179207 4.239 2009     2601      1     3.292
## 2510 79960797583 3.922 2011     2601      1     3.152
## 2736 84863007463 4.290 2012     2601      1     3.410
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
UniqueAuthors +
##     Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##     k.max = 1000))
##  \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.55946 -0.52012 -0.00564  0.49270  3.42303
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)           0.7526     0.0948    7.94  4.0e-15 ***
## FirstAuthorFemale1    0.1492     0.0686    2.18    0.030 *
## LastAuthorFemale1     0.0301     0.0692    0.44    0.663
## UniqueAuthors2        0.3257     0.0463    7.04  3.0e-12 ***
## UniqueAuthors3        0.4330     0.0742    5.84  6.5e-09 ***
## UniqueAuthors4        0.9130     0.1248    7.31  4.3e-13 ***
## UniqueAuthors5        0.6539     0.1297    5.04  5.2e-07 ***
```

```
## Year1997             0.0472     0.1417    0.33    0.739
## Year1998             0.2896     0.1241    2.33    0.020 *
## Year1999             0.2364     0.1390    1.70    0.089 .
## Year2000             0.2554     0.1168    2.19    0.029 *
## Year2001            -0.0260     0.1218   -0.21    0.831
## Year2002            -0.0077     0.1256   -0.06    0.951
## Year2003            -0.0398     0.1348   -0.30    0.768
## Year2004             0.2119     0.1239    1.71    0.088 .
## Year2005             0.0686     0.1130    0.61    0.544
## Year2006             0.1334     0.1372    0.97    0.331
## Year2007             0.1075     0.1209    0.89    0.374
## Year2008             0.2243     0.1186    1.89    0.059 .
## Year2009             0.1945     0.1185    1.64    0.101
## Year2010            -0.1061     0.1142   -0.93    0.353
## Year2011             0.0170     0.1168    0.15    0.884
## Year2012             0.1275     0.1218    1.05    0.295
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.742
## Multiple R-squared:  0.115,  Adjusted R-squared:  0.101
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  122 weights are ~= 1. The remaining 1340 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0008  0.8760  0.9470  0.9140  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          6.84e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.550  1           1.245
## LastAuthorFemale  1.510  1           1.229
## Year              1.073 16           1.002
```



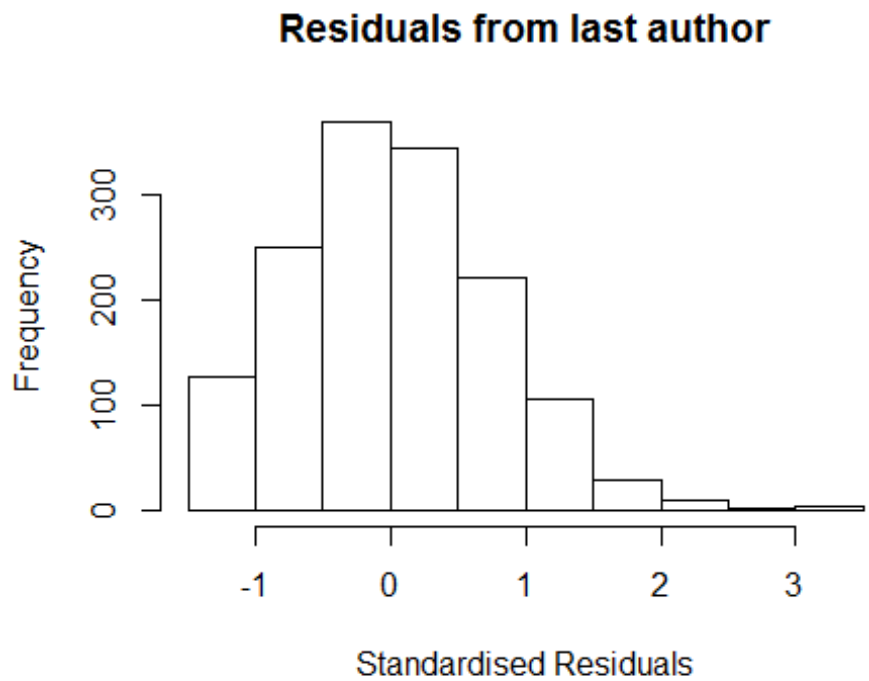

```
## [1] "List of  5 outliers with residuals above  2.5"
##         ScopusId  NLCS Year OneField Fields residuals
## 87   21344474293 3.593 1996     2601      2     2.528
## 1486 33646521094 4.309 2006     2601      1     3.294
## 2067 68149179207 4.239 2009     2601      1     3.167
## 2510 79960797583 3.922 2011     2601      1     3.033
## 2736 84863007463 4.290 2012     2601      1     3.253
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.4138 -0.5124 -0.0207  0.5037  3.2936
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         0.84658    0.10248    8.26  3.2e-16 ***
## FirstAuthorFemale1  0.21834    0.07365    2.96   0.0031 **
## LastAuthorFemale1   0.03469    0.07669    0.45   0.6511
## Year1997            0.12804    0.14843    0.86   0.3885
## Year1998            0.30794    0.13088    2.35   0.0188 *
## Year1999            0.31418    0.15363    2.05   0.0410 *
## Year2000            0.30532    0.12393    2.46   0.0139 *
## Year2001            0.00571    0.13212    0.04   0.9656
```

```
## Year2002            0.03934    0.13320    0.30   0.7678
## Year2003            0.01385    0.13504    0.10   0.9183
## Year2004            0.30472    0.12744    2.39   0.0169 *
## Year2005            0.08178    0.11952    0.68   0.4940
## Year2006            0.16879    0.14321    1.18   0.2387
## Year2007            0.19215    0.12741    1.51   0.1318
## Year2008            0.25067    0.12561    2.00   0.0462 *
## Year2009            0.22562    0.12557    1.80   0.0726 .
## Year2010           -0.05971    0.12327   -0.48   0.6282
## Year2011            0.04284    0.12593    0.34   0.7337
## Year2012            0.19066    0.13169    1.45   0.1479
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.769
## Multiple R-squared:  0.0396, Adjusted R-squared:  0.0276
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
##  123 weights are ~= 1. The remaining 1339 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0269  0.8760  0.9510  0.9130  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          6.84e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.058  1           1.028
## Year              1.058 16           1.002
```




```
## [1] "List of  5 outliers with residuals above  2.5"
##         ScopusId  NLCS Year OneField Fields residuals
## 87   21344474293 3.593 1996     2601      2     2.528
## 1486 33646521094 4.309 2006     2601      1     3.294
## 2067 68149179207 4.239 2009     2601      1     3.167
## 2510 79960797583 3.922 2011     2601      1     3.033
## 2736 84863007463 4.290 2012     2601      1     3.253
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.4014 -0.5143 -0.0235  0.5023  3.2920
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         0.84679    0.10254    8.26  3.3e-16 ***
## FirstAuthorFemale1  0.23868    0.06070    3.93  8.8e-05 ***
## Year1997            0.12974    0.14901    0.87    0.384
## Year1998            0.31129    0.13103    2.38    0.018 *
## Year1999            0.31594    0.15384    2.05    0.040 *
## Year2000            0.30464    0.12398    2.46    0.014 *
## Year2001            0.00782    0.13216    0.06    0.953
## Year2002            0.04047    0.13314    0.30    0.761
```

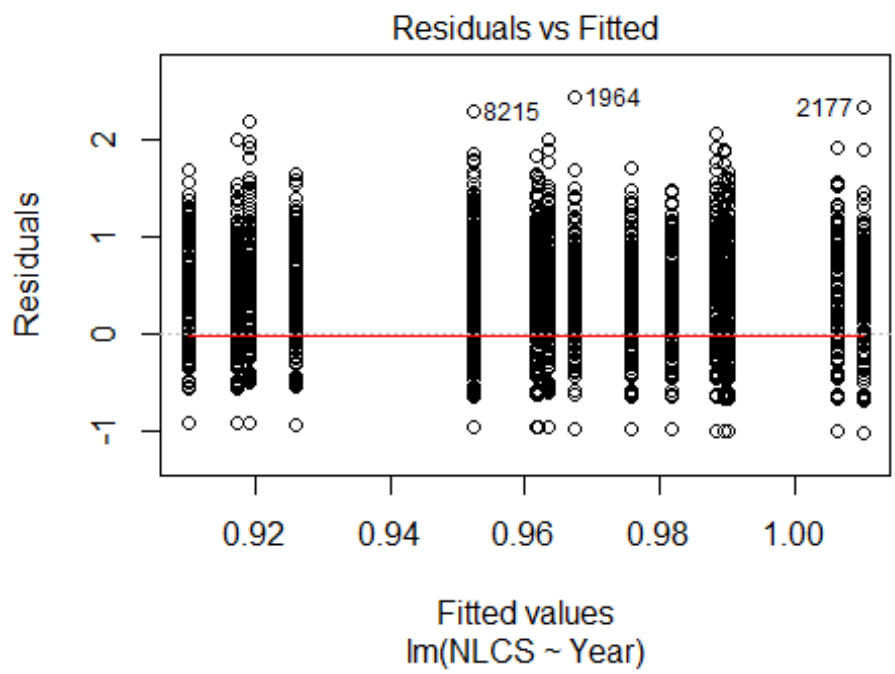
```
## Year2003            0.01499    0.13523    0.11    0.912
## Year2004            0.30765    0.12759    2.41    0.016 *
## Year2005            0.08353    0.11964    0.70    0.485
## Year2006            0.17021    0.14330    1.19    0.235
## Year2007            0.19531    0.12755    1.53    0.126
## Year2008            0.25251    0.12558    2.01    0.045 *
## Year2009            0.22544    0.12565    1.79    0.073 .
## Year2010           -0.05832    0.12336   -0.47    0.636
## Year2011            0.04360    0.12599    0.35    0.729
## Year2012            0.19239    0.13181    1.46    0.145
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.768
## Multiple R-squared:  0.0394, Adjusted R-squared:  0.0281
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
##  122 weights are ~= 1. The remaining 1340 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0268  0.8780  0.9520  0.9130  0.9860  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          6.84e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                    psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##                   GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.035  1           1.017
## Year             1.035 16           1.001
```



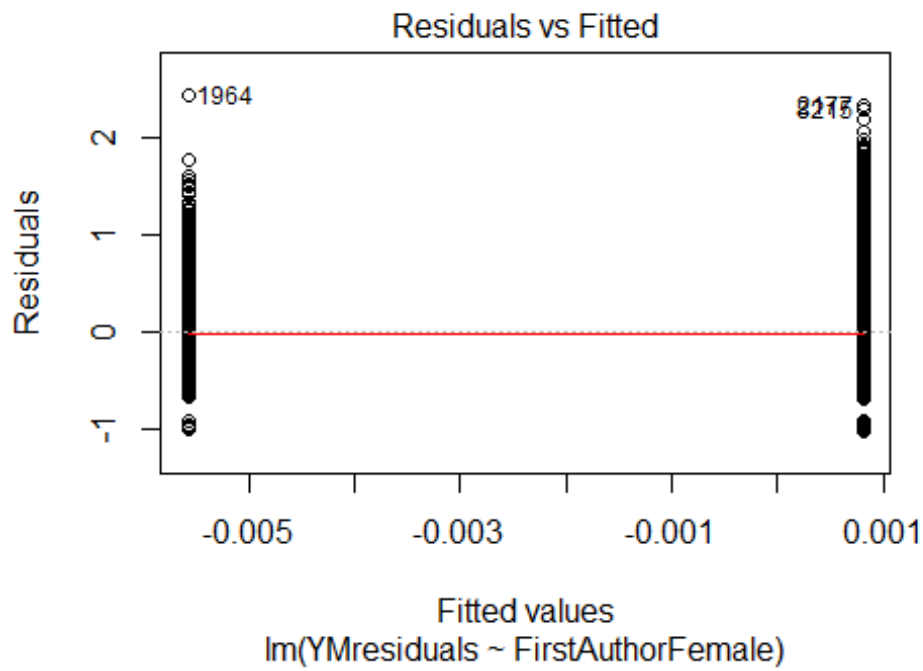

```
## [1] "List of  5 outliers with residuals above  2.5"
##         ScopusId  NLCS Year OneField Fields residuals
## 87   21344474293 3.593 1996     2601      2     2.528
## 1486 33646521094 4.309 2006     2601      1     3.294
## 2067 68149179207 4.239 2009     2601      1     3.167
## 2510 79960797583 3.922 2011     2601      1     3.033
## 2736 84863007463 4.290 2012     2601      1     3.253
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.3613 -0.5254 -0.0163  0.5006  3.2807
##
## Coefficients:
##                   Estimate Std. Error t value Pr(>|t|)
## (Intercept)        0.87068    0.10137    8.59   <2e-16 ***
## LastAuthorFemale1  0.17293    0.06199    2.79   0.0053 **
## Year1997           0.12259    0.14731    0.83   0.4054
## Year1998           0.29566    0.12937    2.29   0.0224 *
## Year1999           0.29424    0.15384    1.91   0.0560 .
## Year2000           0.31766    0.12358    2.57   0.0103 *
## Year2001          -0.01225    0.13023   -0.09   0.9251
## Year2002           0.03600    0.13285    0.27   0.7864
```

```
## Year2003           0.00169    0.13449    0.01   0.9900
## Year2004           0.28397    0.12598    2.25   0.0243 *
## Year2005           0.07697    0.11896    0.65   0.5177
## Year2006           0.15764    0.14185    1.11   0.2666
## Year2007           0.18134    0.12650    1.43   0.1519
## Year2008           0.24492    0.12521    1.96   0.0507 .
## Year2009           0.22723    0.12543    1.81   0.0703 .
## Year2010          -0.07365    0.12301   -0.60   0.5494
## Year2011           0.03302    0.12532    0.26   0.7922
## Year2012           0.17059    0.13084    1.30   0.1925
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.774
## Multiple R-squared:  0.0325, Adjusted R-squared:  0.0211
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
##  121 weights are ~= 1. The remaining 1341 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0333  0.8790  0.9510  0.9140  0.9860  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          6.84e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Sample size for the above analysis:  1462"
## [1] ""
## [1] ""
## [1] "##################################"
## [1] "Analysis of AJSC 2602"
## [1] "##################################"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  556  564  569  512  588  604  532  534  592  596  605  642  662  748  708
## 2011 2012
##  695  760
```

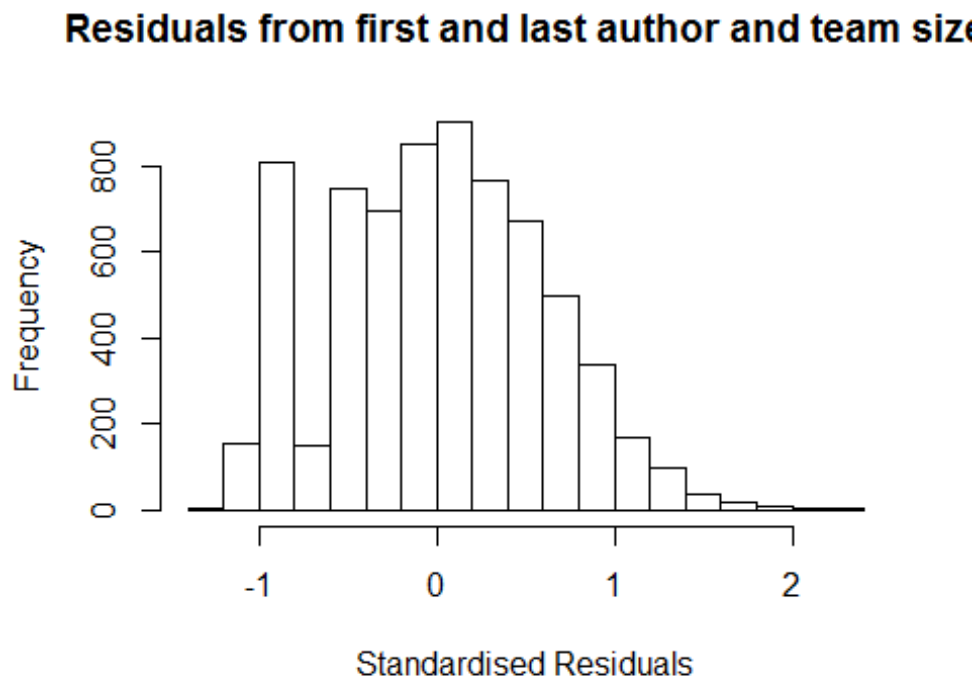
```
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  387  382  389  361  407  412  382  385  442  440  436  479  477  543  502
## 2011 2012
##  483  569
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  363  356  363  343  378  392  358  364  409  399  405  453  438  502  461
## 2011 2012
##  437  503
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
##  Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 28, df = 16, p-value = 0.03
```




```
##
##  Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 0.059, df = 1, p-value = 0.8
```

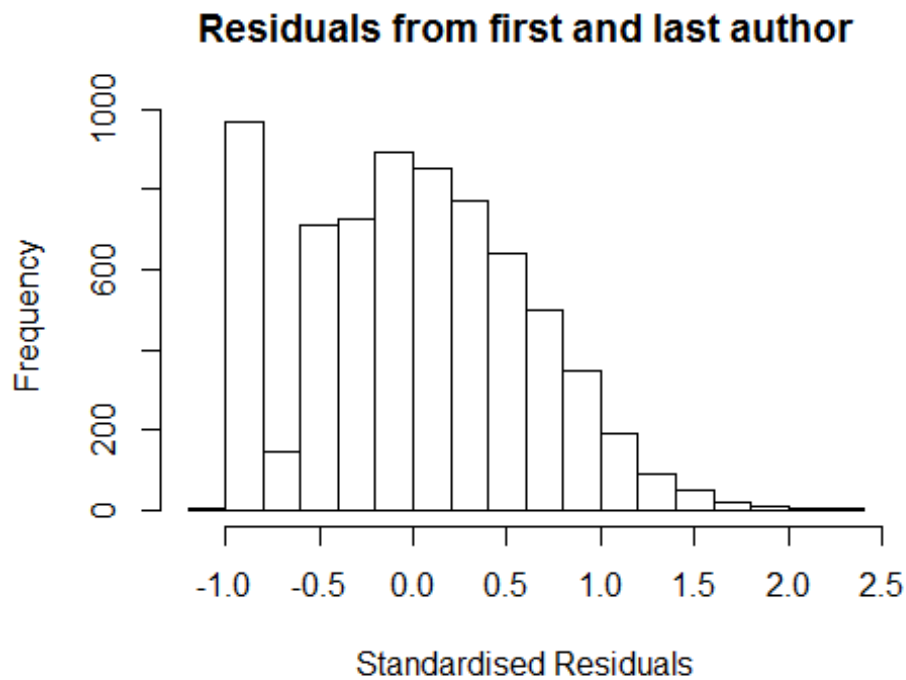



```
## [1] "Female first author team size 2018 geometric mean: 1.78189972704166"
## [1] "Male first author team size 2018 geometric mean: 1.38257396204995"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 15000, p-value = 1e-04
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 1.78408247671274"
## [1] "Male last author team size 2018 geometric mean: 1.38750970297835"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 14000, p-value = 4e-04
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##                     GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.719  1           1.311
## LastAuthorFemale  1.721  1           1.312
## UniqueAuthors     1.052  4           1.006
## Year              1.063 16           1.002
```



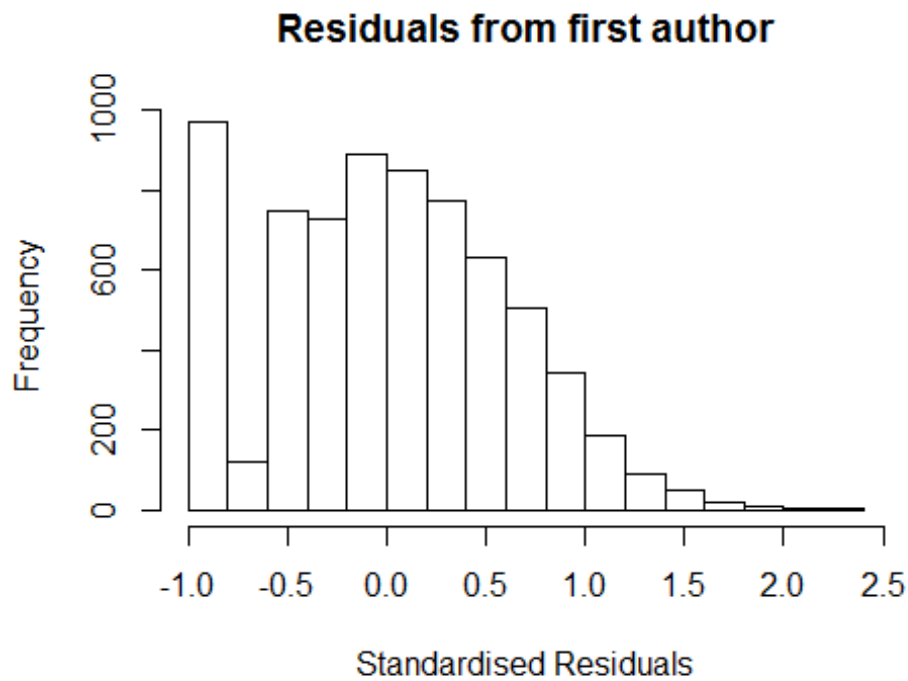

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
UniqueAuthors +
##     Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##     k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.2562 -0.4268  0.0171  0.4242  2.2591
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)          0.9553     0.0372   25.70  < 2e-16 ***
## FirstAuthorFemale1   0.0362     0.0302    1.20  0.22993
## LastAuthorFemale1   -0.0640     0.0293   -2.18  0.02898 *
## UniqueAuthors2       0.1445     0.0182    7.93  2.6e-15 ***
## UniqueAuthors3       0.2646     0.0319    8.30  < 2e-16 ***
## UniqueAuthors4       0.1262     0.0667    1.89  0.05854 .
## UniqueAuthors5       0.3180     0.0919    3.46  0.00054 ***
## Year1997            -0.0337     0.0505   -0.67  0.50430
## Year1998            -0.0556     0.0495   -1.12  0.26092
## Year1999            -0.0079     0.0500   -0.16  0.87450
```

```
## Year2000            -0.0477     0.0493   -0.97  0.33316
## Year2001            -0.0181     0.0486   -0.37  0.70972
## Year2002            -0.0652     0.0502   -1.30  0.19379
## Year2003            -0.0699     0.0491   -1.42  0.15435
## Year2004            -0.0783     0.0471   -1.66  0.09616 .
## Year2005            -0.0394     0.0468   -0.84  0.39936
## Year2006            -0.0555     0.0480   -1.16  0.24785
## Year2007            -0.0840     0.0473   -1.78  0.07585 .
## Year2008            -0.1043     0.0470   -2.22  0.02648 *
## Year2009            -0.1035     0.0452   -2.29  0.02209 *
## Year2010            -0.0843     0.0484   -1.74  0.08180 .
## Year2011            -0.1178     0.0481   -2.45  0.01427 *
## Year2012            -0.1418     0.0483   -2.94  0.00332 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.634
## Multiple R-squared:  0.0215, Adjusted R-squared:  0.0184
## Convergence in 10 IRWLS iterations
##
## Robustness weights:
##  560 weights are ~= 1. The remaining 6364 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.177   0.855   0.951   0.916   0.985   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.44e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.723  1           1.313
## LastAuthorFemale  1.722  1           1.312
## Year              1.016 16           1.000
```



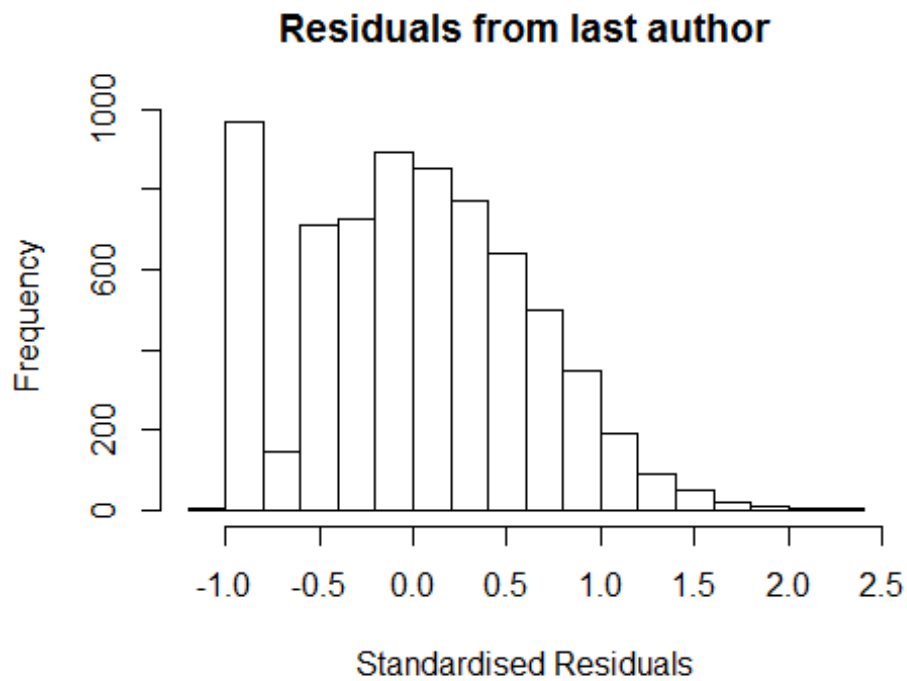

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.01666 -0.43365  0.00413  0.43479  2.35635
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)          0.99214    0.03736   26.55   <2e-16 ***
## FirstAuthorFemale1   0.02453    0.03046    0.81    0.421
## LastAuthorFemale1   -0.04301    0.02950   -1.46    0.145
## Year1997            -0.02327    0.05120   -0.45    0.649
## Year1998            -0.04222    0.04987   -0.85    0.397
## Year1999             0.00251    0.05031    0.05    0.960
## Year2000            -0.03991    0.04973   -0.80    0.422
## Year2001            -0.00793    0.04896   -0.16    0.871
## Year2002            -0.04974    0.05060   -0.98    0.326
## Year2003            -0.06105    0.04959   -1.23    0.218
## Year2004            -0.06093    0.04744   -1.28    0.199
## Year2005            -0.02421    0.04738   -0.51    0.609
```

```
## Year2006           -0.02919    0.04855   -0.60    0.548
## Year2007           -0.06571    0.04768   -1.38    0.168
## Year2008           -0.08981    0.04755   -1.89    0.059 .
## Year2009           -0.08457    0.04556   -1.86    0.063 .
## Year2010           -0.06171    0.04872   -1.27    0.205
## Year2011           -0.08588    0.04823   -1.78    0.075 .
## Year2012           -0.11259    0.04855   -2.32    0.020 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.64
## Multiple R-squared:  0.00308,    Adjusted R-squared:  0.00048
## Convergence in 10 IRWLS iterations
##
## Robustness weights:
##  583 weights are ~= 1. The remaining 6341 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.146   0.850   0.951   0.916   0.986   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.44e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.008  1           1.004
## Year              1.008 16           1.000
```




```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.99167 -0.43904  0.00293  0.43654  2.35933
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         0.98996    0.03737   26.49   <2e-16 ***
## FirstAuthorFemale1 -0.00433    0.02339   -0.19    0.853
## Year1997           -0.02331    0.05121   -0.46    0.649
## Year1998           -0.04253    0.04988   -0.85    0.394
## Year1999            0.00171    0.05034    0.03    0.973
## Year2000           -0.03964    0.04983   -0.80    0.426
## Year2001           -0.00750    0.04898   -0.15    0.878
## Year2002           -0.04931    0.05060   -0.97    0.330
## Year2003           -0.06097    0.04961   -1.23    0.219
## Year2004           -0.05980    0.04747   -1.26    0.208
## Year2005           -0.02375    0.04737   -0.50    0.616
## Year2006           -0.02963    0.04851   -0.61    0.541
```

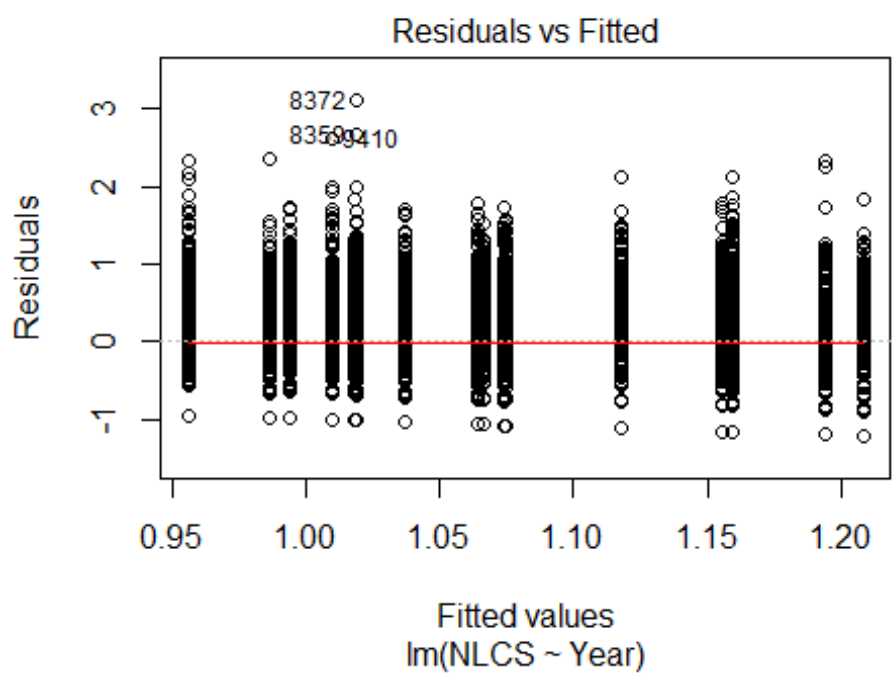
```
## Year2007           -0.06605    0.04771   -1.38    0.166
## Year2008           -0.08943    0.04756   -1.88    0.060 .
## Year2009           -0.08485    0.04559   -1.86    0.063 .
## Year2010           -0.06189    0.04874   -1.27    0.204
## Year2011           -0.08692    0.04822   -1.80    0.072 .
## Year2012           -0.11244    0.04857   -2.31    0.021 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.639
## Multiple R-squared:  0.00275,    Adjusted R-squared:  0.000299
## Convergence in 10 IRWLS iterations
##
## Robustness weights:
##  585 weights are ~= 1. The remaining 6339 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.144   0.849   0.951   0.916   0.987   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.44e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##                   GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.008  1           1.004
## Year             1.008 16           1.000
```



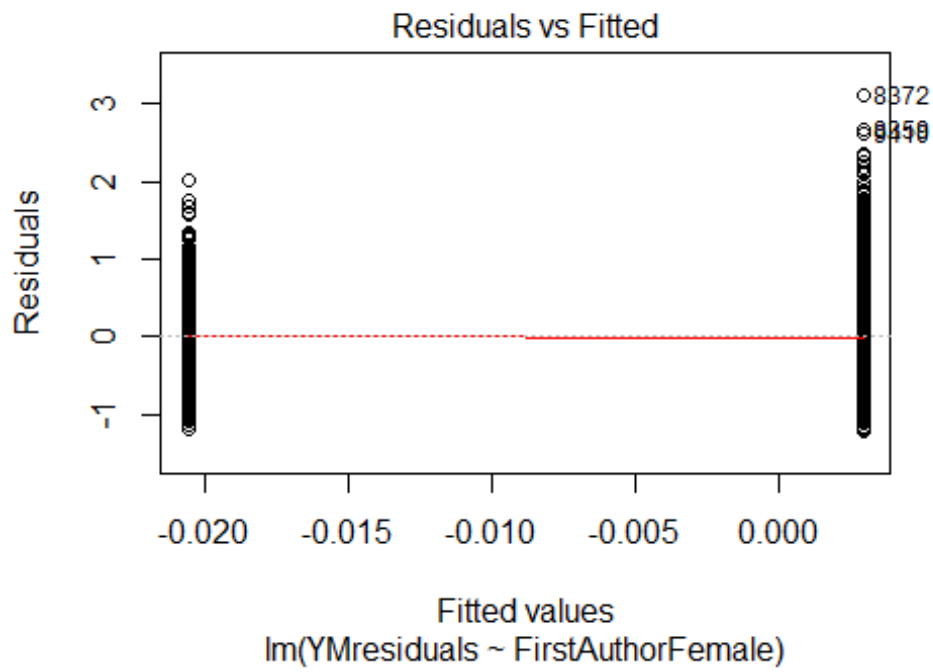

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.99604 -0.43617  0.00348  0.43398  2.35496
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         0.99293    0.03733   26.60   <2e-16 ***
## LastAuthorFemale1 -0.02753    0.02254   -1.22    0.222
## Year1997          -0.02341    0.05119   -0.46    0.648
## Year1998          -0.04203    0.04987   -0.84    0.399
## Year1999           0.00311    0.05030    0.06    0.951
## Year2000          -0.03972    0.04976   -0.80    0.425
## Year2001          -0.00788    0.04895   -0.16    0.872
## Year2002          -0.04924    0.05058   -0.97    0.330
## Year2003          -0.06085    0.04958   -1.23    0.220
## Year2004          -0.06044    0.04744   -1.27    0.203
## Year2005          -0.02411    0.04735   -0.51    0.611
## Year2006          -0.02942    0.04852   -0.61    0.544
```

```
## Year2007          -0.06563    0.04766   -1.38    0.169
## Year2008          -0.08927    0.04753   -1.88    0.060 .
## Year2009          -0.08451    0.04554   -1.86    0.064 .
## Year2010          -0.06123    0.04871   -1.26    0.209
## Year2011          -0.08603    0.04822   -1.78    0.074 .
## Year2012          -0.11224    0.04854   -2.31    0.021 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.639
## Multiple R-squared:  0.00298,    Adjusted R-squared:  0.000526
## Convergence in 10 IRWLS iterations
##
## Robustness weights:
##  590 weights are ~= 1. The remaining 6334 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.146   0.848   0.950   0.916   0.987   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.44e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Sample size for the above analysis:  6924"
## [1] ""
## [1] ""
## [1] "####################################"
## [1] "Analysis of AJSC 2603"
## [1] "####################################"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  594  646  506  470  555  525  533  467  557  579  647  688  737  907  828
## 2011 2012
##  773  836
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  330  362  227  245  287  252  300  266  299  350  377  419  477  551  496
## 2011 2012
```

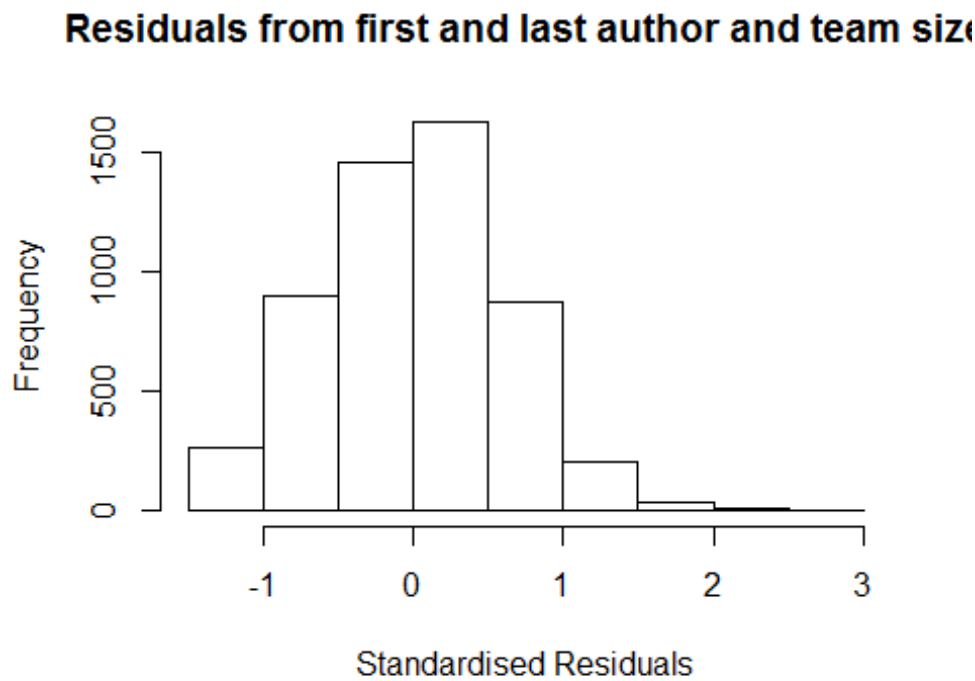
```
##  488  494
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  288  308  204  200  251  217  266  240  263  309  319  352  420  470  413
## 2011 2012
##  408  419
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
##  Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 41, df = 16, p-value = 6e-04
```




```
##
##  Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 1.5, df = 1, p-value = 0.2
```

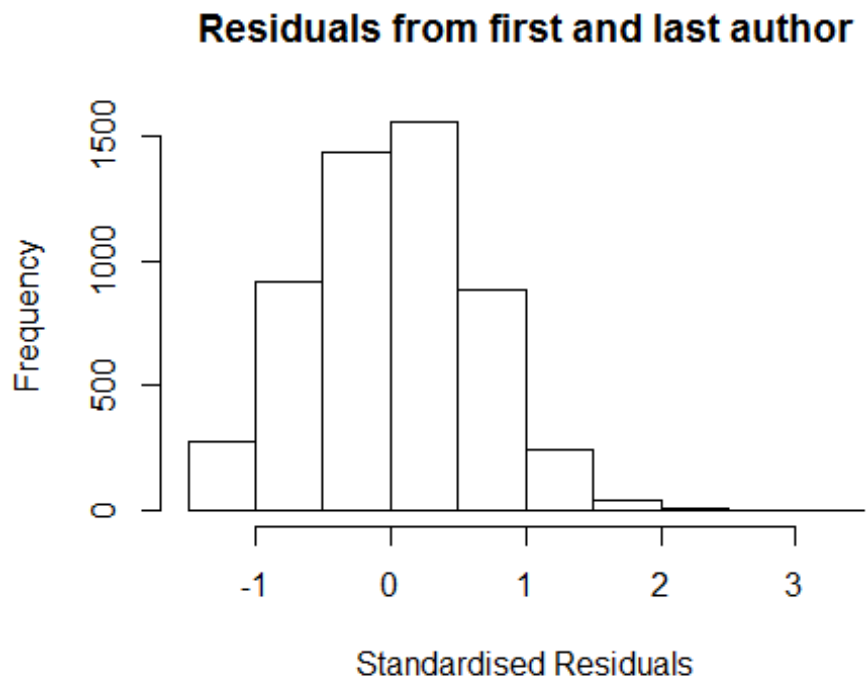



```
## [1] "Female first author team size 2018 geometric mean: 2.09128893329362"
## [1] "Male first author team size 2018 geometric mean: 1.45820666007821"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 8800, p-value = 3e-04
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 1.87690802708404"
## [1] "Male last author team size 2018 geometric mean: 1.48339818837811"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 7000, p-value = 0.02
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##                     GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.554  1           1.247
## LastAuthorFemale  1.540  1           1.241
## UniqueAuthors     1.082  4           1.010
## Year              1.085 16           1.003
```



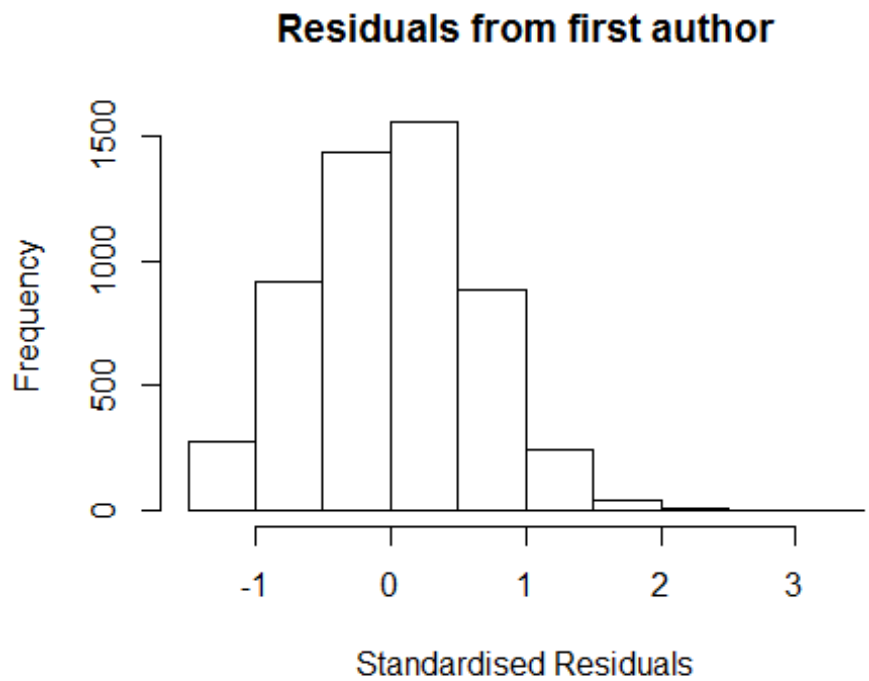

```
## [1] "List of  2 outliers with residuals above  2.5"
##         ScopusId  NLCS Year OneField Fields residuals
## 8372 57349174008 4.128 2008     2600      3     2.908
## 9410 71049116435 3.639 2009     1703      4     2.525
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
UniqueAuthors +
##     Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##     k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.3918 -0.4273  0.0207  0.4225  2.9077
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.03025    0.04288   24.03  < 2e-16 ***
## FirstAuthorFemale1 -0.03777    0.03171   -1.19   0.2337
## LastAuthorFemale1  -0.01139    0.03240   -0.35   0.7252
## UniqueAuthors2      0.22842    0.01982   11.53  < 2e-16 ***
## UniqueAuthors3      0.33482    0.03364    9.95  < 2e-16 ***
## UniqueAuthors4      0.32591    0.07507    4.34  1.4e-05 ***
## UniqueAuthors5      0.53287    0.10037    5.31  1.1e-07 ***
## Year1997            0.02670    0.06111    0.44   0.6622
## Year1998            0.08274    0.06414    1.29   0.1971
```

```
## Year1999            0.07279    0.06073    1.20   0.2308
## Year2000            0.00940    0.05797    0.16   0.8712
## Year2001            0.00904    0.05665    0.16   0.8732
## Year2002           -0.06570    0.05722   -1.15   0.2509
## Year2003           -0.08013    0.05609   -1.43   0.1532
## Year2004           -0.07346    0.05579   -1.32   0.1880
## Year2005           -0.10433    0.05332   -1.96   0.0504 .
## Year2006           -0.13700    0.05444   -2.52   0.0119 *
## Year2007           -0.15866    0.05385   -2.95   0.0032 **
## Year2008           -0.14478    0.05223   -2.77   0.0056 **
## Year2009           -0.14493    0.05063   -2.86   0.0042 **
## Year2010           -0.13925    0.05172   -2.69   0.0071 **
## Year2011           -0.18750    0.05294   -3.54   0.0004 ***
## Year2012           -0.22925    0.05446   -4.21  2.6e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.627
## Multiple R-squared:  0.057,  Adjusted R-squared:  0.0531
## Convergence in 10 IRWLS iterations
##
## Robustness weights:
##  452 weights are ~= 1. The remaining 4895 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0004  0.8670  0.9500  0.9130  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.87e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.513  1           1.230
## LastAuthorFemale  1.519  1           1.232
## Year              1.022 16           1.001
```



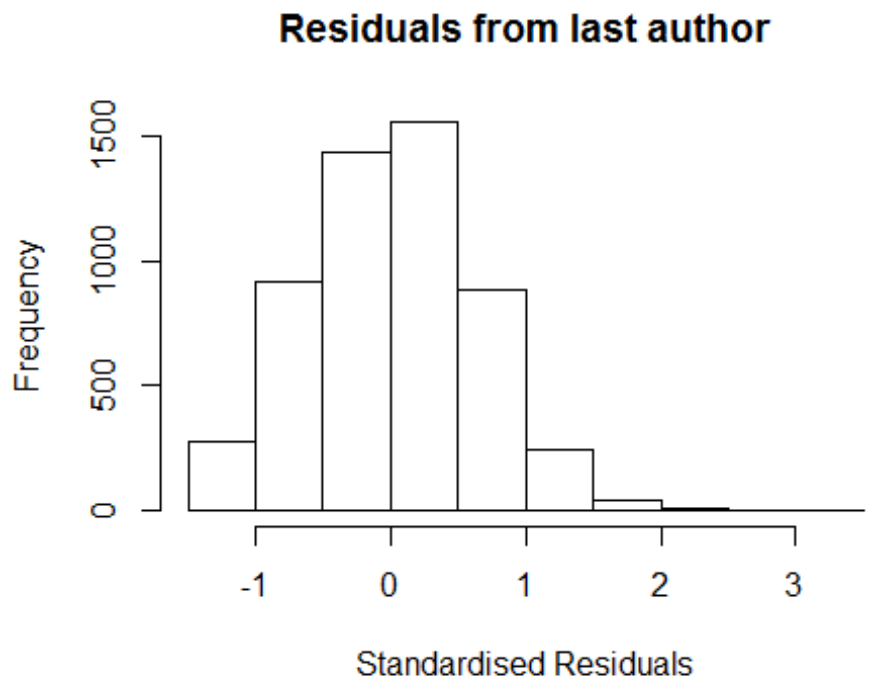

```
## [1] "List of  3 outliers with residuals above  2.5"
##         ScopusId  NLCS Year OneField Fields residuals
## 8359 55649115527 3.676 2008     2600      3     2.697
## 8372 57349174008 4.128 2008     2600      3     3.149
## 9410 71049116435 3.639 2009     1703      4     2.653
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.1854 -0.4276  0.0181  0.4311  3.1491
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.11286    0.04311   25.81  < 2e-16 ***
## FirstAuthorFemale1 -0.00212    0.03230   -0.07  0.94761
## LastAuthorFemale1  -0.01042    0.03294   -0.32  0.75183
## Year1997            0.01917    0.06187    0.31  0.75664
## Year1998            0.07256    0.06506    1.12  0.26478
## Year1999            0.06159    0.06113    1.01  0.31370
## Year2000            0.01898    0.05954    0.32  0.74994
## Year2001            0.00666    0.05817    0.11  0.90892
## Year2002           -0.04115    0.05945   -0.69  0.48891
## Year2003           -0.07677    0.05683   -1.35  0.17682
```

```
## Year2004           -0.07598    0.05685   -1.34  0.18144
## Year2005           -0.08555    0.05457   -1.57  0.11704
## Year2006           -0.11646    0.05535   -2.10  0.03543 *
## Year2007           -0.15936    0.05463   -2.92  0.00355 **
## Year2008           -0.13394    0.05298   -2.53  0.01149 *
## Year2009           -0.12700    0.05175   -2.45  0.01416 *
## Year2010           -0.10993    0.05274   -2.08  0.03719 *
## Year2011           -0.16000    0.05372   -2.98  0.00291 **
## Year2012           -0.20380    0.05508   -3.70  0.00022 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.642
## Multiple R-squared:  0.0151, Adjusted R-squared:  0.0118
## Convergence in 10 IRWLS iterations
##
## Robustness weights:
##  observation 3271 is an outlier with |weight| = 0 ( < 1.9e-05);
##  456 weights are ~= 1. The remaining 4890 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0385  0.8720  0.9500  0.9140  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.87e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                    psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.009  1           1.004
## Year              1.009 16           1.000
```




```
## [1] "List of  3 outliers with residuals above  2.5"
##         ScopusId  NLCS Year OneField Fields residuals
## 8359 55649115527 3.676 2008     2600      3     2.697
## 8372 57349174008 4.128 2008     2600      3     3.149
## 9410 71049116435 3.639 2009     1703      4     2.653
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.1847 -0.4273  0.0182  0.4313  3.1496
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.11254    0.04310   25.81  < 2e-16 ***
## FirstAuthorFemale1 -0.00818    0.02640   -0.31  0.75668
## Year1997            0.01925    0.06186    0.31  0.75565
## Year1998            0.07218    0.06502    1.11  0.26702
## Year1999            0.06123    0.06112    1.00  0.31643
## Year2000            0.01903    0.05954    0.32  0.74933
## Year2001            0.00672    0.05815    0.12  0.90801
## Year2002           -0.04103    0.05945   -0.69  0.49016
## Year2003           -0.07697    0.05683   -1.35  0.17565
## Year2004           -0.07610    0.05684   -1.34  0.18070
```

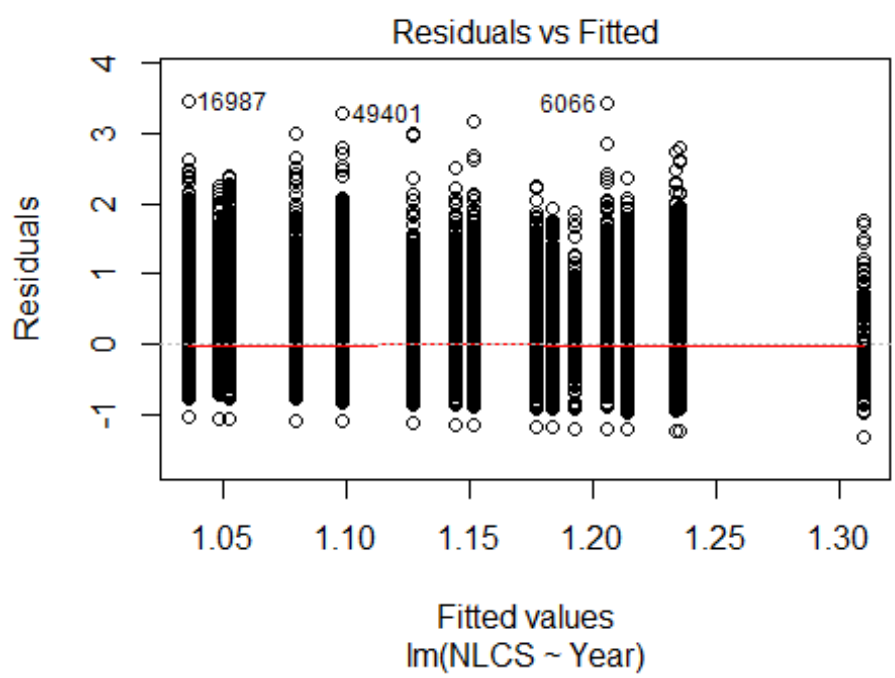
```
## Year2005           -0.08582    0.05456   -1.57  0.11578
## Year2006           -0.11706    0.05525   -2.12  0.03416 *
## Year2007           -0.15957    0.05462   -2.92  0.00350 **
## Year2008           -0.13411    0.05297   -2.53  0.01138 *
## Year2009           -0.12716    0.05174   -2.46  0.01402 *
## Year2010           -0.11026    0.05272   -2.09  0.03653 *
## Year2011           -0.15979    0.05372   -2.97  0.00295 **
## Year2012           -0.20395    0.05507   -3.70  0.00021 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.642
## Multiple R-squared:  0.0151, Adjusted R-squared:  0.012
## Convergence in 10 IRWLS iterations
##
## Robustness weights:
##  observation 3271 is an outlier with |weight| = 0 ( < 1.9e-05);
##  454 weights are ~= 1. The remaining 4892 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0385  0.8720  0.9500  0.9140  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.87e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##                   GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.012  1           1.006
## Year             1.012 16           1.000
```



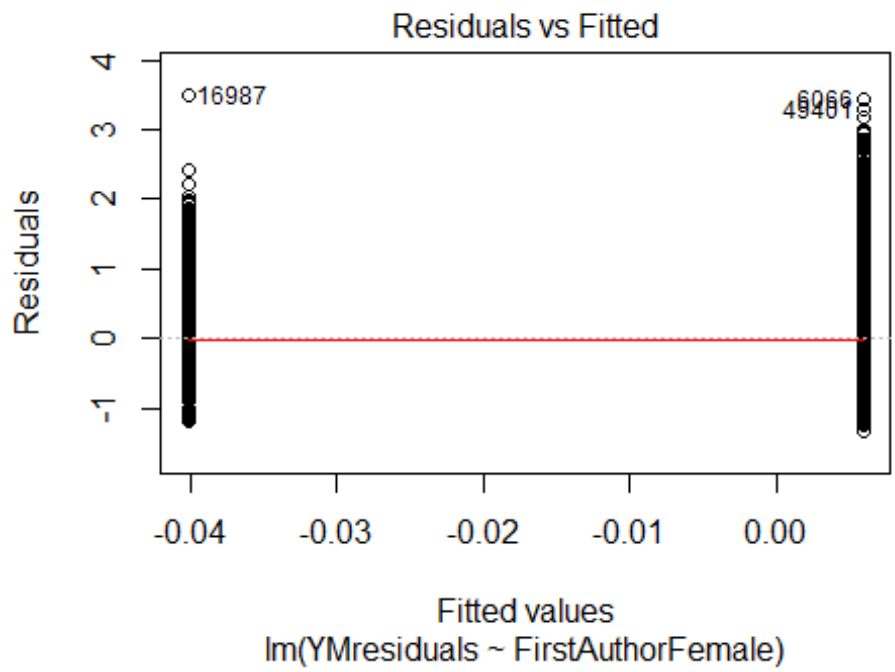

```
## [1] "List of  3 outliers with residuals above  2.5"
##         ScopusId  NLCS Year OneField Fields residuals
## 8359 55649115527 3.676 2008     2600      3     2.697
## 8372 57349174008 4.128 2008     2600      3     3.149
## 9410 71049116435 3.639 2009     1703      4     2.653
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.1853 -0.4283  0.0183  0.4312  3.1492
##
## Coefficients:
##                   Estimate Std. Error t value Pr(>|t|)
## (Intercept)        1.11277    0.04308   25.83  < 2e-16 ***
## LastAuthorFemale1 -0.01173    0.02692   -0.44  0.66305
## Year1997           0.01918    0.06186    0.31  0.75659
## Year1998           0.07258    0.06506    1.12  0.26464
## Year1999           0.06161    0.06113    1.01  0.31360
## Year2000           0.01894    0.05953    0.32  0.75034
## Year2001           0.00669    0.05817    0.12  0.90841
## Year2002          -0.04116    0.05944   -0.69  0.48875
## Year2003          -0.07679    0.05683   -1.35  0.17666
## Year2004          -0.07597    0.05685   -1.34  0.18149
```

```
## Year2005          -0.08557    0.05457   -1.57  0.11690
## Year2006          -0.11643    0.05533   -2.10  0.03541 *
## Year2007          -0.15936    0.05462   -2.92  0.00354 **
## Year2008          -0.13399    0.05296   -2.53  0.01143 *
## Year2009          -0.12700    0.05175   -2.45  0.01414 *
## Year2010          -0.10994    0.05274   -2.08  0.03717 *
## Year2011          -0.16008    0.05372   -2.98  0.00290 **
## Year2012          -0.20380    0.05507   -3.70  0.00022 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.642
## Multiple R-squared:  0.0151, Adjusted R-squared:  0.012
## Convergence in 10 IRWLS iterations
##
## Robustness weights:
##  observation 3271 is an outlier with |weight| = 0 ( < 1.9e-05);
##  455 weights are ~= 1. The remaining 4891 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0388  0.8720  0.9500  0.9140  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.87e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Sample size for the above analysis:  5347"
## [1] ""
## [1] ""
## [1] "####################################"
## [1] "Analysis of AJSC 2604"
## [1] "####################################"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 3070 3061 2909  196  361 4902 4206 2499 2653 2974 3422 3890 3815 4104 3902
## 2011 2012
## 3795 3797
##
```

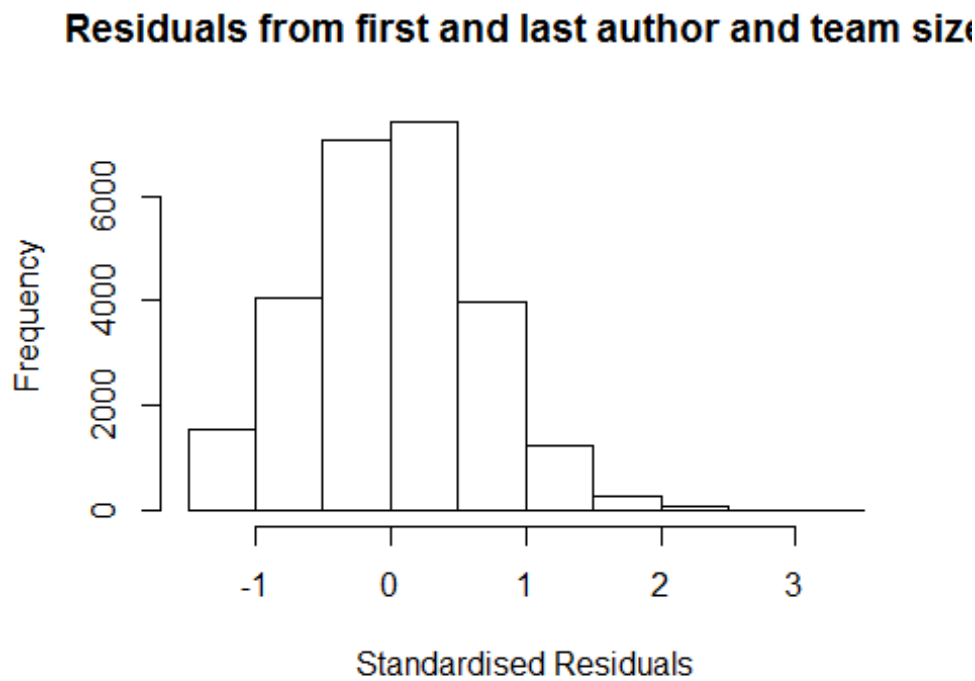
```
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 1716 1665 1321  111  167 2542 2541 1438 1531 1720 1965 2251 2277 2432 2313
## 2011 2012
## 2260 2294
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 1479 1410 1139   97  140 2118 2154 1241 1290 1438 1658 1854 1903 2059 1916
## 2011 2012
## 1845 1901
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
##  Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 270, df = 16, p-value <2e-16
```




```
##
##  Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 30, df = 1, p-value = 5e-08
```



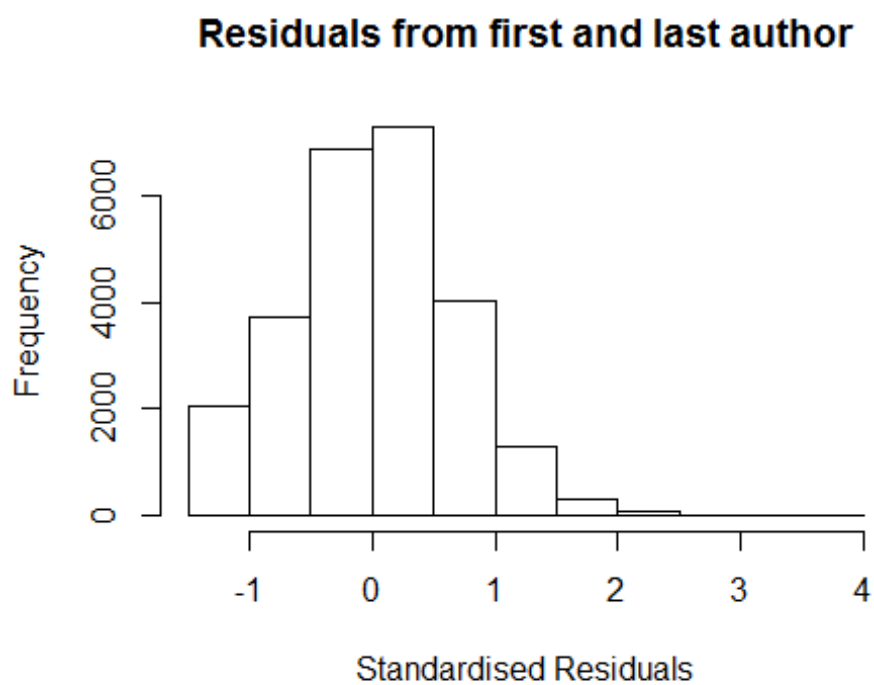

```
## [1] "Female first author team size 2018 geometric mean: 2.30476516186485"
## [1] "Male first author team size 2018 geometric mean: 1.98896297737054"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 180000, p-value = 1e-04
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 2.31082754538105"
## [1] "Male last author team size 2018 geometric mean: 1.9930407119793"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 170000, p-value = 0.002
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.181  1           1.087
## LastAuthorFemale  1.170  1           1.082
## UniqueAuthors     1.073  4           1.009
## Year              1.067 16           1.002
```




```
## [1] "List of  18 outliers with residuals above  2.5"
##          ScopusId  NLCS Year OneField Fields residuals
## 523    0030142764 3.833 1996     1702      4     2.742
## 1576  30244489068 4.025 1996     2604      1     2.934
## 2612   0030287048 3.862 1996     1711      3     2.771
## 4178   0031185845 4.068 1997     1702      5     2.700
## 6066   0031211090 4.640 1997     1703      4     3.337
## 8910   0032131292 3.979 1998     2604      2     2.597
## 16987 33646589837 4.485 2002     1706      3     3.427
## 18426  0035997407 3.336 2002     2600      2     2.526
## 32911 33745604236 4.326 2006     2600      2     3.043
## 33495 33646365077 3.822 2006     2600      2     2.846
## 33999 29144523061 3.760 2006     1700      7     2.542
## 40077 85032751965 3.652 2007     1711      3     2.687
## 40544 57349174008 4.128 2008     2600      3     2.860
## 44704 85032750937 4.080 2008     1711      3     2.877
## 48066 70049091784 3.562 2009     2600      2     2.635
## 49401 84969334819 4.392 2009     2600      2     3.223
## 58419 79953048649 3.731 2011     1702      5     2.530
## 58444 79551550744 4.060 2011     2600      2     2.863
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
UniqueAuthors +
##     Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##     k.max = 1000))
```

```
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.4019 -0.4341  0.0113  0.4299  3.4266
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.09082    0.02065   52.83  < 2e-16 ***
## FirstAuthorFemale1 -0.05859    0.01260   -4.65  3.3e-06 ***
## LastAuthorFemale1  -0.03478    0.01330   -2.62   0.0089 **
## UniqueAuthors2      0.24168    0.00979   24.68  < 2e-16 ***
## UniqueAuthors3      0.30691    0.01230   24.96  < 2e-16 ***
## UniqueAuthors4      0.31107    0.01749   17.78  < 2e-16 ***
## UniqueAuthors5      0.26365    0.01741   15.14  < 2e-16 ***
## Year1997           -0.02938    0.02766   -1.06   0.2880
## Year1998           -0.01574    0.02910   -0.54   0.5887
## Year1999            0.02657    0.07324    0.36   0.7167
## Year2000           -0.08127    0.06563   -1.24   0.2156
## Year2001           -0.26404    0.02642  -10.00  < 2e-16 ***
## Year2002           -0.28070    0.02581  -10.88  < 2e-16 ***
## Year2003           -0.06643    0.02706   -2.45   0.0141 *
## Year2004           -0.05079    0.02608   -1.95   0.0515 .
## Year2005           -0.07923    0.02592   -3.06   0.0022 **
## Year2006           -0.11456    0.02530   -4.53  6.0e-06 ***
## Year2007           -0.12615    0.02445   -5.16  2.5e-07 ***
## Year2008           -0.12985    0.02443   -5.31  1.1e-07 ***
## Year2009           -0.16362    0.02428   -6.74  1.6e-11 ***
## Year2010           -0.16520    0.02477   -6.67  2.6e-11 ***
## Year2011           -0.20072    0.02480   -8.09  6.1e-16 ***
## Year2012           -0.23301    0.02509   -9.29  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.641
## Multiple R-squared:  0.0516, Adjusted R-squared:  0.0508
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  4 observations c(2537,6512,13111,19785)
##   are outliers with |weight| = 0 ( < 3.9e-06);
##  2156 weights are ~= 1. The remaining 23482 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0019  0.8690  0.9510  0.9090  0.9860  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          3.90e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
```

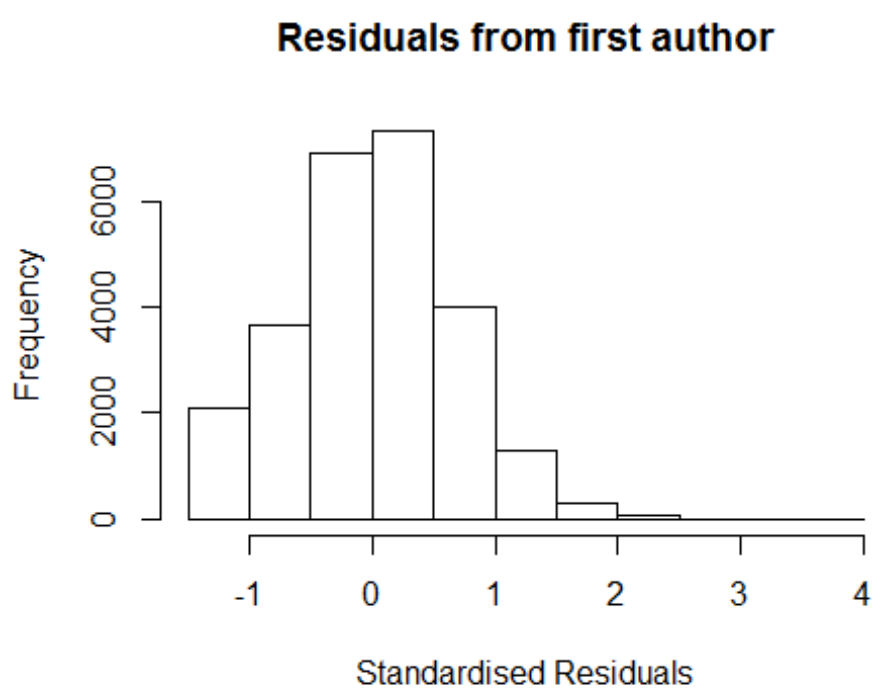
```
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.160  1           1.077
## LastAuthorFemale  1.159  1           1.076
## Year              1.005 16           1.000
```




```
## [1] "List of  20 outliers with residuals above  2.5"
##          ScopusId  NLCS Year OneField Fields residuals
## 523    0030142764 3.833 1996     1702      4     2.623
## 1576  30244489068 4.025 1996     2604      1     2.815
## 2612   0030287048 3.862 1996     1711      3     2.652
## 4178   0031185845 4.068 1997     1702      5     2.890
## 6066   0031211090 4.640 1997     1703      4     3.462
## 8910   0032131292 3.979 1998     2604      2     2.767
## 16987 33646589837 4.485 2002     1706      3     3.515
## 19817 85032751930 3.522 2002     1711      3     2.525
## 32911 33745604236 4.326 2006     2600      2     3.195
```

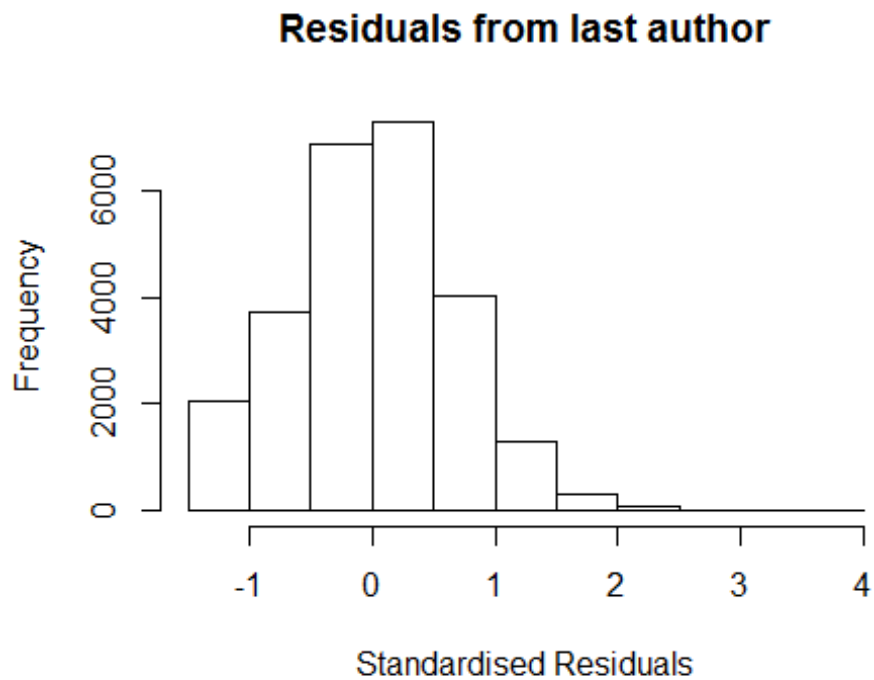
```
## 33495 33646365077 3.822 2006     2600      2     2.691
## 33999 29144523061 3.760 2006     1700      7     2.629
## 40077 85032751965 3.652 2007     1711      3     2.529
## 40544 57349174008 4.128 2008     2600      3     3.015
## 44704 85032750937 4.080 2008     1711      3     2.967
## 45339 71049116435 3.639 2009     1703      4     2.558
## 49401 84969334819 4.392 2009     2600      2     3.311
## 50448 78549288866 3.635 2010     2604      3     2.547
## 56959 79959527478 3.598 2011     1702      5     2.544
## 58419 79953048649 3.731 2011     1702      5     2.677
## 58444 79551550744 4.060 2011     2600      2     3.006
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.3090 -0.4403  0.0112  0.4392  3.5150
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.21022    0.02071   58.43  < 2e-16 ***
## FirstAuthorFemale1 -0.02690    0.01273   -2.11  0.03458 *
## LastAuthorFemale1  -0.03058    0.01349   -2.27  0.02336 *
## Year1997           -0.03244    0.02833   -1.15  0.25221
## Year1998            0.00132    0.02997    0.04  0.96500
## Year1999            0.09875    0.07498    1.32  0.18785
## Year2000           -0.02604    0.06295   -0.41  0.67906
## Year2001           -0.19188    0.02618   -7.33  2.4e-13 ***
## Year2002           -0.21331    0.02574   -8.29  < 2e-16 ***
## Year2003           -0.04678    0.02770   -1.69  0.09122 .
## Year2004           -0.02212    0.02674   -0.83  0.40818
## Year2005           -0.05503    0.02658   -2.07  0.03845 *
## Year2006           -0.07880    0.02602   -3.03  0.00246 **
## Year2007           -0.08754    0.02503   -3.50  0.00047 ***
## Year2008           -0.09743    0.02504   -3.89  1.0e-04 ***
## Year2009           -0.12950    0.02493   -5.20  2.1e-07 ***
## Year2010           -0.12227    0.02549   -4.80  1.6e-06 ***
## Year2011           -0.15605    0.02545   -6.13  8.9e-10 ***
## Year2012           -0.18393    0.02574   -7.15  9.2e-13 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.657
## Multiple R-squared:  0.0113, Adjusted R-squared:  0.0106
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
```

```
##  4 observations c(2537,6512,13111,19785)
##   are outliers with |weight| = 0 ( < 3.9e-06);
##  2175 weights are ~= 1. The remaining 23463 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0017  0.8690  0.9510  0.9100  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          3.90e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.003  1           1.002
## Year              1.003 16           1.000
```



```
## [1] "List of  20 outliers with residuals above  2.5"
##          ScopusId  NLCS Year OneField Fields residuals
## 523    0030142764 3.833 1996     1702      4     2.623
## 1576  30244489068 4.025 1996     2604      1     2.815
## 2612   0030287048 3.862 1996     1711      3     2.652
## 4178   0031185845 4.068 1997     1702      5     2.890
## 6066   0031211090 4.640 1997     1703      4     3.462
## 8910   0032131292 3.979 1998     2604      2     2.767
## 16987 33646589837 4.485 2002     1706      3     3.515
## 19817 85032751930 3.522 2002     1711      3     2.525
## 32911 33745604236 4.326 2006     2600      2     3.195
## 33495 33646365077 3.822 2006     2600      2     2.691
## 33999 29144523061 3.760 2006     1700      7     2.629
## 40077 85032751965 3.652 2007     1711      3     2.529
## 40544 57349174008 4.128 2008     2600      3     3.015
## 44704 85032750937 4.080 2008     1711      3     2.967
## 45339 71049116435 3.639 2009     1703      4     2.558
## 49401 84969334819 4.392 2009     2600      2     3.311
## 50448 78549288866 3.635 2010     2604      3     2.547
## 56959 79959527478 3.598 2011     1702      5     2.544
## 58419 79953048649 3.731 2011     1702      5     2.677
## 58444 79551550744 4.060 2011     2600      2     3.006
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.3072 -0.4382  0.0105  0.4400  3.5282
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.20831    0.02069   58.41  < 2e-16 ***
## FirstAuthorFemale1 -0.03797    0.01188   -3.20  0.00140 **
## Year1997           -0.03220    0.02833   -1.14  0.25562
## Year1998            0.00121    0.02997    0.04  0.96781
## Year1999            0.09888    0.07500    1.32  0.18738
## Year2000           -0.02547    0.06300   -0.40  0.68595
## Year2001           -0.19215    0.02617   -7.34  2.2e-13 ***
## Year2002           -0.21358    0.02573   -8.30  < 2e-16 ***
## Year2003           -0.04724    0.02769   -1.71  0.08801 .
## Year2004           -0.02228    0.02674   -0.83  0.40483
## Year2005           -0.05568    0.02658   -2.09  0.03622 *
## Year2006           -0.07874    0.02602   -3.03  0.00248 **
## Year2007           -0.08790    0.02503   -3.51  0.00045 ***
## Year2008           -0.09745    0.02504   -3.89  9.9e-05 ***
## Year2009           -0.12928    0.02493   -5.19  2.2e-07 ***
## Year2010           -0.12232    0.02549   -4.80  1.6e-06 ***
```

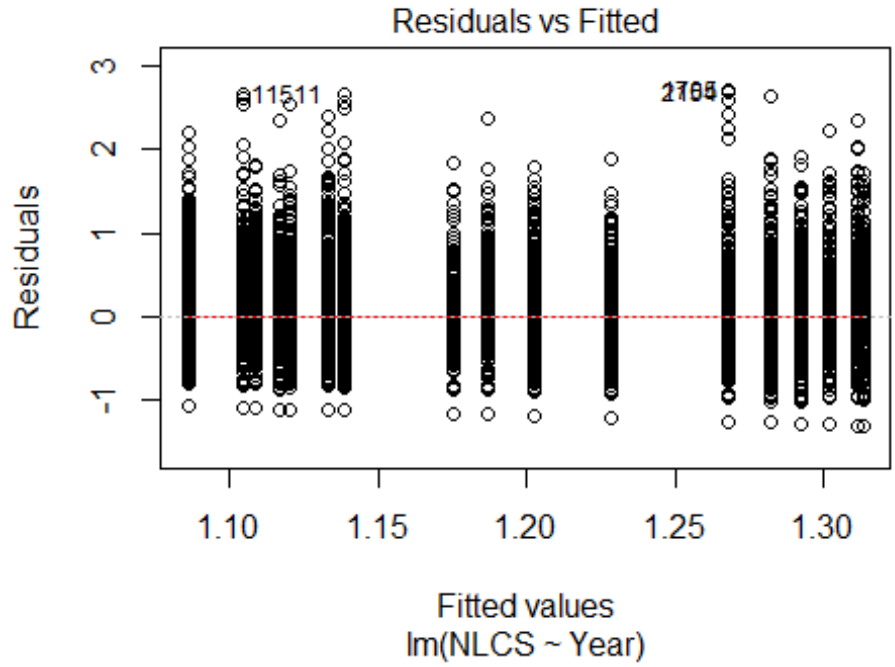
```
## Year2011           -0.15634    0.02545   -6.14  8.2e-10 ***
## Year2012           -0.18397    0.02574   -7.15  9.0e-13 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.657
## Multiple R-squared:  0.0112, Adjusted R-squared:  0.0105
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  4 observations c(2537,6512,13111,19785)
##   are outliers with |weight| = 0 ( < 3.9e-06);
##  2142 weights are ~= 1. The remaining 23496 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0016  0.8690  0.9510  0.9100  0.9860  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          3.90e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##                   GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.002  1           1.001
## Year             1.002 16           1.000
```



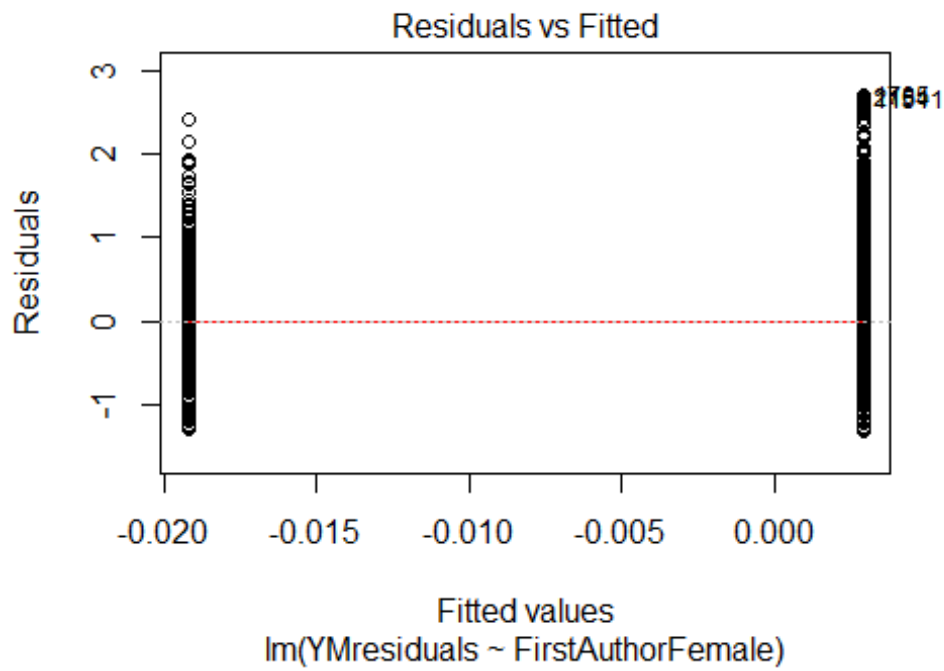

```
## [1] "List of  20 outliers with residuals above  2.5"
##          ScopusId  NLCS Year OneField Fields residuals
## 523    0030142764 3.833 1996     1702      4     2.623
## 1576  30244489068 4.025 1996     2604      1     2.815
## 2612   0030287048 3.862 1996     1711      3     2.652
## 4178   0031185845 4.068 1997     1702      5     2.890
## 6066   0031211090 4.640 1997     1703      4     3.462
## 8910   0032131292 3.979 1998     2604      2     2.767
## 16987 33646589837 4.485 2002     1706      3     3.515
## 19817 85032751930 3.522 2002     1711      3     2.525
## 32911 33745604236 4.326 2006     2600      2     3.195
## 33495 33646365077 3.822 2006     2600      2     2.691
## 33999 29144523061 3.760 2006     1700      7     2.629
## 40077 85032751965 3.652 2007     1711      3     2.529
## 40544 57349174008 4.128 2008     2600      3     3.015
## 44704 85032750937 4.080 2008     1711      3     2.967
## 45339 71049116435 3.639 2009     1703      4     2.558
## 49401 84969334819 4.392 2009     2600      2     3.311
## 50448 78549288866 3.635 2010     2604      3     2.547
## 56959 79959527478 3.598 2011     1702      5     2.544
## 58419 79953048649 3.731 2011     1702      5     2.677
## 58444 79551550744 4.060 2011     2600      2     3.006
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
```

```
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.3073 -0.4385  0.0106  0.4399  3.4900
##
## Coefficients:
##                   Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.2081     0.0207   58.48  < 2e-16 ***
## LastAuthorFemale1  -0.0416     0.0126   -3.30  0.00096 ***
## Year1997           -0.0318     0.0283   -1.12  0.26102
## Year1998            0.0010     0.0300    0.03  0.97331
## Year1999            0.0992     0.0750    1.32  0.18593
## Year2000           -0.0272     0.0630   -0.43  0.66535
## Year2001           -0.1916     0.0262   -7.32  2.6e-13 ***
## Year2002           -0.2131     0.0257   -8.28  < 2e-16 ***
## Year2003           -0.0469     0.0277   -1.69  0.09054 .
## Year2004           -0.0221     0.0267   -0.83  0.40851
## Year2005           -0.0549     0.0266   -2.06  0.03902 *
## Year2006           -0.0789     0.0260   -3.03  0.00241 **
## Year2007           -0.0880     0.0250   -3.52  0.00044 ***
## Year2008           -0.0976     0.0250   -3.90  9.7e-05 ***
## Year2009           -0.1299     0.0249   -5.21  1.9e-07 ***
## Year2010           -0.1227     0.0255   -4.82  1.5e-06 ***
## Year2011           -0.1567     0.0254   -6.16  7.6e-10 ***
## Year2012           -0.1845     0.0257   -7.17  7.7e-13 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.657
## Multiple R-squared:  0.0112, Adjusted R-squared:  0.0105
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  4 observations c(2537,6512,13111,19785)
##   are outliers with |weight| = 0 ( < 3.9e-06);
##  2187 weights are ~= 1. The remaining 23451 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0016  0.8690  0.9510  0.9100  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          3.90e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                    psi           subsampling                   cov
```

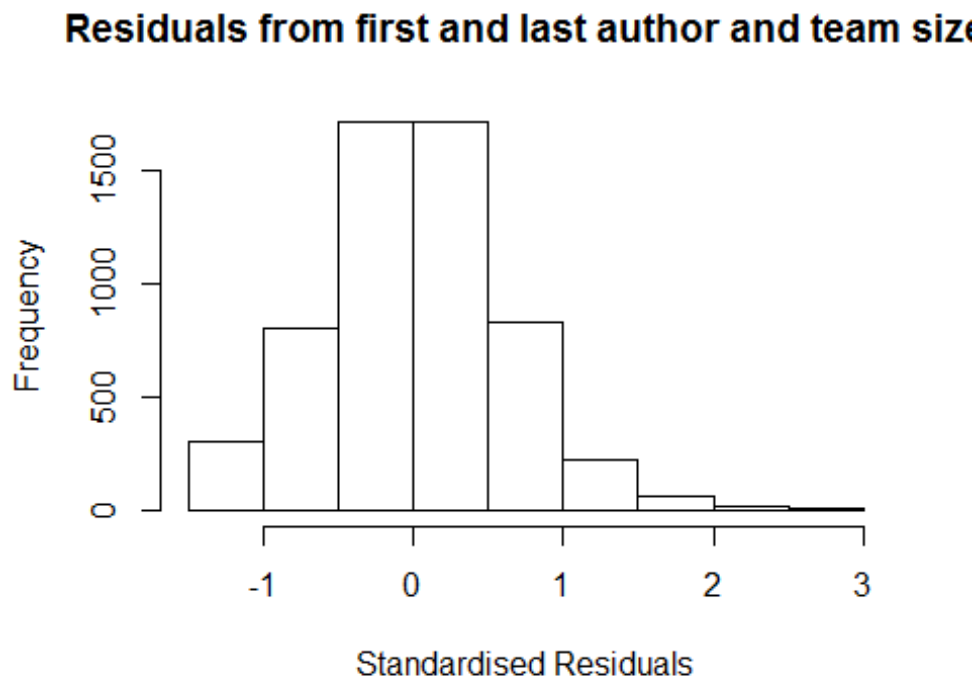
```
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Sample size for the above analysis:  25642"
## [1] ""
## [1] ""
## [1] "##################################"
## [1] "Analysis of AJSC 2605"
## [1] "##################################"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  625  570  692  512  618  617  605  604  697  865  860  947  888 1007  960
## 2011 2012
##  938  978
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  334  299  339  192  193  289  342  333  369  444  481  538  487  578  517
## 2011 2012
##  581  567
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  291  253  293  160  169  245  289  289  314  353  401  428  396  478  412
## 2011 2012
##  462  454
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
##  Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 82, df = 16, p-value = 7e-11
```




```
##
##  Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 16, df = 1, p-value = 7e-05
```



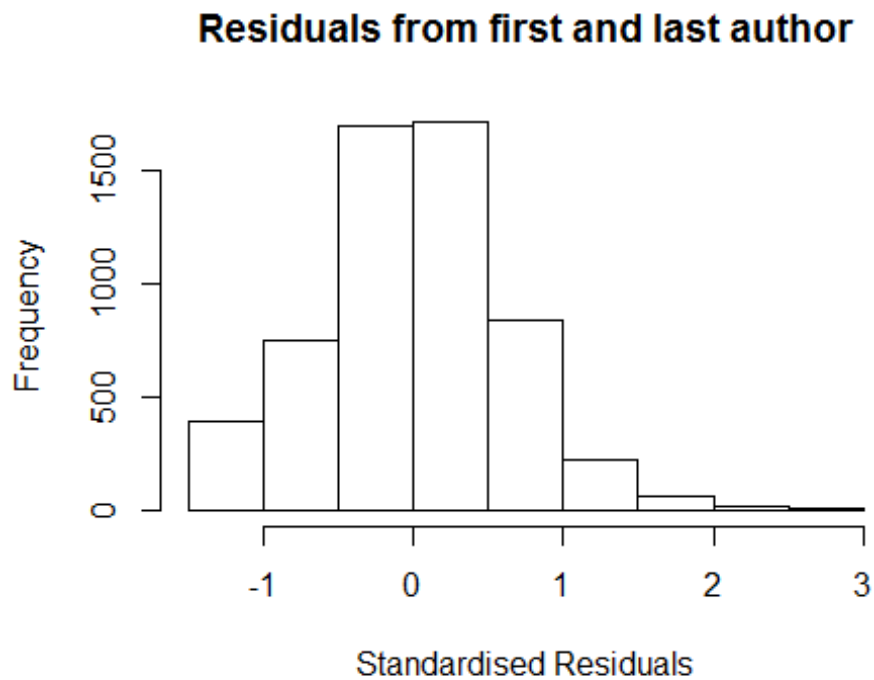

```
## [1] "Female first author team size 2018 geometric mean: 2.82434260152991"
## [1] "Male first author team size 2018 geometric mean: 2.31013946231325"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 16000, p-value = 0.008
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 2.754216996604"
## [1] "Male last author team size 2018 geometric mean: 2.32674628091961"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 14000, p-value = 0.07
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.153  1           1.074
## LastAuthorFemale  1.149  1           1.072
## UniqueAuthors     1.108  4           1.013
## Year              1.115 16           1.003
```




```
## [1] "List of  6 outliers with residuals above  2.5"
##          ScopusId  NLCS Year OneField Fields residuals
## 308    0037571112 3.672 1996     1600      2     2.500
## 1705   0032131292 3.979 1998     2604      2     2.631
## 2154   0031773680 3.959 1998     1303      6     2.670
## 10850 71049116435 3.639 2009     1703      4     2.540
## 11511 65449136284 3.759 2009     1303      7     2.601
## 13028 76149120388 3.793 2010     1600      2     2.680
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
UniqueAuthors +
##     Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##     k.max = 1000))
##  \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.48573 -0.38683  0.00526  0.39413  2.68032
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)          1.1719     0.0449   26.11  < 2e-16 ***
## FirstAuthorFemale1  -0.0642     0.0235   -2.74   0.0062 **
## LastAuthorFemale1    0.0270     0.0253    1.07   0.2858
## UniqueAuthors2       0.1893     0.0206    9.18  < 2e-16 ***
## UniqueAuthors3       0.2479     0.0244   10.16  < 2e-16 ***
```

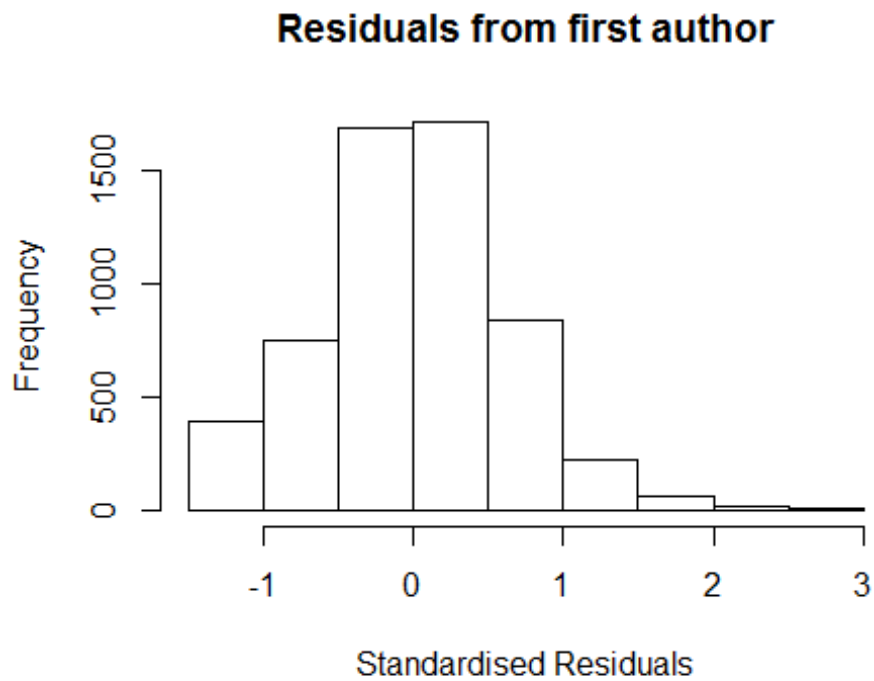
```
## UniqueAuthors4       0.3138     0.0305   10.30  < 2e-16 ***
## UniqueAuthors5       0.3799     0.0318   11.95  < 2e-16 ***
## Year1997            -0.1414     0.0619   -2.29   0.0223 *
## Year1998            -0.0723     0.0581   -1.24   0.2134
## Year1999            -0.0467     0.0737   -0.63   0.5260
## Year2000            -0.1482     0.0625   -2.37   0.0177 *
## Year2001             0.0303     0.0616    0.49   0.6227
## Year2002            -0.0588     0.0576   -1.02   0.3075
## Year2003            -0.0821     0.0559   -1.47   0.1423
## Year2004            -0.0960     0.0561   -1.71   0.0871 .
## Year2005            -0.1376     0.0539   -2.55   0.0108 *
## Year2006            -0.2031     0.0519   -3.91  9.2e-05 ***
## Year2007            -0.2286     0.0526   -4.35  1.4e-05 ***
## Year2008            -0.2103     0.0529   -3.98  7.1e-05 ***
## Year2009            -0.2623     0.0507   -5.17  2.4e-07 ***
## Year2010            -0.2485     0.0527   -4.72  2.4e-06 ***
## Year2011            -0.2475     0.0523   -4.73  2.3e-06 ***
## Year2012            -0.2848     0.0530   -5.38  7.9e-08 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.588
## Multiple R-squared:  0.058,  Adjusted R-squared:  0.0543
## Convergence in 10 IRWLS iterations
##
## Robustness weights:
##  502 weights are ~= 1. The remaining 5185 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0029  0.8640  0.9510  0.9010  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.76e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.119  1           1.058
```

```
## LastAuthorFemale  1.124  1           1.060
## Year              1.019 16           1.001
```



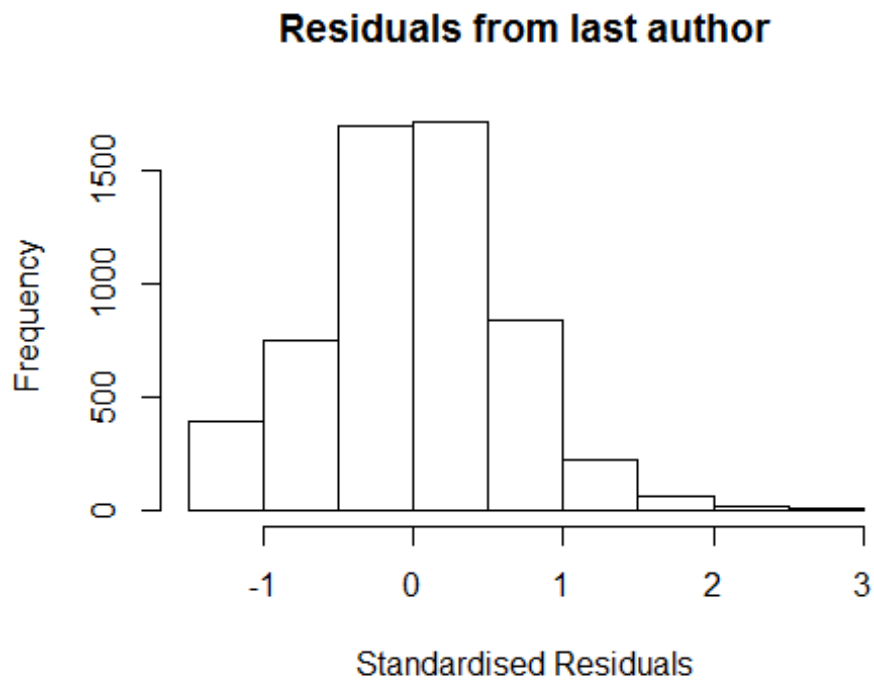

```
## [1] "List of  10 outliers with residuals above  2.5"
##          ScopusId  NLCS Year OneField Fields residuals
## 1454  11644261806 3.849 1998     1600      2     2.633
## 1705   0032131292 3.979 1998     2604      2     2.763
## 2154   0031773680 3.959 1998     1303      6     2.743
## 5871   4444221565 3.927 2004     1600      2     2.680
## 9766  55649115527 3.676 2008     2600      3     2.558
## 10817 70349932423 3.722 2009     1600      2     2.640
## 10850 71049116435 3.639 2009     1703      4     2.557
## 11511 65449136284 3.759 2009     1303      7     2.677
## 12159 78549288866 3.635 2010     2604      3     2.536
## 13028 76149120388 3.793 2010     1600      2     2.694
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.37601 -0.38954  0.00526  0.40626  2.76348
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
```

```
## (Intercept)          1.2826     0.0463   27.70  < 2e-16 ***
## FirstAuthorFemale1  -0.0279     0.0237   -1.17  0.24027
## LastAuthorFemale1    0.0390     0.0259    1.51  0.13204
## Year1997            -0.1534     0.0643   -2.39  0.01705 *
## Year1998            -0.0671     0.0605   -1.11  0.26748
## Year1999            -0.0465     0.0752   -0.62  0.53666
## Year2000            -0.1196     0.0650   -1.84  0.06567 .
## Year2001             0.0544     0.0636    0.86  0.39250
## Year2002            -0.0171     0.0605   -0.28  0.77680
## Year2003            -0.0470     0.0585   -0.80  0.42183
## Year2004            -0.0359     0.0583   -0.62  0.53824
## Year2005            -0.0758     0.0557   -1.36  0.17385
## Year2006            -0.1630     0.0538   -3.03  0.00247 **
## Year2007            -0.2002     0.0546   -3.67  0.00025 ***
## Year2008            -0.1642     0.0548   -3.00  0.00274 **
## Year2009            -0.2008     0.0528   -3.80  0.00015 ***
## Year2010            -0.1832     0.0550   -3.33  0.00086 ***
## Year2011            -0.1727     0.0544   -3.18  0.00150 **
## Year2012            -0.2139     0.0556   -3.85  0.00012 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.6
## Multiple R-squared:  0.0166, Adjusted R-squared:  0.0135
## Convergence in 10 IRWLS iterations
##
## Robustness weights:
##  472 weights are ~= 1. The remaining 5215 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0012  0.8640  0.9520  0.9010  0.9860  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.76e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.008  1           1.004
## Year              1.008 16           1.000
```




```
## [1] "List of  10 outliers with residuals above  2.5"
##          ScopusId  NLCS Year OneField Fields residuals
## 1454  11644261806 3.849 1998     1600      2     2.633
## 1705   0032131292 3.979 1998     2604      2     2.763
## 2154   0031773680 3.959 1998     1303      6     2.743
## 5871   4444221565 3.927 2004     1600      2     2.680
## 9766  55649115527 3.676 2008     2600      3     2.558
## 10817 70349932423 3.722 2009     1600      2     2.640
## 10850 71049116435 3.639 2009     1703      4     2.557
## 11511 65449136284 3.759 2009     1303      7     2.677
## 12159 78549288866 3.635 2010     2604      3     2.536
## 13028 76149120388 3.793 2010     1600      2     2.694
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.33907 -0.39128  0.00698  0.40529  2.76038
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)          1.2847     0.0463   27.76  < 2e-16 ***
## FirstAuthorFemale1  -0.0166     0.0225   -0.74  0.46063
## Year1997            -0.1538     0.0643   -2.39  0.01687 *
```

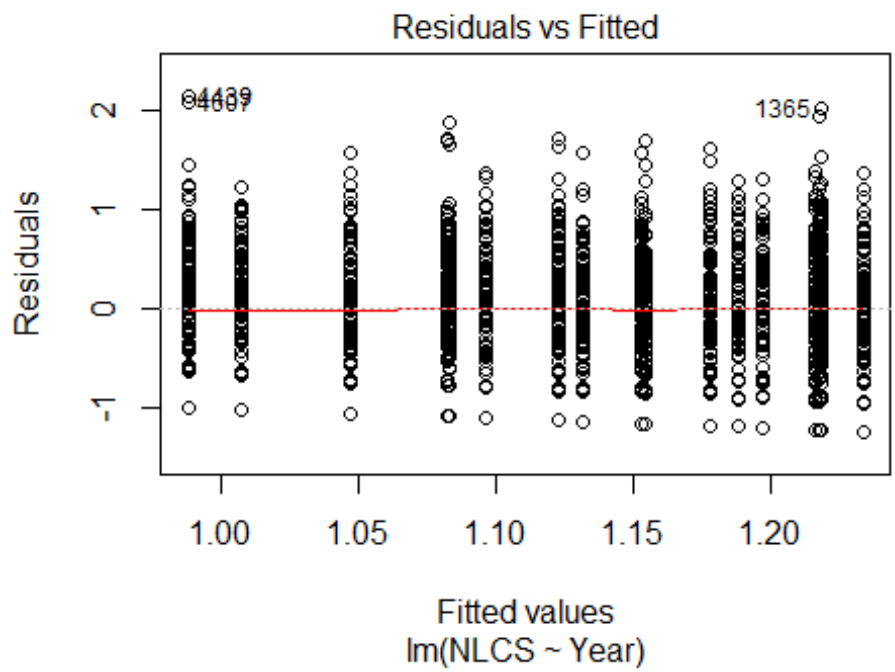
```
## Year1998            -0.0661     0.0605   -1.09  0.27421
## Year1999            -0.0453     0.0753   -0.60  0.54729
## Year2000            -0.1177     0.0649   -1.81  0.06979 .
## Year2001             0.0543     0.0636    0.85  0.39300
## Year2002            -0.0167     0.0605   -0.28  0.78228
## Year2003            -0.0467     0.0585   -0.80  0.42475
## Year2004            -0.0363     0.0584   -0.62  0.53373
## Year2005            -0.0749     0.0557   -1.34  0.17937
## Year2006            -0.1627     0.0539   -3.02  0.00254 **
## Year2007            -0.1989     0.0546   -3.64  0.00028 ***
## Year2008            -0.1635     0.0548   -2.98  0.00287 **
## Year2009            -0.1993     0.0528   -3.77  0.00016 ***
## Year2010            -0.1812     0.0550   -3.30  0.00099 ***
## Year2011            -0.1704     0.0543   -3.14  0.00172 **
## Year2012            -0.2124     0.0556   -3.82  0.00014 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.6
## Multiple R-squared:  0.0162, Adjusted R-squared:  0.0133
## Convergence in 10 IRWLS iterations
##
## Robustness weights:
##  469 weights are ~= 1. The remaining 5218 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0013  0.8640  0.9520  0.9010  0.9860  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.76e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##                   GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.012  1           1.006
## Year             1.012 16           1.000
```




```
## [1] "List of  10 outliers with residuals above  2.5"
##          ScopusId  NLCS Year OneField Fields residuals
## 1454  11644261806 3.849 1998     1600      2     2.633
## 1705   0032131292 3.979 1998     2604      2     2.763
## 2154   0031773680 3.959 1998     1303      6     2.743
## 5871   4444221565 3.927 2004     1600      2     2.680
## 9766  55649115527 3.676 2008     2600      3     2.558
## 10817 70349932423 3.722 2009     1600      2     2.640
## 10850 71049116435 3.639 2009     1703      4     2.557
## 11511 65449136284 3.759 2009     1303      7     2.677
## 12159 78549288866 3.635 2010     2604      3     2.536
## 13028 76149120388 3.793 2010     1600      2     2.694
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.36535 -0.39093  0.00433  0.40602  2.76608
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)          1.2793     0.0462   27.70  < 2e-16 ***
## LastAuthorFemale1    0.0297     0.0247    1.20  0.22852
## Year1997            -0.1528     0.0643   -2.38  0.01755 *
```

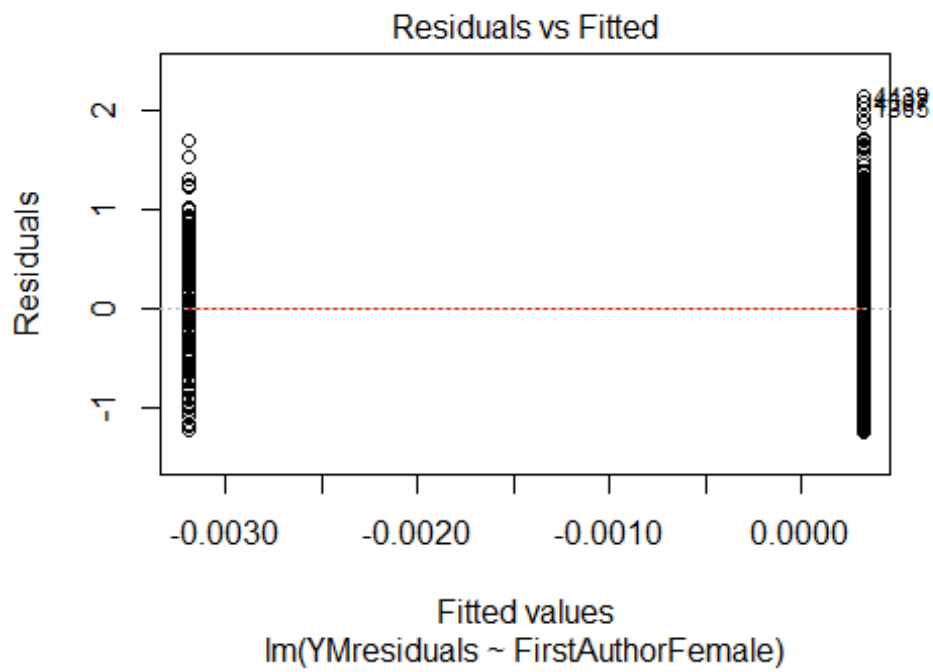
```
## Year1998           -0.0664     0.0605   -1.10  0.27269
## Year1999           -0.0462     0.0752   -0.61  0.53940
## Year2000           -0.1187     0.0649   -1.83  0.06738 .
## Year2001            0.0564     0.0636    0.89  0.37510
## Year2002           -0.0165     0.0604   -0.27  0.78524
## Year2003           -0.0466     0.0584   -0.80  0.42525
## Year2004           -0.0364     0.0583   -0.62  0.53242
## Year2005           -0.0756     0.0558   -1.36  0.17527
## Year2006           -0.1626     0.0538   -3.02  0.00253 **
## Year2007           -0.1998     0.0546   -3.66  0.00026 ***
## Year2008           -0.1640     0.0548   -2.99  0.00278 **
## Year2009           -0.2005     0.0528   -3.79  0.00015 ***
## Year2010           -0.1821     0.0550   -3.31  0.00093 ***
## Year2011           -0.1727     0.0543   -3.18  0.00149 **
## Year2012           -0.2132     0.0556   -3.83  0.00013 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.6
## Multiple R-squared:  0.0164, Adjusted R-squared:  0.0135
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  473 weights are ~= 1. The remaining 5214 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0011  0.8650  0.9520  0.9010  0.9860  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.76e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Sample size for the above analysis:  5687"
## [1] ""
## [1] ""
## [1] "####################################"
## [1] "Analysis of AJSC 2606"
## [1] "####################################"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
```

```
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  212  234  224  200  250  273  242  163  166  170  215  220  252  245  224
## 2011 2012
##  228  241
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  108  121  102   93  110  149  137   84   90  101  123  118  142  138  120
## 2011 2012
##  136  144
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##   97  100   92   83   91  122  122   64   77   75  106   94  118  112  101
## 2011 2012
##  110  116
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
##  Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 24, df = 16, p-value = 0.09
```



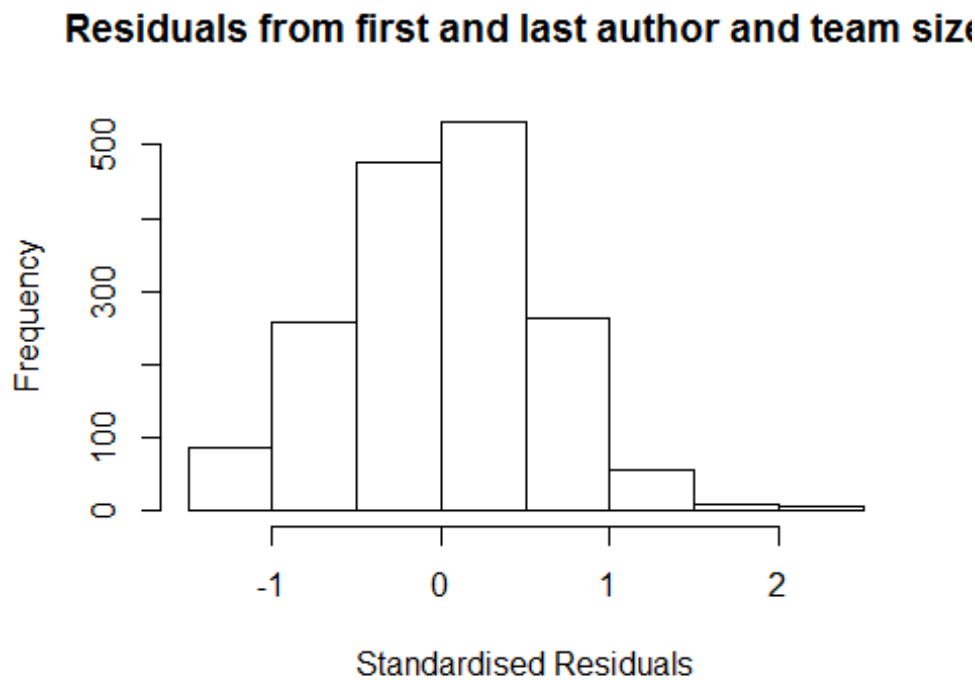

```
##
##  Bartlett test of homogeneity of variances
##
```

```
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 0.0048, df = 1, p-value = 0.9
```



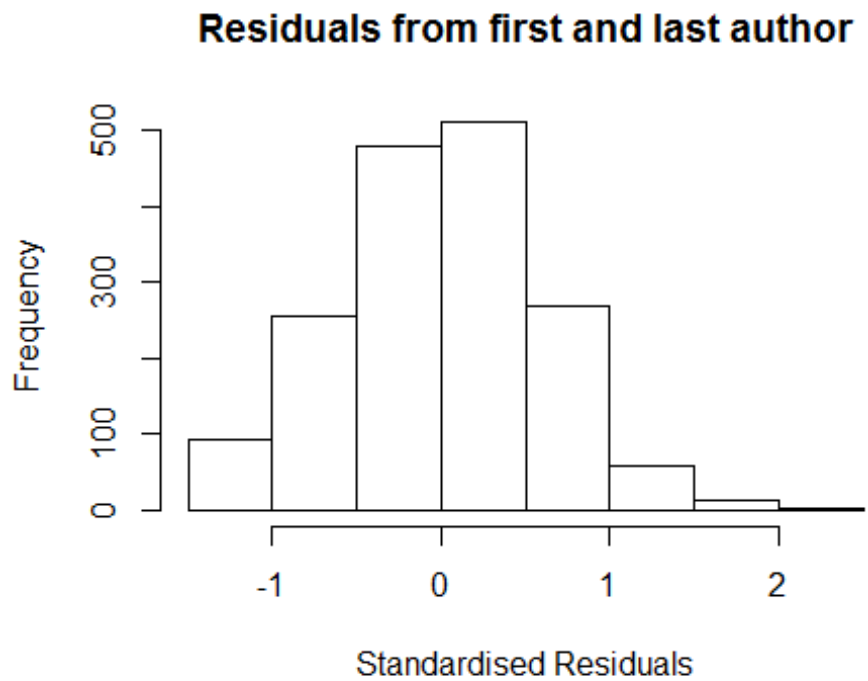

```
## [1] "Female first author team size 2018 geometric mean: 2.80215562036908"
## [1] "Male first author team size 2018 geometric mean: 2.2997563060958"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 3700, p-value = 0.08
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 2.44108327352057"
## [1] "Male last author team size 2018 geometric mean: 2.35943055997565"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 2700, p-value = 0.5
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.340  1           1.157
## LastAuthorFemale  1.363  1           1.168
```

```
## UniqueAuthors     1.167  4           1.020
## Year              1.245 16           1.007
```



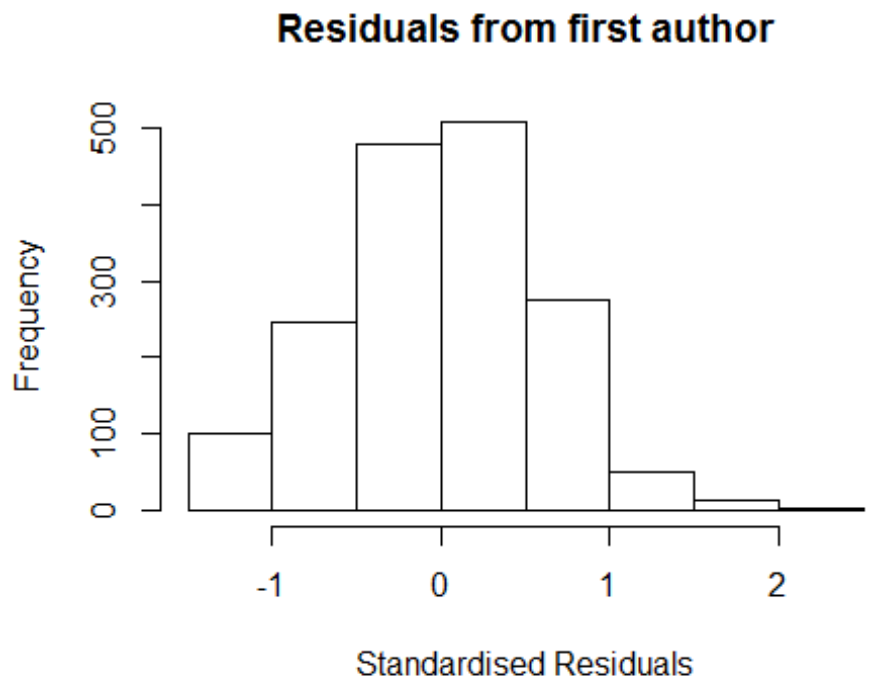

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
UniqueAuthors +
##     Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##     k.max = 1000))
##  \--> method = "MM"
## Residuals:
##    Min     1Q Median     3Q    Max
## -1.395 -0.407  0.024  0.392  2.185
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)           1.0639     0.0699   15.22  < 2e-16 ***
## FirstAuthorFemale1    0.0592     0.0616    0.96  0.33684
## LastAuthorFemale1    -0.1672     0.0613   -2.73  0.00647 **
## UniqueAuthors2        0.1838     0.0342    5.38  8.5e-08 ***
## UniqueAuthors3        0.2382     0.0478    4.98  7.1e-07 ***
## UniqueAuthors4        0.2945     0.0690    4.27  2.1e-05 ***
## UniqueAuthors5        0.3402     0.1020    3.33  0.00087 ***
```

```
## Year1997             0.0222     0.0982    0.23  0.82094
## Year1998             0.1009     0.0942    1.07  0.28422
## Year1999             0.0488     0.1018    0.48  0.63168
## Year2000            -0.1116     0.1011   -1.10  0.26989
## Year2001            -0.0069     0.0908   -0.08  0.93942
## Year2002            -0.0588     0.0900   -0.65  0.51352
## Year2003             0.0332     0.1006    0.33  0.74118
## Year2004             0.0201     0.0964    0.21  0.83463
## Year2005            -0.0275     0.0947   -0.29  0.77186
## Year2006            -0.0952     0.0900   -1.06  0.29061
## Year2007            -0.0877     0.0933   -0.94  0.34724
## Year2008            -0.0472     0.0835   -0.56  0.57237
## Year2009            -0.0476     0.0885   -0.54  0.59049
## Year2010            -0.1711     0.0876   -1.95  0.05112 .
## Year2011            -0.0710     0.0908   -0.78  0.43440
## Year2012            -0.2031     0.0969   -2.10  0.03627 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.603
## Multiple R-squared:  0.05,   Adjusted R-squared:  0.0374
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  144 weights are ~= 1. The remaining 1536 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.162   0.875   0.953   0.911   0.985   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          5.95e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.297  1           1.139
## LastAuthorFemale  1.342  1           1.159
## Year              1.101 16           1.003
```



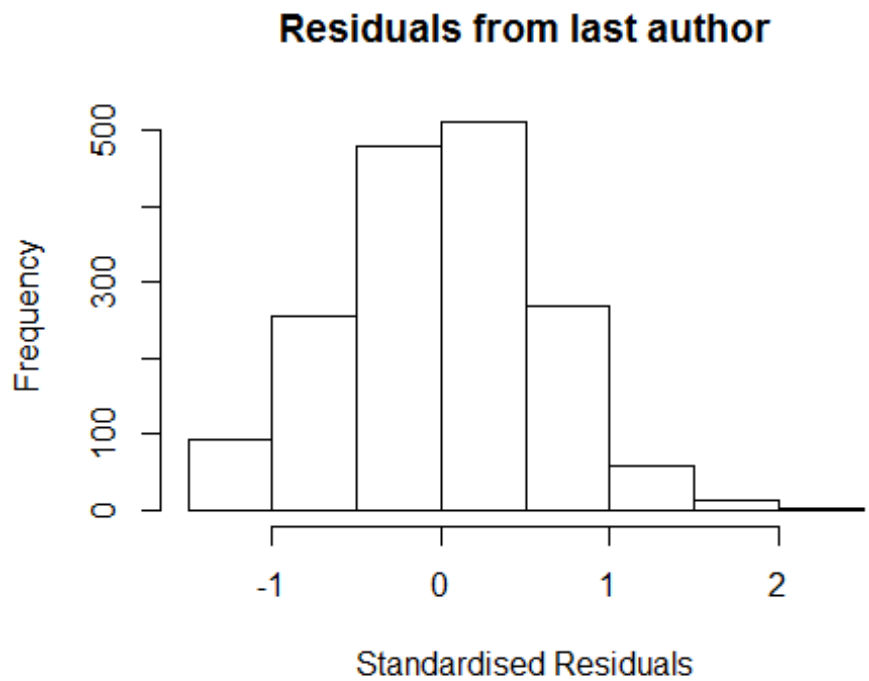

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.28400 -0.41613  0.00924  0.41812  2.14081
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.15749    0.06967   16.61   <2e-16 ***
## FirstAuthorFemale1  0.07422    0.06130    1.21   0.2261
## LastAuthorFemale1  -0.16849    0.06071   -2.78   0.0056 **
## Year1997            0.04680    0.09756    0.48   0.6315
## Year1998            0.09750    0.09320    1.05   0.2956
## Year1999            0.07942    0.10219    0.78   0.4372
## Year2000           -0.06507    0.09941   -0.65   0.5128
## Year2001            0.05251    0.09065    0.58   0.5625
## Year2002           -0.03338    0.09070   -0.37   0.7129
## Year2003            0.05229    0.09902    0.53   0.5975
## Year2004            0.03089    0.09888    0.31   0.7548
## Year2005            0.00253    0.09456    0.03   0.9787
```

```
## Year2006           -0.07545    0.09190   -0.82   0.4117
## Year2007           -0.04707    0.09239   -0.51   0.6105
## Year2008           -0.01497    0.08487   -0.18   0.8600
## Year2009           -0.00121    0.08871   -0.01   0.9891
## Year2010           -0.14687    0.08898   -1.65   0.0990 .
## Year2011           -0.03096    0.09310   -0.33   0.7395
## Year2012           -0.16729    0.09867   -1.70   0.0902 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.612
## Multiple R-squared:  0.0189, Adjusted R-squared:  0.00827
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
##  142 weights are ~= 1. The remaining 1538 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.196   0.868   0.950   0.911   0.986   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          5.95e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.039  1           1.019
## Year              1.039 16           1.001
```




```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.23763 -0.41709  0.00969  0.42503  2.16301
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.15e+00   6.96e-02   16.55   <2e-16 ***
## FirstAuthorFemale1 -2.65e-03   5.52e-02   -0.05    0.962
## Year1997            4.06e-02   9.71e-02    0.42    0.676
## Year1998            8.65e-02   9.28e-02    0.93    0.351
## Year1999            7.87e-02   1.04e-01    0.76    0.447
## Year2000           -6.36e-02   9.92e-02   -0.64    0.521
## Year2001            5.07e-02   9.08e-02    0.56    0.577
## Year2002           -3.18e-02   9.04e-02   -0.35    0.725
## Year2003            4.59e-02   9.97e-02    0.46    0.645
## Year2004            2.71e-02   9.96e-02    0.27    0.786
## Year2005           -1.86e-05   9.41e-02    0.00    1.000
## Year2006           -7.76e-02   9.14e-02   -0.85    0.396
```

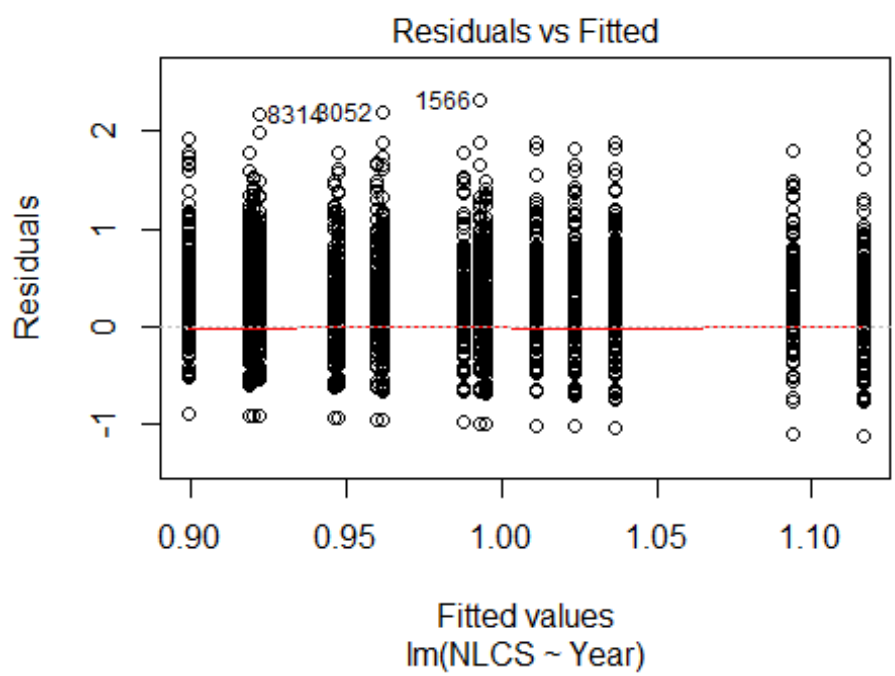
```
## Year2007           -4.68e-02   9.19e-02   -0.51    0.610
## Year2008           -1.24e-02   8.48e-02   -0.15    0.884
## Year2009            4.45e-03   8.86e-02    0.05    0.960
## Year2010           -1.49e-01   8.90e-02   -1.67    0.094 .
## Year2011           -3.16e-02   9.34e-02   -0.34    0.735
## Year2012           -1.83e-01   9.91e-02   -1.85    0.065 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.615
## Multiple R-squared:  0.0139, Adjusted R-squared:  0.00379
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  142 weights are ~= 1. The remaining 1538 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.191   0.869   0.950   0.911   0.986   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          5.95e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##                   GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.066  1           1.033
## Year             1.066 16           1.002
```



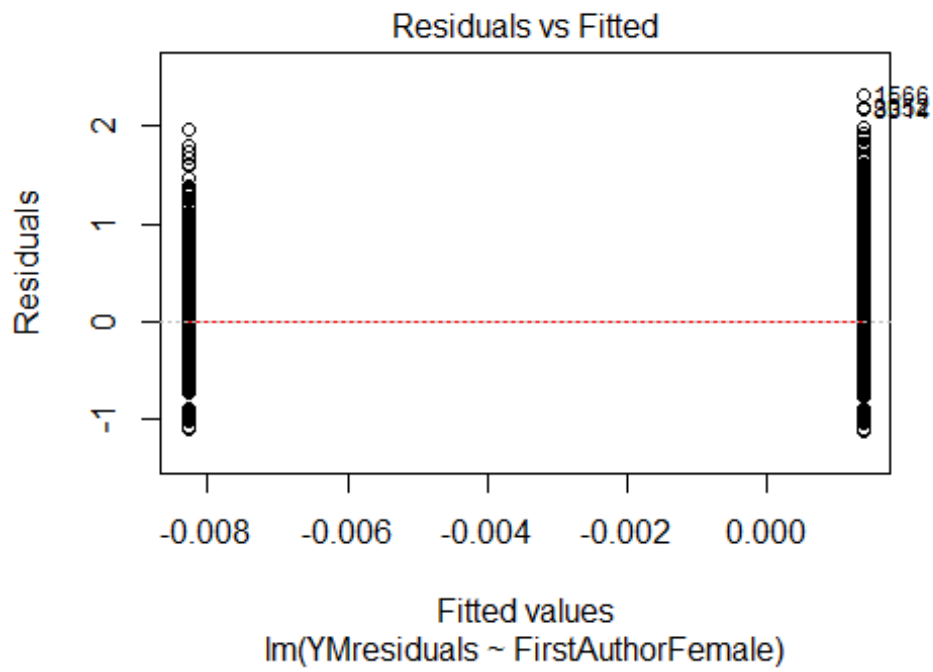

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.2589 -0.4196  0.0137  0.4154  2.1390
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.15922    0.06976   16.62   <2e-16 ***
## LastAuthorFemale1 -0.13515    0.05425   -2.49    0.013 *
## Year1997            0.04786    0.09748    0.49    0.624
## Year1998            0.09971    0.09326    1.07    0.285
## Year1999            0.08293    0.10223    0.81    0.417
## Year2000           -0.06475    0.09966   -0.65    0.516
## Year2001            0.05381    0.09087    0.59    0.554
## Year2002           -0.03064    0.09068   -0.34    0.736
## Year2003            0.05283    0.09914    0.53    0.594
## Year2004            0.03153    0.09934    0.32    0.751
## Year2005            0.00450    0.09438    0.05    0.962
## Year2006           -0.07409    0.09209   -0.80    0.421
```

```
## Year2007          -0.04479    0.09221   -0.49    0.627
## Year2008          -0.01121    0.08478   -0.13    0.895
## Year2009           0.00197    0.08869    0.02    0.982
## Year2010          -0.14236    0.08879   -1.60    0.109
## Year2011          -0.02963    0.09321   -0.32    0.751
## Year2012          -0.16724    0.09865   -1.70    0.090 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.614
## Multiple R-squared:  0.0179, Adjusted R-squared:  0.00788
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
##  144 weights are ~= 1. The remaining 1536 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.200   0.867   0.950   0.911   0.986   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          5.95e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Sample size for the above analysis:  1680"
## [1] ""
## [1] ""
## [1] "####################################"
## [1] "Analysis of AJSC 2607"
## [1] "####################################"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  315  367  355  357  429  308  386  299  369  356  392  433  456  501  482
## 2011 2012
##  451  425
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  184  231  215  219  223  207  273  191  267  237  281  310  327  343  344
## 2011 2012
```

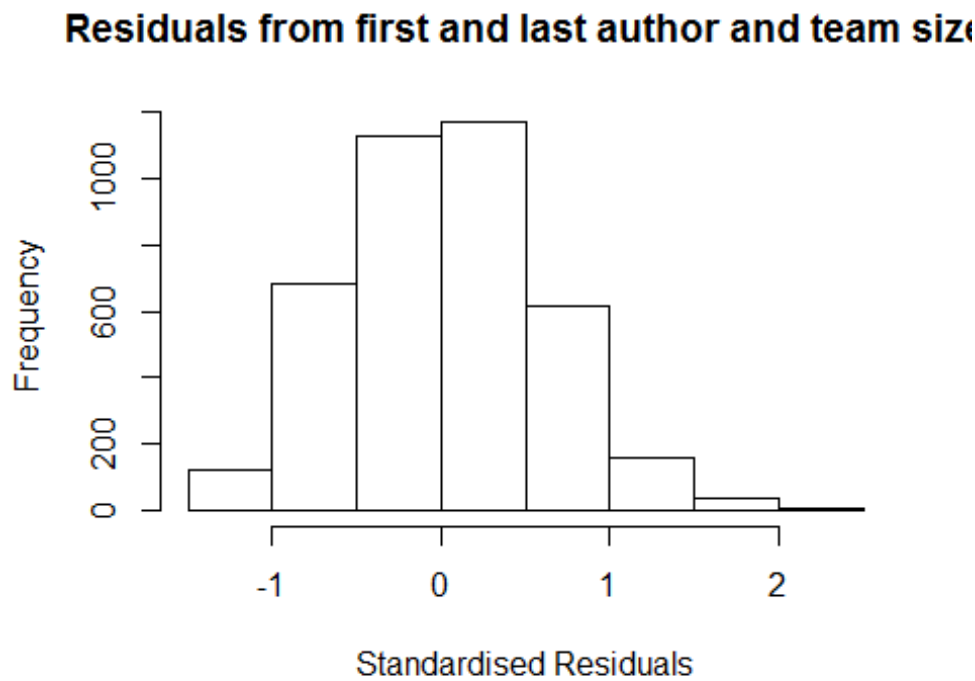
```
##  331  322
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  156  199  195  200  199  176  239  165  243  205  247  270  285  306  293
## 2011 2012
##  271  264
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
##  Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 24, df = 16, p-value = 0.09
```




```
##
##  Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 0.053, df = 1, p-value = 0.8
```

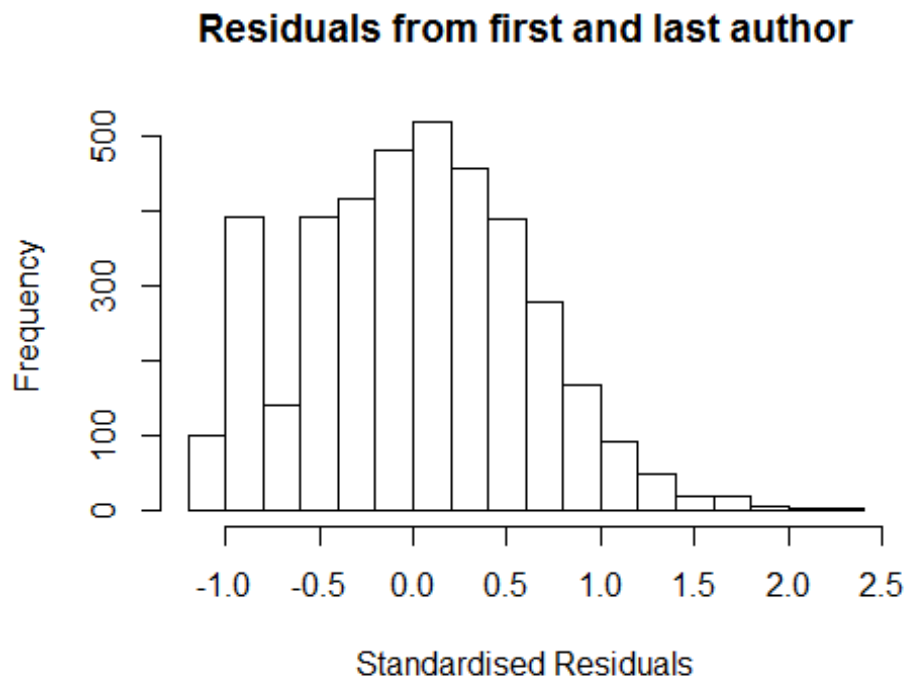



```
## [1] "Female first author team size 2018 geometric mean: 2.19879613093644"
## [1] "Male first author team size 2018 geometric mean: 1.68027455047536"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 8800, p-value = 0.002
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 2.29526790622924"
## [1] "Male last author team size 2018 geometric mean: 1.66801253934441"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 9000, p-value = 1e-04
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.391  1           1.180
## LastAuthorFemale  1.372  1           1.171
## UniqueAuthors     1.098  4           1.012
## Year              1.116 16           1.003
```



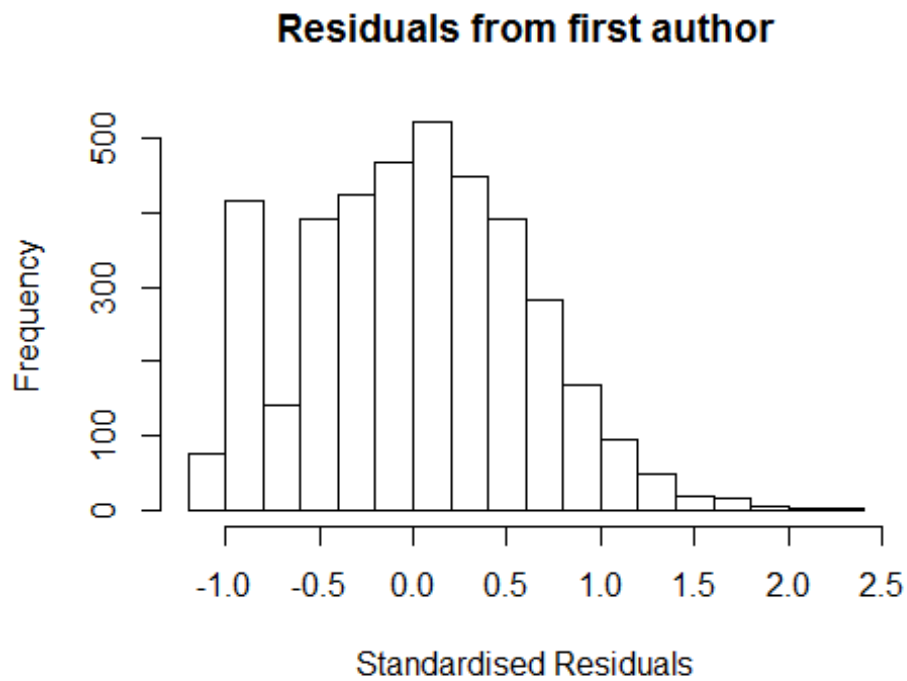

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
UniqueAuthors +
##     Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##     k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.3096 -0.4083  0.0112  0.4128  2.4173
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.00462    0.05657   17.76  < 2e-16 ***
## FirstAuthorFemale1  0.00779    0.03299    0.24   0.8134
## LastAuthorFemale1  -0.04793    0.03376   -1.42   0.1558
## UniqueAuthors2      0.13302    0.02234    5.96  2.8e-09 ***
## UniqueAuthors3      0.17405    0.03516    4.95  7.7e-07 ***
## UniqueAuthors4      0.19129    0.04737    4.04  5.5e-05 ***
## UniqueAuthors5      0.38699    0.09236    4.19  2.9e-05 ***
## Year1997           -0.05565    0.07175   -0.78   0.4380
## Year1998           -0.05041    0.07081   -0.71   0.4765
## Year1999           -0.10993    0.06971   -1.58   0.1149
```

```
## Year2000            0.00569    0.07017    0.08   0.9354
## Year2001           -0.08203    0.07435   -1.10   0.2700
## Year2002           -0.15653    0.06930   -2.26   0.0240 *
## Year2003           -0.12566    0.07005   -1.79   0.0729 .
## Year2004           -0.12463    0.06536   -1.91   0.0566 .
## Year2005           -0.10474    0.06914   -1.52   0.1299
## Year2006           -0.02958    0.06920   -0.43   0.6691
## Year2007           -0.08214    0.06666   -1.23   0.2180
## Year2008           -0.15853    0.06724   -2.36   0.0184 *
## Year2009           -0.13439    0.06424   -2.09   0.0365 *
## Year2010           -0.20247    0.06783   -2.98   0.0029 **
## Year2011           -0.19845    0.06827   -2.91   0.0037 **
## Year2012           -0.20686    0.06886   -3.00   0.0027 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.611
## Multiple R-squared:  0.0274, Adjusted R-squared:  0.0219
## Convergence in 10 IRWLS iterations
##
## Robustness weights:
##  336 weights are ~= 1. The remaining 3577 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0828  0.8650  0.9490  0.9130  0.9840  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          2.56e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.369  1           1.170
## LastAuthorFemale  1.360  1           1.166
## Year              1.032 16           1.001
```



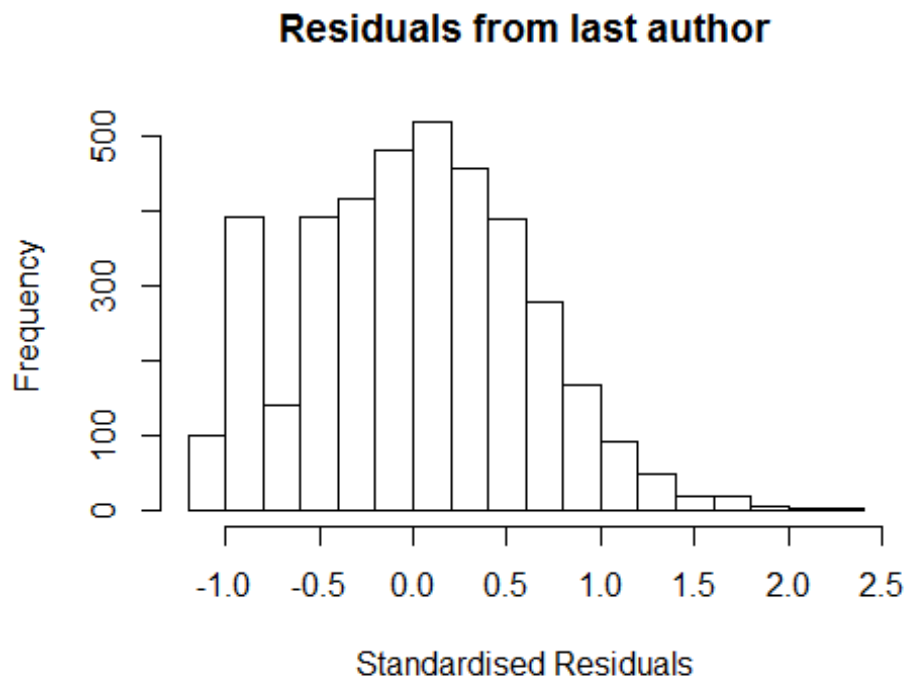

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.0857 -0.4192  0.0164  0.4198  2.3650
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)          1.05488    0.05776   18.26   <2e-16 ***
## FirstAuthorFemale1  0.01823    0.03333    0.55    0.584
## LastAuthorFemale1  -0.04142    0.03431   -1.21    0.227
## Year1997           -0.05264    0.07365   -0.71    0.475
## Year1998           -0.04629    0.07287   -0.64    0.525
## Year1999           -0.10791    0.07136   -1.51    0.131
## Year2000            0.01257    0.07183    0.17    0.861
## Year2001           -0.06861    0.07584   -0.90    0.366
## Year2002           -0.14366    0.07081   -2.03    0.043 *
## Year2003           -0.11012    0.07189   -1.53    0.126
## Year2004           -0.10869    0.06702   -1.62    0.105
## Year2005           -0.09404    0.07069   -1.33    0.183
```

```
## Year2006           -0.00772    0.07058   -0.11    0.913
## Year2007           -0.05991    0.06788   -0.88    0.378
## Year2008           -0.13710    0.06888   -1.99    0.047 *
## Year2009           -0.10968    0.06589   -1.66    0.096 .
## Year2010           -0.16226    0.06931   -2.34    0.019 *
## Year2011           -0.16641    0.06954   -2.39    0.017 *
## Year2012           -0.16874    0.07014   -2.41    0.016 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.618
## Multiple R-squared:  0.00882,    Adjusted R-squared:  0.00424
## Convergence in 10 IRWLS iterations
##
## Robustness weights:
##  355 weights are ~= 1. The remaining 3558 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.111   0.867   0.950   0.913   0.985   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          2.56e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.018  1           1.009
## Year              1.018 16           1.001
```




```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.0638 -0.4177  0.0125  0.4205  2.3682
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.05267    0.05755   18.29   <2e-16 ***
## FirstAuthorFemale1 -0.00217    0.02891   -0.07    0.940
## Year1997           -0.05398    0.07349   -0.73    0.463
## Year1998           -0.04719    0.07257   -0.65    0.516
## Year1999           -0.10885    0.07126   -1.53    0.127
## Year2000            0.01116    0.07169    0.16    0.876
## Year2001           -0.07035    0.07559   -0.93    0.352
## Year2002           -0.14391    0.07069   -2.04    0.042 *
## Year2003           -0.11021    0.07177   -1.54    0.125
## Year2004           -0.11054    0.06679   -1.65    0.098 .
## Year2005           -0.09548    0.07052   -1.35    0.176
## Year2006           -0.00702    0.07040   -0.10    0.921
```

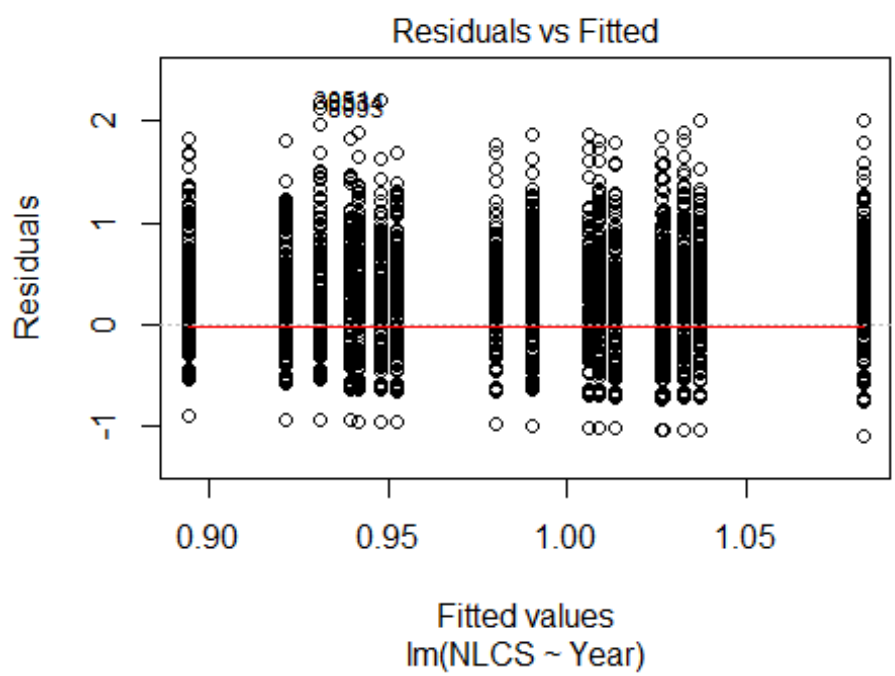
```
## Year2007           -0.05953    0.06773   -0.88    0.380
## Year2008           -0.13705    0.06874   -1.99    0.046 *
## Year2009           -0.10969    0.06580   -1.67    0.096 .
## Year2010           -0.16216    0.06915   -2.34    0.019 *
## Year2011           -0.16680    0.06941   -2.40    0.016 *
## Year2012           -0.16937    0.07003   -2.42    0.016 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.62
## Multiple R-squared:  0.00838,    Adjusted R-squared:  0.00406
## Convergence in 10 IRWLS iterations
##
## Robustness weights:
##  361 weights are ~= 1. The remaining 3552 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.112   0.866   0.950   0.913   0.985   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          2.56e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##                   GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.011  1           1.006
## Year             1.011 16           1.000
```



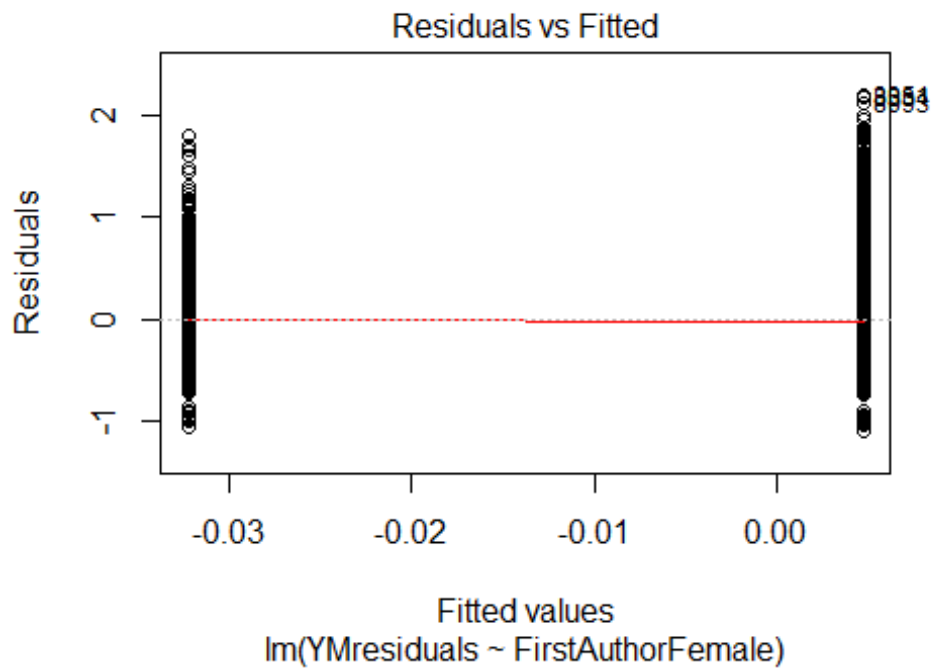

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.0687 -0.4203  0.0152  0.4187  2.3643
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.05663    0.05740   18.41   <2e-16 ***
## LastAuthorFemale1 -0.03218    0.02954   -1.09    0.276
## Year1997          -0.05322    0.07352   -0.72    0.469
## Year1998          -0.04681    0.07273   -0.64    0.520
## Year1999          -0.10889    0.07119   -1.53    0.126
## Year2000           0.01212    0.07168    0.17    0.866
## Year2001          -0.06992    0.07555   -0.93    0.355
## Year2002          -0.14422    0.07065   -2.04    0.041 *
## Year2003          -0.11047    0.07176   -1.54    0.124
## Year2004          -0.10983    0.06676   -1.65    0.100
## Year2005          -0.09440    0.07054   -1.34    0.181
## Year2006          -0.00831    0.07044   -0.12    0.906
```

```
## Year2007          -0.05999    0.06776   -0.89    0.376
## Year2008          -0.13764    0.06872   -2.00    0.045 *
## Year2009          -0.10942    0.06579   -1.66    0.096 .
## Year2010          -0.16278    0.06916   -2.35    0.019 *
## Year2011          -0.16660    0.06941   -2.40    0.016 *
## Year2012          -0.16875    0.07006   -2.41    0.016 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.618
## Multiple R-squared:  0.00874,    Adjusted R-squared:  0.00442
## Convergence in 10 IRWLS iterations
##
## Robustness weights:
##  356 weights are ~= 1. The remaining 3557 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.112   0.867   0.950   0.913   0.985   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          2.56e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Sample size for the above analysis:  3913"
## [1] ""
## [1] ""
## [1] "####################################"
## [1] "Analysis of AJSC 2608"
## [1] "####################################"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  337  406  377  368  364  338  351  293  382  373  412  419  462  534  543
## 2011 2012
##  559  527
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  234  261  237  245  216  232  255  208  265  260  277  306  334  404  414
## 2011 2012
```

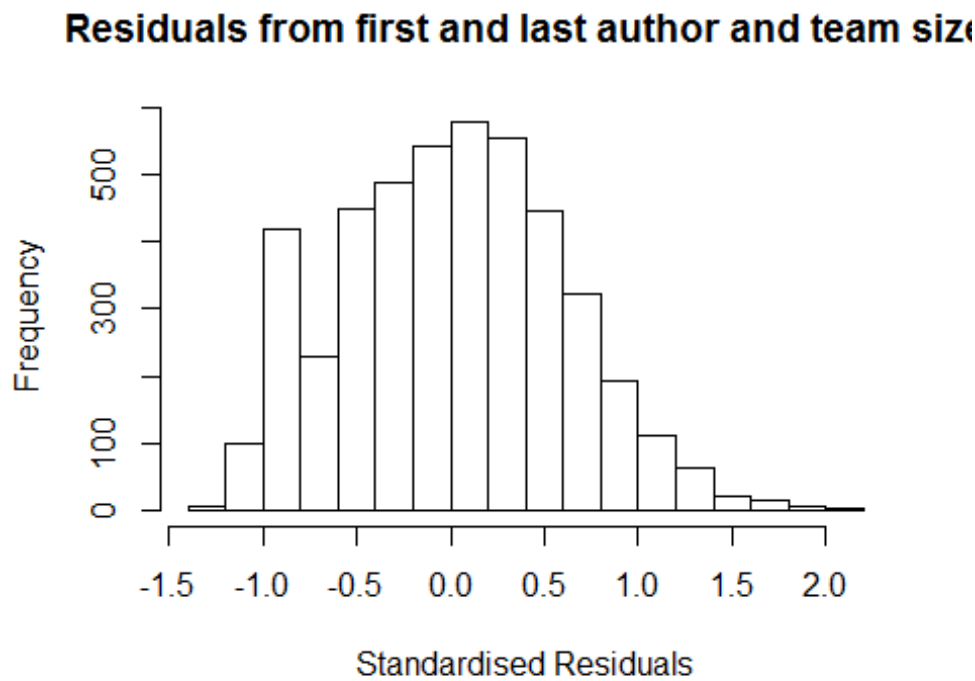
```
##  429  386
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  212  242  221  223  199  216  240  199  246  233  255  281  304  367  373
## 2011 2012
##  388  343
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
##  Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 30, df = 16, p-value = 0.02
```




```
##
##  Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 8.4, df = 1, p-value = 0.004
```

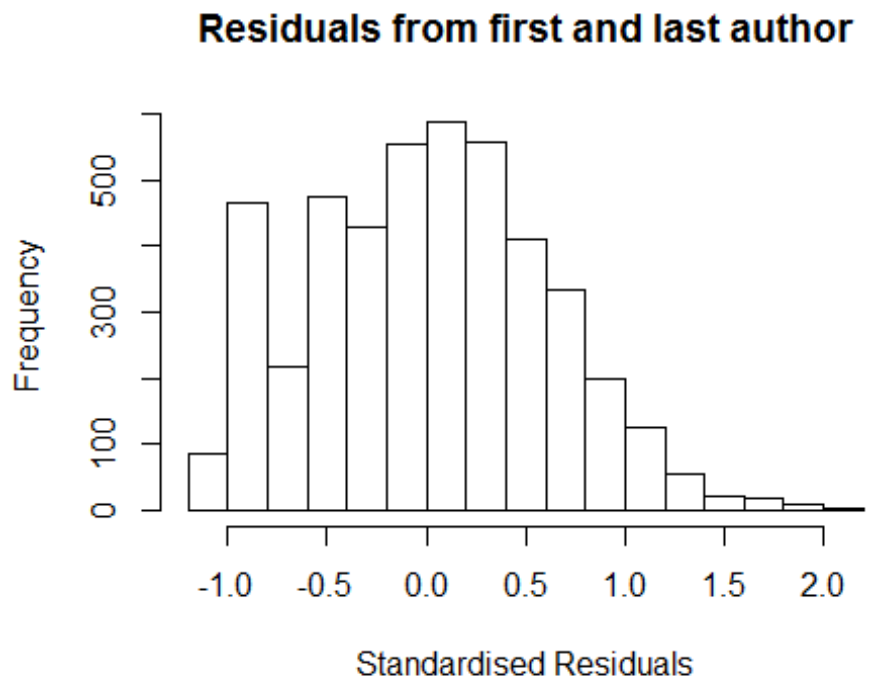



```
## [1] "Female first author team size 2018 geometric mean: 1.69864737943302"
## [1] "Male first author team size 2018 geometric mean: 1.4074647959324"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 9400, p-value = 0.02
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 1.75488120729826"
## [1] "Male last author team size 2018 geometric mean: 1.40102316283799"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 9400, p-value = 0.01
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.748  1           1.322
## LastAuthorFemale  1.750  1           1.323
## UniqueAuthors     1.077  4           1.009
## Year              1.101 16           1.003
```



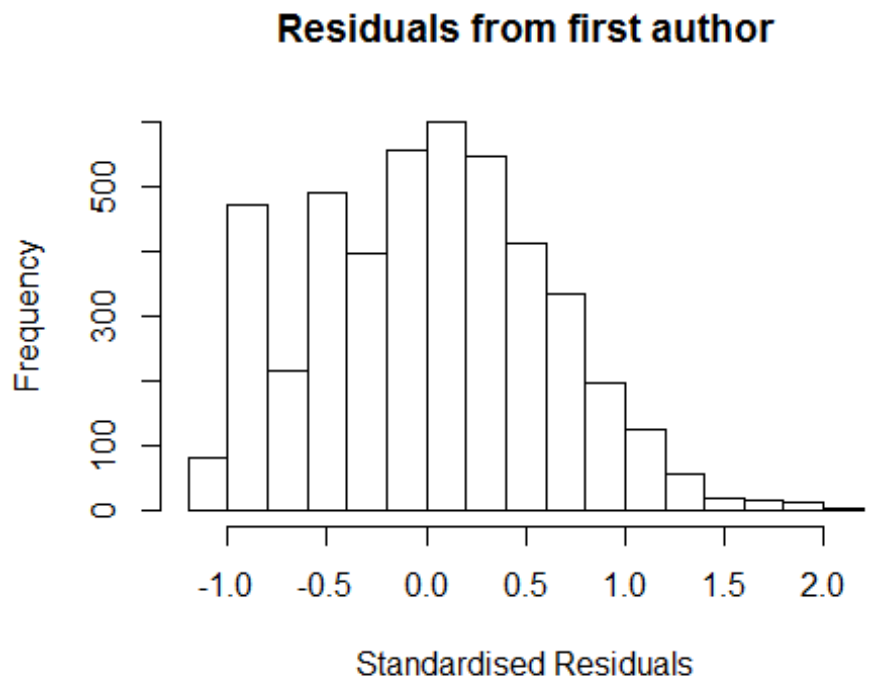

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
UniqueAuthors +
##     Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##     k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.3295 -0.4282  0.0159  0.4155  2.1097
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)           0.9954     0.0501   19.89  < 2e-16 ***
## FirstAuthorFemale1    0.0238     0.0333    0.71  0.47486
## LastAuthorFemale1    -0.0873     0.0339   -2.57  0.01012 *
## UniqueAuthors2        0.1788     0.0213    8.38  < 2e-16 ***
## UniqueAuthors3        0.2571     0.0370    6.95  4.2e-12 ***
## UniqueAuthors4        0.2358     0.0647    3.64  0.00027 ***
## UniqueAuthors5        0.4170     0.0926    4.51  6.8e-06 ***
## Year1997             -0.0434     0.0666   -0.65  0.51479
## Year1998             -0.0233     0.0624   -0.37  0.70952
## Year1999             -0.0621     0.0662   -0.94  0.34859
```

```
## Year2000            -0.0829     0.0634   -1.31  0.19144
## Year2001            -0.1024     0.0637   -1.61  0.10797
## Year2002            -0.1237     0.0631   -1.96  0.04994 *
## Year2003            -0.0963     0.0652   -1.48  0.13992
## Year2004            -0.0874     0.0656   -1.33  0.18269
## Year2005            -0.0498     0.0617   -0.81  0.41950
## Year2006            -0.1285     0.0630   -2.04  0.04159 *
## Year2007            -0.1237     0.0608   -2.04  0.04187 *
## Year2008            -0.1074     0.0619   -1.74  0.08277 .
## Year2009            -0.1585     0.0579   -2.74  0.00622 **
## Year2010            -0.0997     0.0594   -1.68  0.09343 .
## Year2011            -0.2110     0.0592   -3.57  0.00037 ***
## Year2012            -0.1623     0.0608   -2.67  0.00759 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.623
## Multiple R-squared:  0.0347, Adjusted R-squared:  0.03
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  341 weights are ~= 1. The remaining 4201 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.228   0.869   0.951   0.917   0.986   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          2.20e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.742  1           1.320
## LastAuthorFemale  1.744  1           1.321
## Year              1.030 16           1.001
```



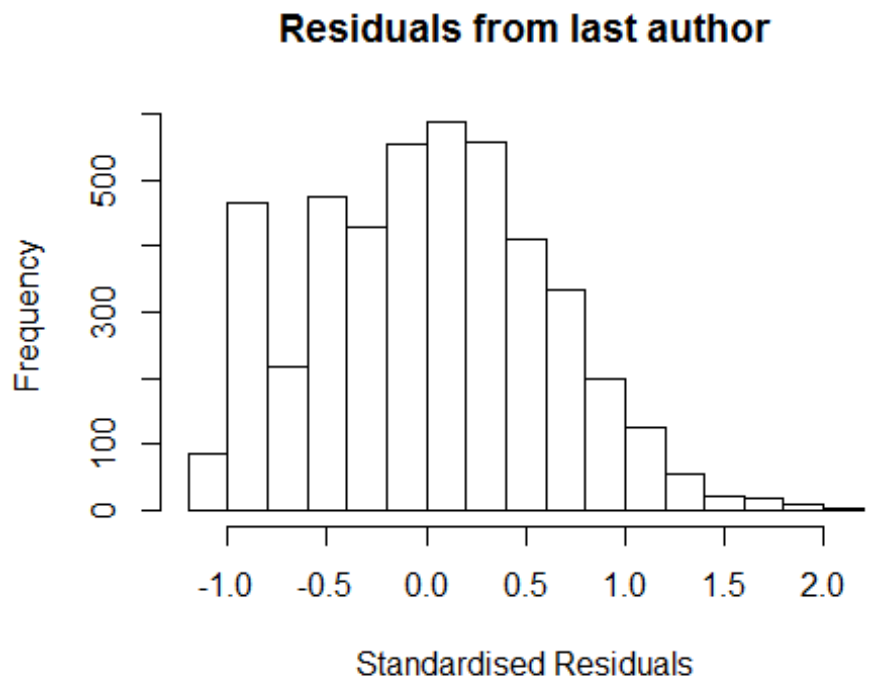

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.0823 -0.4384  0.0154  0.4242  2.1930
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)          1.05133    0.05046   20.83   <2e-16 ***
## FirstAuthorFemale1  0.03102    0.03385    0.92   0.3594
## LastAuthorFemale1  -0.08695    0.03450   -2.52   0.0118 *
## Year1997           -0.03423    0.06777   -0.51   0.6134
## Year1998           -0.00848    0.06371   -0.13   0.8941
## Year1999           -0.05521    0.06749   -0.82   0.4134
## Year2000           -0.07133    0.06391   -1.12   0.2645
## Year2001           -0.07278    0.06454   -1.13   0.2595
## Year2002           -0.11591    0.06354   -1.82   0.0682 .
## Year2003           -0.06748    0.06611   -1.02   0.3074
## Year2004           -0.06471    0.06670   -0.97   0.3320
## Year2005           -0.04227    0.06224   -0.68   0.4971
```

```
## Year2006           -0.10767    0.06370   -1.69   0.0911 .
## Year2007           -0.11457    0.06159   -1.86   0.0629 .
## Year2008           -0.10111    0.06282   -1.61   0.1076
## Year2009           -0.13958    0.05883   -2.37   0.0177 *
## Year2010           -0.07243    0.06015   -1.20   0.2286
## Year2011           -0.17759    0.05966   -2.98   0.0029 **
## Year2012           -0.13628    0.06105   -2.23   0.0256 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.631
## Multiple R-squared:  0.00749,    Adjusted R-squared:  0.00354
## Convergence in 10 IRWLS iterations
##
## Robustness weights:
##  385 weights are ~= 1. The remaining 4157 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.202   0.876   0.950   0.916   0.985   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          2.20e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.015  1           1.008
## Year              1.015 16           1.000
```




```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.0498 -0.4358  0.0185  0.4179  2.1970
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)           1.0498     0.0504   20.81   <2e-16 ***
## FirstAuthorFemale1   -0.0252     0.0261   -0.96   0.3348
## Year1997             -0.0344     0.0679   -0.51   0.6123
## Year1998             -0.0122     0.0637   -0.19   0.8474
## Year1999             -0.0568     0.0675   -0.84   0.4006
## Year2000             -0.0767     0.0639   -1.20   0.2300
## Year2001             -0.0756     0.0645   -1.17   0.2409
## Year2002             -0.1175     0.0636   -1.85   0.0649 .
## Year2003             -0.0689     0.0661   -1.04   0.2976
## Year2004             -0.0691     0.0667   -1.04   0.3002
## Year2005             -0.0444     0.0622   -0.71   0.4756
## Year2006             -0.1097     0.0636   -1.72   0.0847 .
```

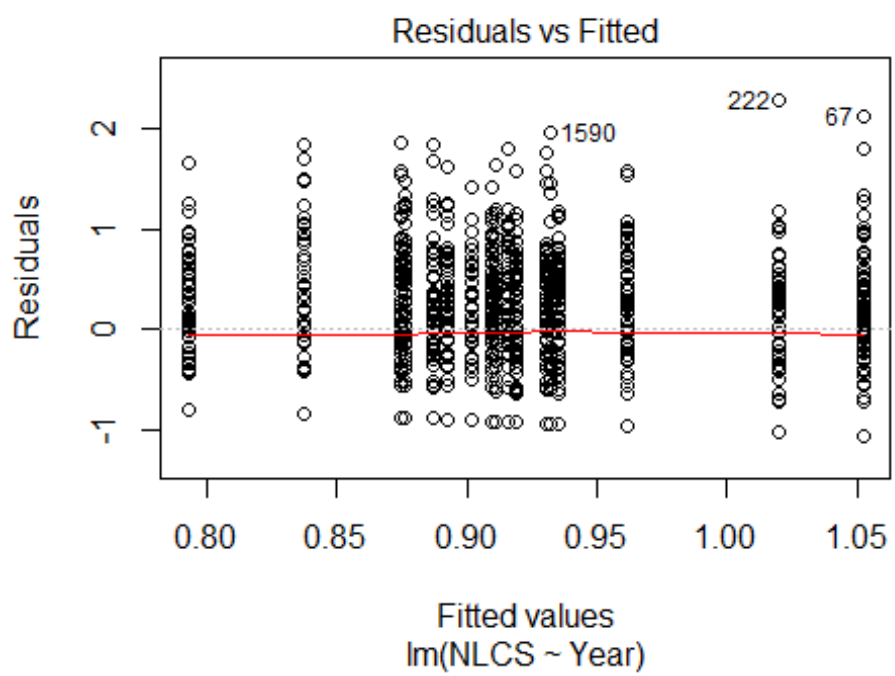
```
## Year2007            -0.1161     0.0616   -1.88   0.0598 .
## Year2008            -0.1054     0.0628   -1.68   0.0936 .
## Year2009            -0.1421     0.0589   -2.42   0.0158 *
## Year2010            -0.0748     0.0601   -1.24   0.2137
## Year2011            -0.1778     0.0597   -2.98   0.0029 **
## Year2012            -0.1388     0.0611   -2.27   0.0231 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.632
## Multiple R-squared:  0.00617,    Adjusted R-squared:  0.00243
## Convergence in 10 IRWLS iterations
##
## Robustness weights:
##  384 weights are ~= 1. The remaining 4158 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.201   0.874   0.950   0.916   0.985   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          2.20e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##                   GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.017  1           1.008
## Year             1.017 16           1.001
```



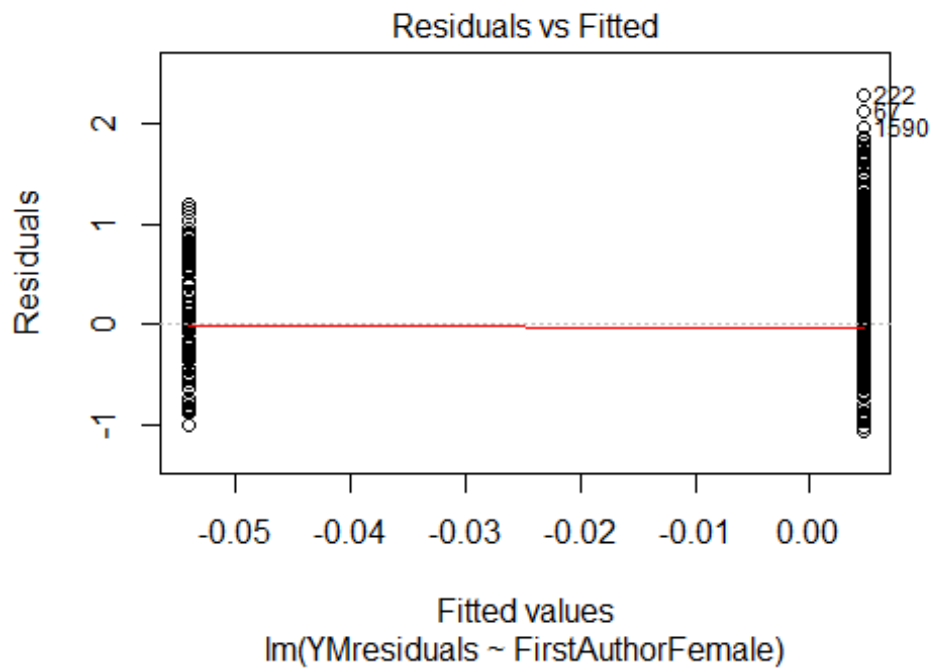

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.0532 -0.4356  0.0157  0.4232  2.1914
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.05317    0.05034   20.92   <2e-16 ***
## LastAuthorFemale1 -0.06624    0.02628   -2.52    0.012 *
## Year1997          -0.03440    0.06770   -0.51    0.611
## Year1998          -0.00866    0.06366   -0.14    0.892
## Year1999          -0.05556    0.06739   -0.82    0.410
## Year2000          -0.07281    0.06384   -1.14    0.254
## Year2001          -0.07412    0.06448   -1.15    0.250
## Year2002          -0.11647    0.06351   -1.83    0.067 .
## Year2003          -0.06764    0.06604   -1.02    0.306
## Year2004          -0.06537    0.06664   -0.98    0.327
## Year2005          -0.04231    0.06217   -0.68    0.496
## Year2006          -0.10789    0.06364   -1.70    0.090 .
```

```
## Year2007          -0.11539    0.06152   -1.88    0.061 .
## Year2008          -0.10145    0.06276   -1.62    0.106
## Year2009          -0.13973    0.05876   -2.38    0.017 *
## Year2010          -0.07252    0.06008   -1.21    0.227
## Year2011          -0.17710    0.05960   -2.97    0.003 **
## Year2012          -0.13653    0.06102   -2.24    0.025 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.631
## Multiple R-squared:  0.00731,    Adjusted R-squared:  0.00358
## Convergence in 10 IRWLS iterations
##
## Robustness weights:
##  387 weights are ~= 1. The remaining 4155 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.203   0.875   0.950   0.916   0.985   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          2.20e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Sample size for the above analysis:  4542"
## [1] ""
## [1] ""
## [1] "####################################"
## [1] "Analysis of AJSC 2609"
## [1] "####################################"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  123  117  130  129  105  126   83   81   95  104  116  110  111  114  121
## 2011 2012
##   93  122
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##   94   87   90   89   78   90   70   65   76   86  102   93   89   95   90
## 2011 2012
```

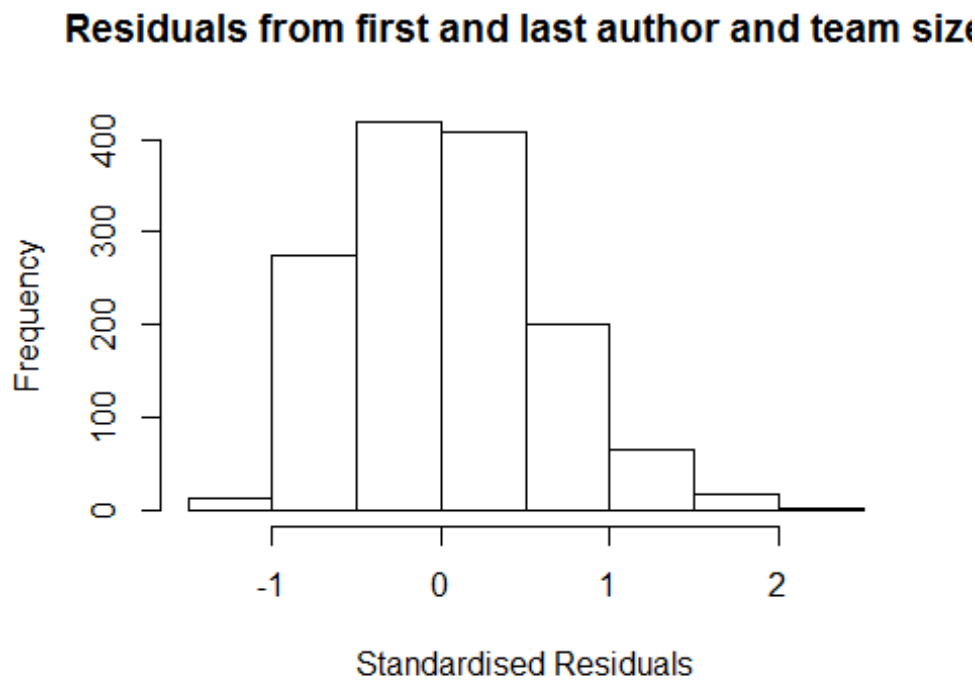
```
##   78   94
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##   86   82   85   87   74   84   67   59   71   82   98   90   84   93   90
## 2011 2012
##   75   92
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
##  Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 18, df = 16, p-value = 0.3
```




```
##
##  Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 0.46, df = 1, p-value = 0.5
```

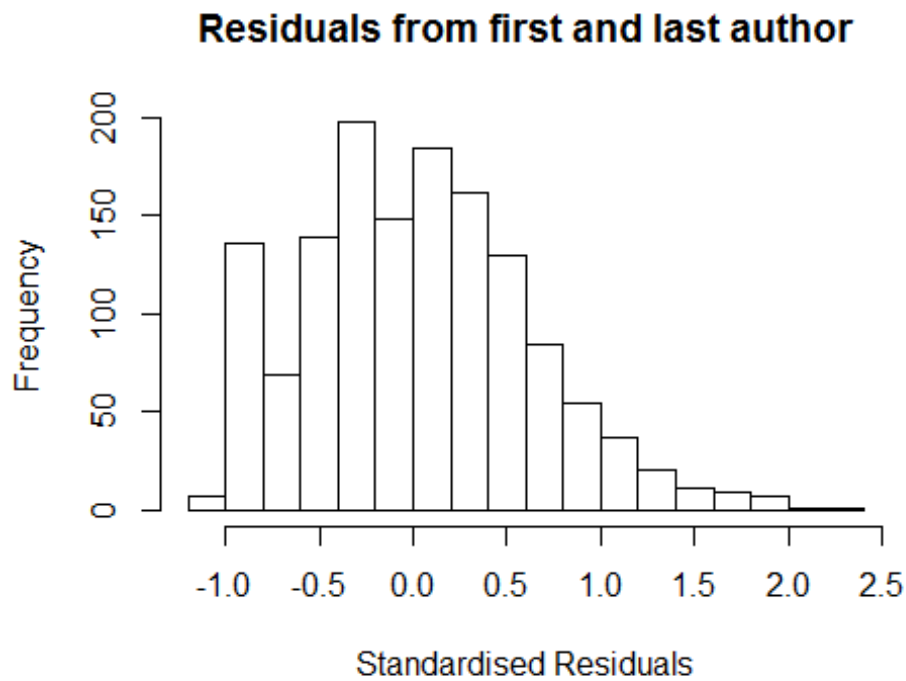



```
## [1] "Female first author team size 2018 geometric mean: 1.34687176593718"
## [1] "Male first author team size 2018 geometric mean: 1.17448314837156"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 570, p-value = 0.1
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 1.33497689498765"
## [1] "Male last author team size 2018 geometric mean: 1.1802381194765"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 460, p-value = 0.5
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##                     GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 2.022  1           1.422
## LastAuthorFemale  2.059  1           1.435
## UniqueAuthors     1.327  4           1.036
## Year              1.275 16           1.008
```



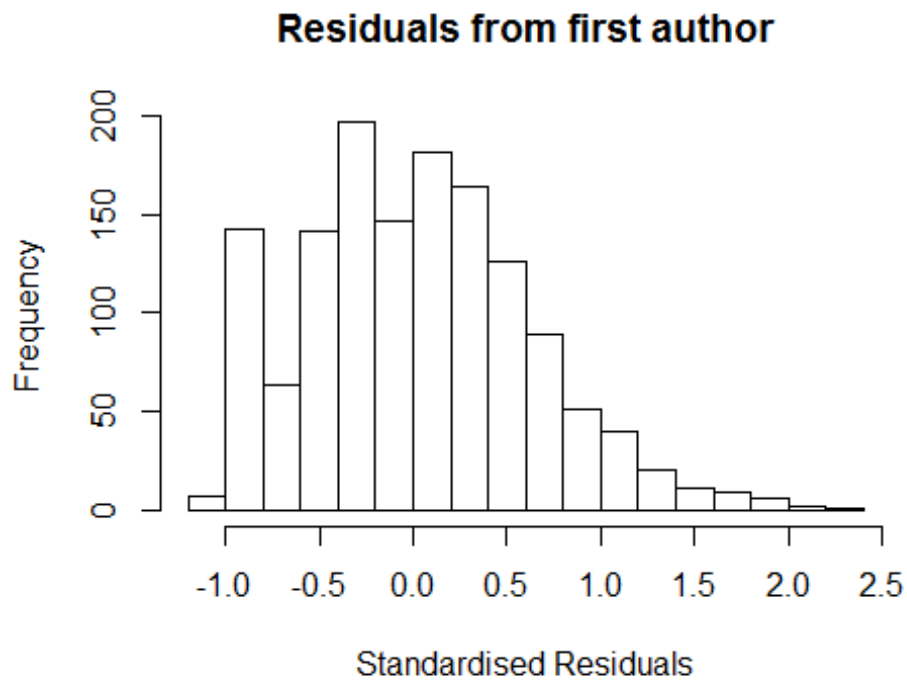

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
UniqueAuthors +
##     Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##     k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.2144 -0.4270 -0.0115  0.3992  2.3484
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         0.95063    0.06801   13.98  < 2e-16 ***
## FirstAuthorFemale1 -0.02407    0.08570   -0.28  0.77886
## LastAuthorFemale1  -0.11216    0.08930   -1.26  0.20935
## UniqueAuthors2      0.14081    0.04579    3.07  0.00215 **
## UniqueAuthors3      0.27986    0.07708    3.63  0.00029 ***
## UniqueAuthors4      0.30138    0.13532    2.23  0.02610 *
## UniqueAuthors5      0.18149    0.18112    1.00  0.31650
## Year1997            0.00794    0.08831    0.09  0.92841
## Year1998           -0.14015    0.10039   -1.40  0.16291
## Year1999           -0.12130    0.09801   -1.24  0.21607
```

```
## Year2000           -0.11257    0.09655   -1.17  0.24384
## Year2001           -0.14888    0.09393   -1.59  0.11317
## Year2002           -0.06248    0.10781   -0.58  0.56233
## Year2003           -0.10516    0.10763   -0.98  0.32873
## Year2004           -0.10064    0.10341   -0.97  0.33060
## Year2005           -0.09642    0.09297   -1.04  0.29987
## Year2006           -0.03759    0.08728   -0.43  0.66678
## Year2007           -0.07767    0.09078   -0.86  0.39234
## Year2008           -0.08639    0.08967   -0.96  0.33553
## Year2009           -0.11413    0.09281   -1.23  0.21903
## Year2010           -0.21908    0.09207   -2.38  0.01747 *
## Year2011           -0.06915    0.08739   -0.79  0.42895
## Year2012           -0.19009    0.10076   -1.89  0.05943 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.616
## Multiple R-squared:  0.0288, Adjusted R-squared:  0.0133
## Convergence in 10 IRWLS iterations
##
## Robustness weights:
##  98 weights are ~= 1. The remaining 1301 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.115   0.866   0.954   0.914   0.986   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          7.15e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.924  1           1.387
## LastAuthorFemale  1.931  1           1.390
## Year              1.056 16           1.002
```



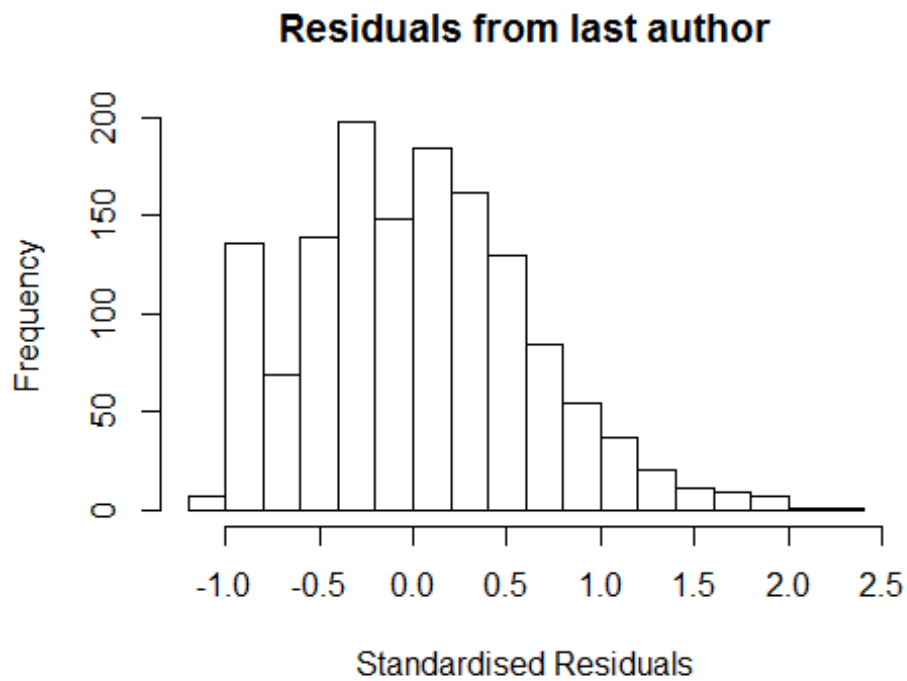

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.01225 -0.40362  0.00269  0.41080  2.31569
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.01225    0.06807   14.87   <2e-16 ***
## FirstAuthorFemale1 -0.00898    0.08471   -0.11    0.916
## LastAuthorFemale1  -0.09615    0.08786   -1.09    0.274
## Year1997           -0.02094    0.09040   -0.23    0.817
## Year1998           -0.17515    0.10044   -1.74    0.081 .
## Year1999           -0.15666    0.09973   -1.57    0.116
## Year2000           -0.12982    0.09853   -1.32    0.188
## Year2001           -0.16922    0.09594   -1.76    0.078 .
## Year2002           -0.06715    0.10880   -0.62    0.537
## Year2003           -0.11124    0.10756   -1.03    0.301
## Year2004           -0.12110    0.10515   -1.15    0.250
## Year2005           -0.10725    0.09544   -1.12    0.261
```

```
## Year2006           -0.06847    0.08905   -0.77    0.442
## Year2007           -0.10849    0.09157   -1.18    0.236
## Year2008           -0.11824    0.08971   -1.32    0.188
## Year2009           -0.12802    0.09310   -1.38    0.169
## Year2010           -0.23101    0.09531   -2.42    0.015 *
## Year2011           -0.09865    0.08852   -1.11    0.265
## Year2012           -0.21367    0.10295   -2.08    0.038 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.622
## Multiple R-squared:  0.0114, Adjusted R-squared:  -0.00149
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  112 weights are ~= 1. The remaining 1287 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.136   0.861   0.952   0.913   0.985   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          7.15e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.022  1           1.011
## Year              1.022 16           1.001
```




```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.0083 -0.4084  0.0035  0.4134  2.3170
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)           1.0083     0.0679   14.85   <2e-16 ***
## FirstAuthorFemale1   -0.0711     0.0627   -1.14    0.257
## Year1997             -0.0183     0.0902   -0.20    0.839
## Year1998             -0.1721     0.1005   -1.71    0.087 .
## Year1999             -0.1525     0.0998   -1.53    0.127
## Year2000             -0.1265     0.0984   -1.29    0.199
## Year2001             -0.1683     0.0961   -1.75    0.080 .
## Year2002             -0.0669     0.1087   -0.62    0.538
## Year2003             -0.1118     0.1081   -1.03    0.301
## Year2004             -0.1186     0.1050   -1.13    0.259
## Year2005             -0.1100     0.0949   -1.16    0.246
## Year2006             -0.0684     0.0886   -0.77    0.440
```

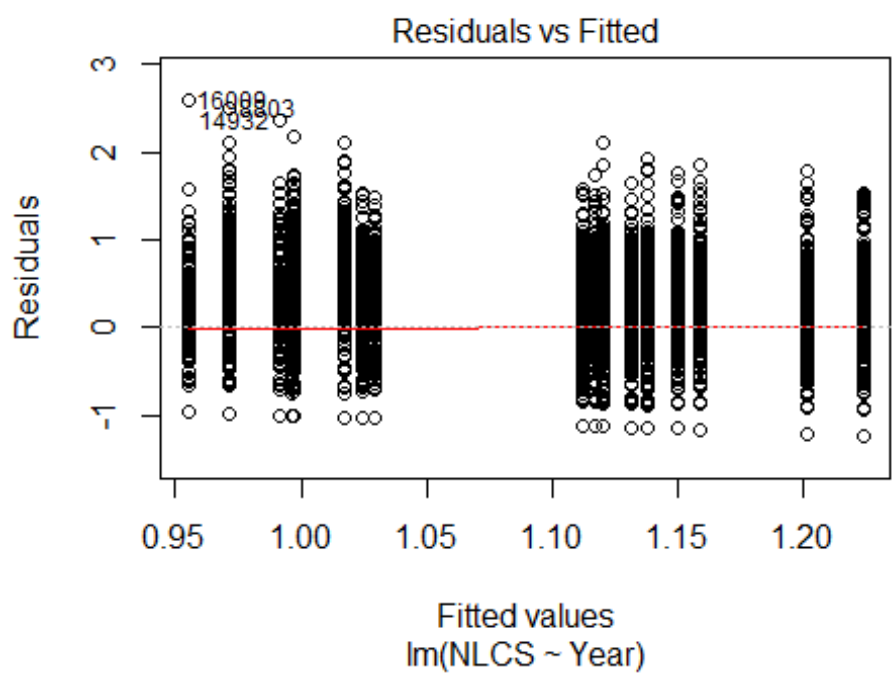
```
## Year2007            -0.1097     0.0914   -1.20    0.231
## Year2008            -0.1167     0.0896   -1.30    0.193
## Year2009            -0.1252     0.0930   -1.35    0.178
## Year2010            -0.2267     0.0954   -2.38    0.018 *
## Year2011            -0.0982     0.0883   -1.11    0.267
## Year2012            -0.2107     0.1034   -2.04    0.042 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.622
## Multiple R-squared:  0.0104, Adjusted R-squared:  -0.00178
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  102 weights are ~= 1. The remaining 1297 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.136   0.861   0.952   0.914   0.986   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          7.15e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##                   GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.028  1           1.014
## Year             1.028 16           1.001
```



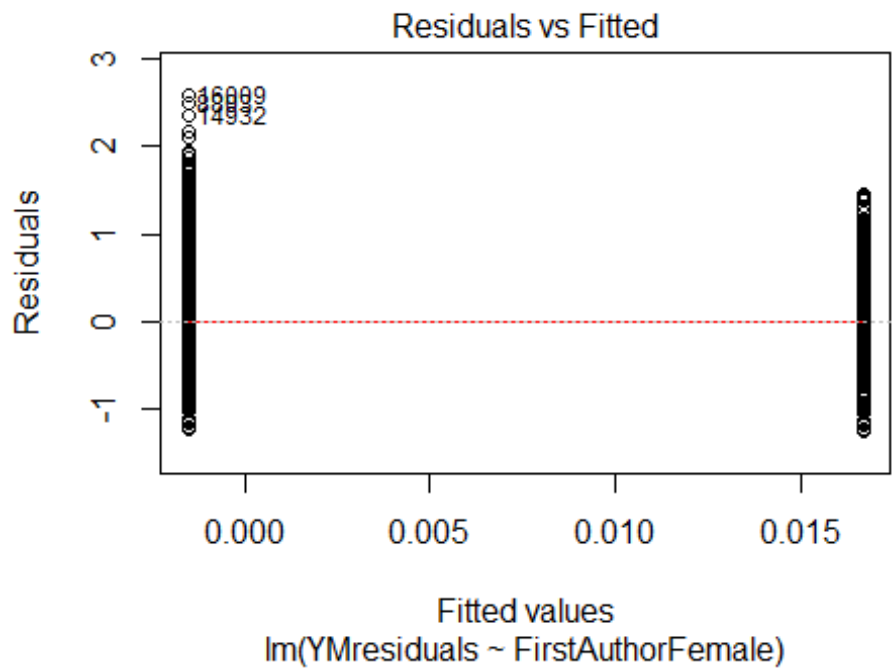

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.01197 -0.40487  0.00307  0.41222  2.31607
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.0120     0.0680   14.88   <2e-16 ***
## LastAuthorFemale1  -0.1021     0.0643   -1.59    0.112
## Year1997           -0.0210     0.0904   -0.23    0.816
## Year1998           -0.1753     0.1004   -1.75    0.081 .
## Year1999           -0.1568     0.0997   -1.57    0.116
## Year2000           -0.1301     0.0985   -1.32    0.187
## Year2001           -0.1691     0.0959   -1.76    0.078 .
## Year2002           -0.0668     0.1087   -0.61    0.539
## Year2003           -0.1113     0.1076   -1.04    0.301
## Year2004           -0.1210     0.1051   -1.15    0.250
## Year2005           -0.1071     0.0955   -1.12    0.262
## Year2006           -0.0684     0.0891   -0.77    0.443
```

```
## Year2007           -0.1082     0.0915   -1.18    0.237
## Year2008           -0.1183     0.0897   -1.32    0.188
## Year2009           -0.1282     0.0931   -1.38    0.169
## Year2010           -0.2312     0.0953   -2.43    0.015 *
## Year2011           -0.0987     0.0885   -1.12    0.265
## Year2012           -0.2141     0.1028   -2.08    0.037 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.621
## Multiple R-squared:  0.0114, Adjusted R-squared:  -0.000762
## Convergence in 10 IRWLS iterations
##
## Robustness weights:
##  111 weights are ~= 1. The remaining 1288 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.135   0.861   0.952   0.913   0.985   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          7.15e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Sample size for the above analysis:  1399"
## [1] ""
## [1] ""
## [1] "####################################"
## [1] "Analysis of AJSC 2610"
## [1] "####################################"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  841  870  826  853  856  936  977 1144  955  767  766  808  717  597  471
## 2011 2012
##  499  528
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  404  408  406  376  408  435  487  556  496  475  455  482  441  361  264
## 2011 2012
```

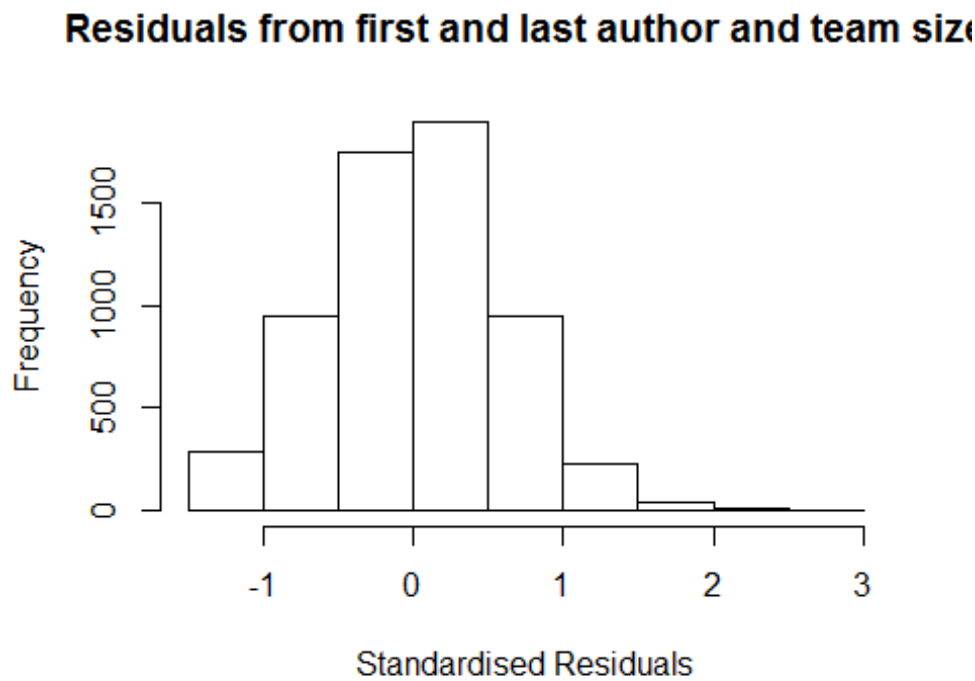
```
##  305  312
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  353  338  339  316  354  376  424  456  416  425  398  432  392  319  232
## 2011 2012
##  277  267
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
##  Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 120, df = 16, p-value <2e-16
```




```
##
##  Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 3.7, df = 1, p-value = 0.06
```

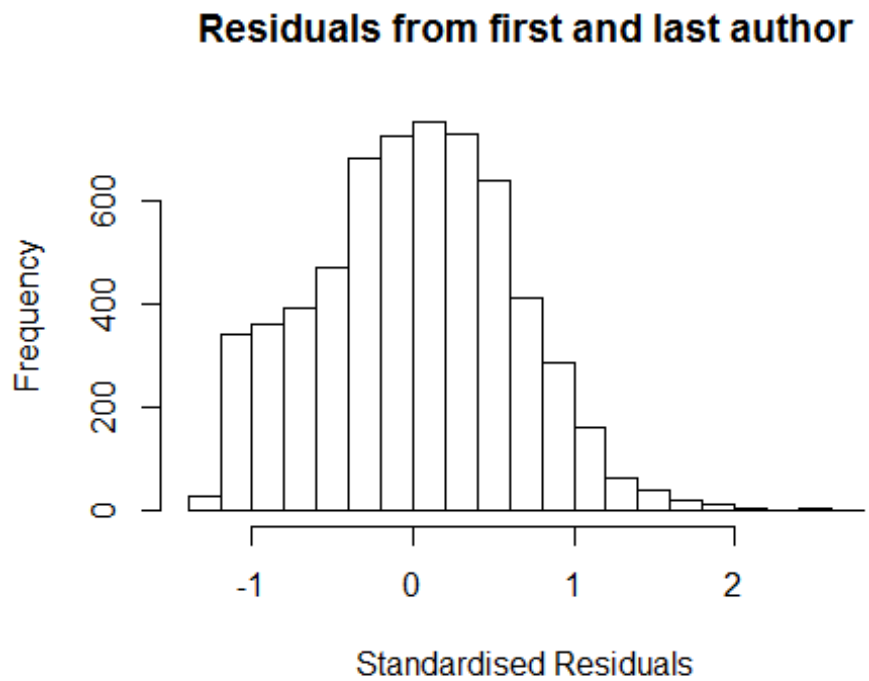



```
## [1] "Female first author team size 2018 geometric mean: 2.2061914938372"
## [1] "Male first author team size 2018 geometric mean: 1.80319377183487"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 4600, p-value = 0.05
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 1.85723506163929"
## [1] "Male last author team size 2018 geometric mean: 1.84071738375705"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 2400, p-value = 0.9
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.185  1           1.088
## LastAuthorFemale  1.169  1           1.081
## UniqueAuthors     1.050  4           1.006
## Year              1.068 16           1.002
```



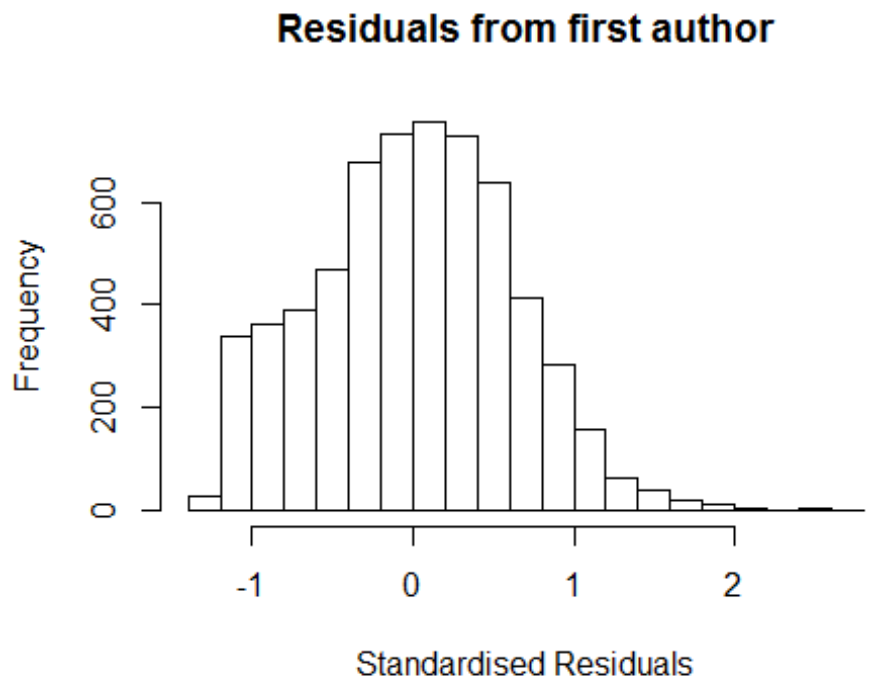

```
## [1] "List of  2 outliers with residuals above  2.5"
##          ScopusId  NLCS Year OneField Fields residuals
## 8803  11244344133 3.457 2004     2610      2     2.686
## 16009 84862158214 3.539 2012     2610      2     2.762
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
UniqueAuthors +
##     Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##     k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.4623 -0.3994  0.0136  0.3973  2.7623
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)          1.0771     0.0372   28.94  < 2e-16 ***
## FirstAuthorFemale1  -0.0143     0.0295   -0.49   0.6264
## LastAuthorFemale1    0.0130     0.0309    0.42   0.6729
## UniqueAuthors2       0.2435     0.0182   13.40  < 2e-16 ***
## UniqueAuthors3       0.3529     0.0235   15.01  < 2e-16 ***
## UniqueAuthors4       0.4838     0.0409   11.83  < 2e-16 ***
## UniqueAuthors5       0.4420     0.0493    8.97  < 2e-16 ***
## Year1997            -0.1407     0.0474   -2.97   0.0030 **
## Year1998            -0.1027     0.0461   -2.23   0.0258 *
```

```
## Year1999            -0.1111     0.0478   -2.32   0.0202 *
## Year2000            -0.0749     0.0445   -1.68   0.0928 .
## Year2001            -0.1964     0.0586   -3.35   0.0008 ***
## Year2002            -0.0986     0.0469   -2.10   0.0355 *
## Year2003            -0.2793     0.0511   -5.47  4.8e-08 ***
## Year2004            -0.3061     0.0522   -5.87  4.6e-09 ***
## Year2005            -0.2370     0.0458   -5.18  2.3e-07 ***
## Year2006            -0.1212     0.0456   -2.66   0.0078 **
## Year2007            -0.1454     0.0450   -3.23   0.0012 **
## Year2008            -0.1480     0.0459   -3.23   0.0013 **
## Year2009            -0.2091     0.0469   -4.46  8.3e-06 ***
## Year2010            -0.2783     0.0520   -5.35  8.9e-08 ***
## Year2011            -0.2159     0.0488   -4.43  9.8e-06 ***
## Year2012            -0.3005     0.0498   -6.03  1.7e-09 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.592
## Multiple R-squared:  0.0831, Adjusted R-squared:  0.0798
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
##  517 weights are ~= 1. The remaining 5597 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0001  0.8610  0.9500  0.9070  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.64e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.152  1           1.073
## LastAuthorFemale  1.145  1           1.070
## Year              1.022 16           1.001
```



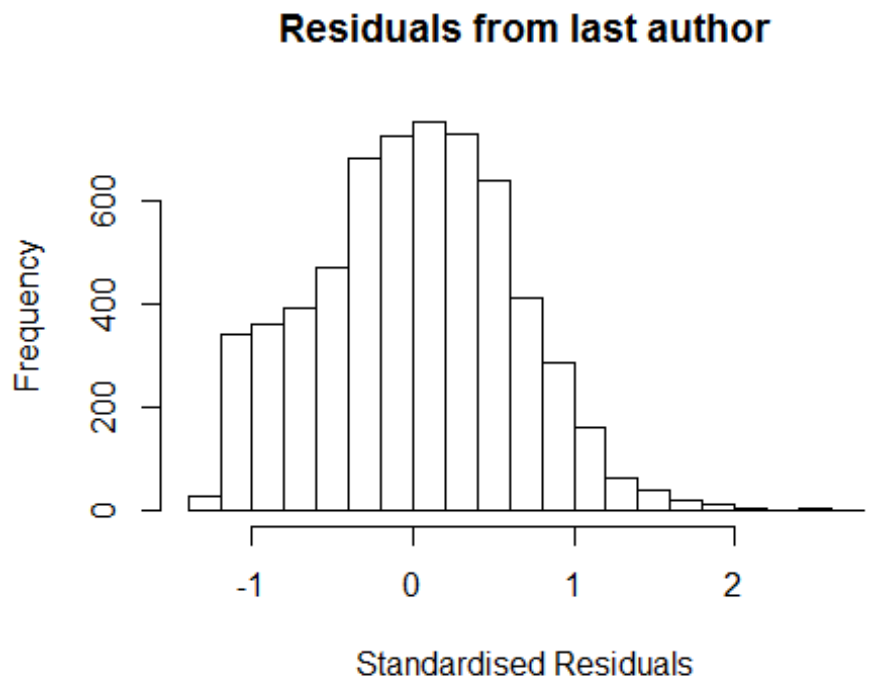

```
## [1] "List of  2 outliers with residuals above  2.5"
##          ScopusId  NLCS Year OneField Fields residuals
## 8803  11244344133 3.457 2004     2610      2     2.511
## 16009 84862158214 3.539 2012     2610      2     2.612
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.2847 -0.4276  0.0168  0.4285  2.6121
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)           1.2366     0.0372   33.27  < 2e-16 ***
## FirstAuthorFemale1    0.0284     0.0311    0.91   0.3618
## LastAuthorFemale1     0.0197     0.0329    0.60   0.5498
## Year1997             -0.1479     0.0497   -2.98   0.0029 **
## Year1998             -0.1128     0.0485   -2.32   0.0202 *
## Year1999             -0.1154     0.0506   -2.28   0.0226 *
## Year2000             -0.0697     0.0465   -1.50   0.1342
## Year2001             -0.2099     0.0580   -3.62   0.0003 ***
## Year2002             -0.0851     0.0482   -1.77   0.0772 .
## Year2003             -0.2514     0.0516   -4.88  1.1e-06 ***
## Year2004             -0.2911     0.0523   -5.57  2.7e-08 ***
```

```
## Year2005            -0.2561     0.0479   -5.35  9.0e-08 ***
## Year2006            -0.1209     0.0476   -2.54   0.0111 *
## Year2007            -0.1398     0.0475   -2.94   0.0033 **
## Year2008            -0.1521     0.0484   -3.14   0.0017 **
## Year2009            -0.2360     0.0489   -4.83  1.4e-06 ***
## Year2010            -0.2980     0.0553   -5.39  7.3e-08 ***
## Year2011            -0.2325     0.0508   -4.57  4.9e-06 ***
## Year2012            -0.3097     0.0509   -6.09  1.2e-09 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.628
## Multiple R-squared:  0.0187, Adjusted R-squared:  0.0158
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  508 weights are ~= 1. The remaining 5606 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.045   0.868   0.950   0.912   0.985   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.64e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.015  1           1.008
## Year              1.015 16           1.000
```




```
## [1] "List of  2 outliers with residuals above  2.5"
##          ScopusId  NLCS Year OneField Fields residuals
## 8803  11244344133 3.457 2004     2610      2     2.511
## 16009 84862158214 3.539 2012     2610      2     2.612
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.2725 -0.4281  0.0186  0.4284  2.6117
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)           1.2374     0.0372   33.30  < 2e-16 ***
## FirstAuthorFemale1    0.0351     0.0291    1.21  0.22752
## Year1997             -0.1476     0.0497   -2.97  0.00297 **
## Year1998             -0.1126     0.0485   -2.32  0.02043 *
## Year1999             -0.1152     0.0506   -2.28  0.02280 *
## Year2000             -0.0696     0.0465   -1.50  0.13466
## Year2001             -0.2095     0.0580   -3.61  0.00031 ***
## Year2002             -0.0853     0.0482   -1.77  0.07659 .
## Year2003             -0.2508     0.0515   -4.86  1.2e-06 ***
## Year2004             -0.2912     0.0523   -5.57  2.7e-08 ***
## Year2005             -0.2559     0.0478   -5.35  9.1e-08 ***
```

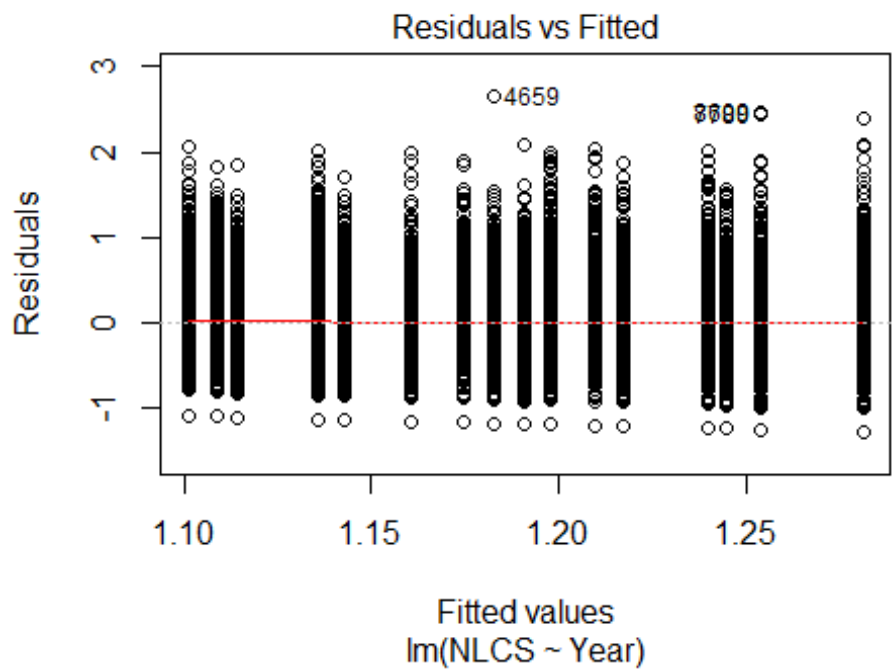
```
## Year2006            -0.1209     0.0476   -2.54  0.01114 *
## Year2007            -0.1395     0.0475   -2.94  0.00331 **
## Year2008            -0.1519     0.0484   -3.14  0.00171 **
## Year2009            -0.2357     0.0489   -4.82  1.5e-06 ***
## Year2010            -0.2977     0.0553   -5.39  7.5e-08 ***
## Year2011            -0.2322     0.0508   -4.57  5.0e-06 ***
## Year2012            -0.3100     0.0509   -6.09  1.2e-09 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.628
## Multiple R-squared:  0.0187, Adjusted R-squared:  0.0159
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  503 weights are ~= 1. The remaining 5611 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0451  0.8690  0.9500  0.9120  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.64e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##                   GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.009  1           1.004
## Year             1.009 16           1.000
```



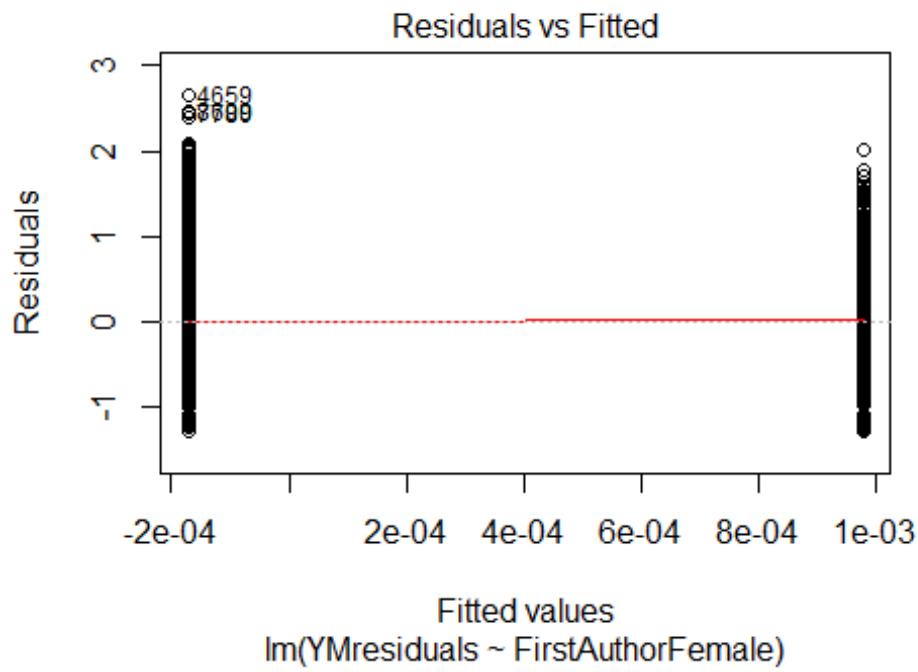

```
## [1] "List of  2 outliers with residuals above  2.5"
##          ScopusId  NLCS Year OneField Fields residuals
## 8803  11244344133 3.457 2004     2610      2     2.511
## 16009 84862158214 3.539 2012     2610      2     2.612
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.2685 -0.4288  0.0206  0.4272  2.6092
##
## Coefficients:
##                   Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.2379     0.0371   33.33  < 2e-16 ***
## LastAuthorFemale1   0.0306     0.0307    1.00  0.31856
## Year1997           -0.1475     0.0497   -2.97  0.00298 **
## Year1998           -0.1127     0.0485   -2.32  0.02020 *
## Year1999           -0.1151     0.0506   -2.28  0.02289 *
## Year2000           -0.0693     0.0465   -1.49  0.13594
## Year2001           -0.2093     0.0580   -3.61  0.00031 ***
## Year2002           -0.0846     0.0482   -1.76  0.07905 .
## Year2003           -0.2514     0.0515   -4.88  1.1e-06 ***
## Year2004           -0.2909     0.0523   -5.56  2.8e-08 ***
## Year2005           -0.2556     0.0478   -5.35  9.3e-08 ***
```

```
## Year2006           -0.1208     0.0476   -2.54  0.01121 *
## Year2007           -0.1389     0.0475   -2.93  0.00343 **
## Year2008           -0.1521     0.0484   -3.14  0.00169 **
## Year2009           -0.2365     0.0488   -4.84  1.3e-06 ***
## Year2010           -0.2975     0.0553   -5.38  7.7e-08 ***
## Year2011           -0.2320     0.0508   -4.57  5.0e-06 ***
## Year2012           -0.3080     0.0508   -6.06  1.4e-09 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.628
## Multiple R-squared:  0.0186, Adjusted R-squared:  0.0159
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  496 weights are ~= 1. The remaining 5618 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0456  0.8690  0.9500  0.9120  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.64e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Sample size for the above analysis:  6114"
## [1] ""
## [1] ""
## [1] "####################################"
## [1] "Analysis of AJSC 2611"
## [1] "####################################"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 1201 1150 1090 1005 1211 1170  956  893  985 1047 1237 1515 1577 1956 1874
## 2011 2012
## 1858 1993
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  600  516  504  485  580  557  474  445  492  488  656  854  936 1164 1070
```

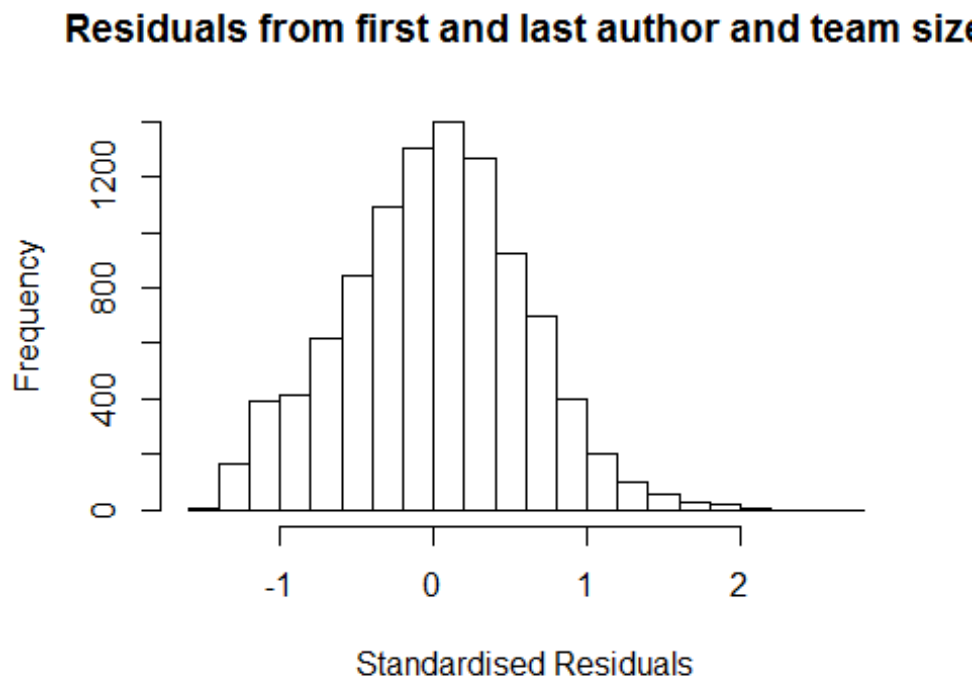
```
## 2011 2012
## 1081 1174
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  503  438  421  410  502  475  392  354  398  388  536  689  767  971  863
## 2011 2012
##  875  955
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
##  Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 140, df = 16, p-value <2e-16
```




```
##
##  Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 3.1, df = 1, p-value = 0.08
```

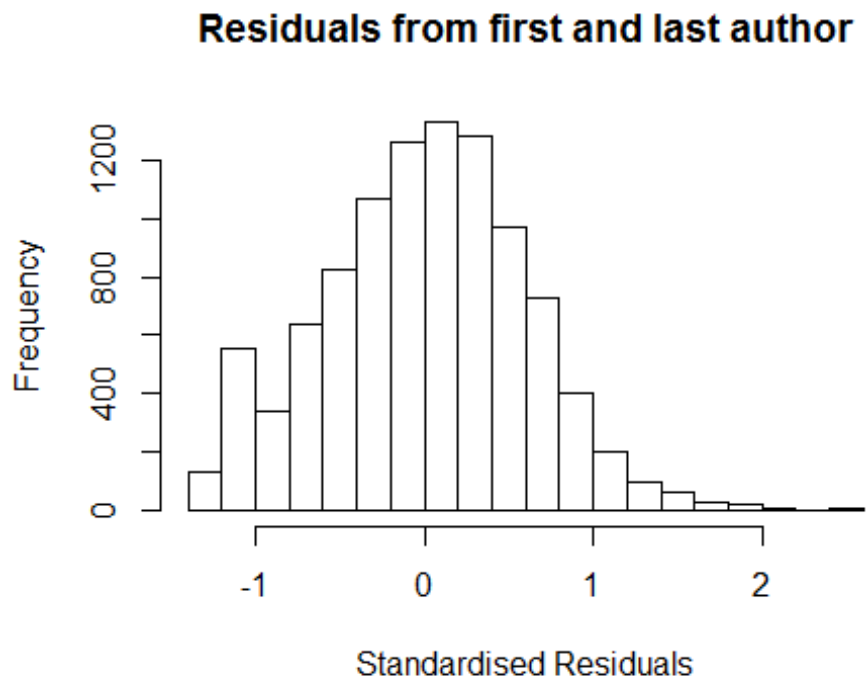



```
## [1] "Female first author team size 2018 geometric mean: 3.00382283050905"
## [1] "Male first author team size 2018 geometric mean: 2.61805715007364"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 110000, p-value = 0.002
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 2.98639777243941"
## [1] "Male last author team size 2018 geometric mean: 2.62998535691198"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 1e+05, p-value = 0.001
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.097  1           1.047
## LastAuthorFemale  1.084  1           1.041
## UniqueAuthors     1.101  4           1.012
## Year              1.127 16           1.004
```



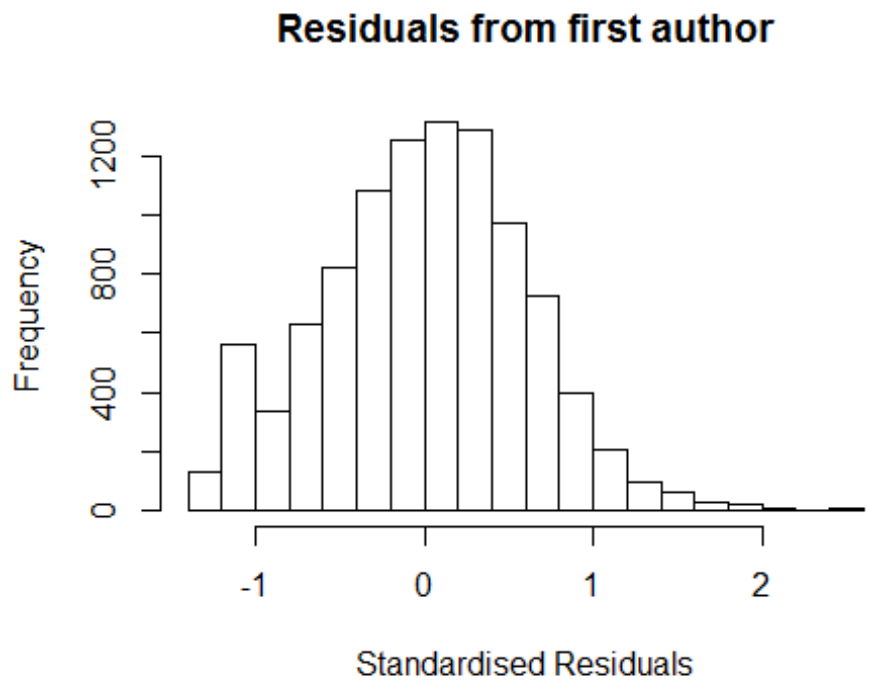

```
## [1] "List of  1 outliers with residuals above  2.5"
##        ScopusId  NLCS Year OneField Fields residuals
## 8609 0036608987 3.709 2002     2207      3     2.613
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
UniqueAuthors +
##     Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##     k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.4241 -0.3904  0.0209  0.3943  2.6134
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.05336    0.03511   30.01   <2e-16 ***
## FirstAuthorFemale1 -0.01300    0.01702   -0.76   0.4448
## LastAuthorFemale1   0.00169    0.01839    0.09   0.9266
## UniqueAuthors2      0.16077    0.01657    9.70   <2e-16 ***
## UniqueAuthors3      0.21900    0.01845   11.87   <2e-16 ***
## UniqueAuthors4      0.29712    0.02329   12.76   <2e-16 ***
## UniqueAuthors5      0.36947    0.02221   16.64   <2e-16 ***
## Year1997            0.02750    0.04815    0.57   0.5678
## Year1998            0.07366    0.04599    1.60   0.1092
## Year1999           -0.00731    0.04574   -0.16   0.8730
```

```
## Year2000           -0.03205    0.04498   -0.71   0.4761
## Year2001            0.07130    0.04604    1.55   0.1215
## Year2002            0.04219    0.04843    0.87   0.3837
## Year2003            0.01629    0.04696    0.35   0.7287
## Year2004           -0.04345    0.04504   -0.96   0.3347
## Year2005           -0.03019    0.04576   -0.66   0.5094
## Year2006            0.00791    0.04238    0.19   0.8519
## Year2007           -0.06599    0.04010   -1.65   0.0999 .
## Year2008           -0.09946    0.03996   -2.49   0.0128 *
## Year2009           -0.08874    0.03899   -2.28   0.0229 *
## Year2010           -0.11903    0.03903   -3.05   0.0023 **
## Year2011           -0.13022    0.03984   -3.27   0.0011 **
## Year2012           -0.12718    0.03898   -3.26   0.0011 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.585
## Multiple R-squared:  0.0408, Adjusted R-squared:  0.0386
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  810 weights are ~= 1. The remaining 9127 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0082  0.8670  0.9510  0.9060  0.9860  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.01e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.081  1           1.040
## LastAuthorFemale  1.069  1           1.034
## Year              1.044 16           1.001
```



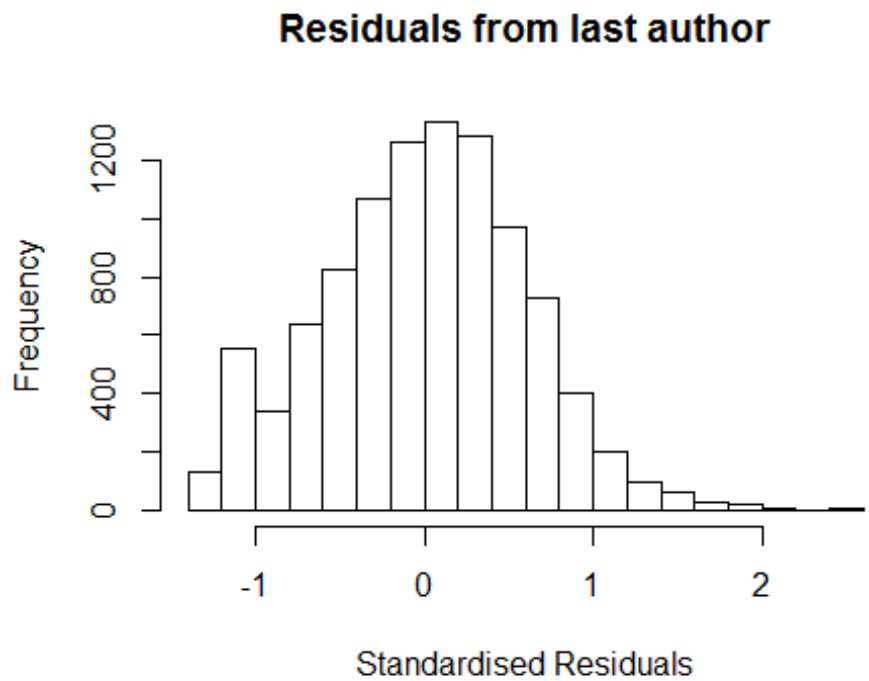

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.2950 -0.4003  0.0221  0.4035  2.4716
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)          1.15586    0.03393   34.06   <2e-16 ***
## FirstAuthorFemale1  0.01172    0.01740    0.67    0.500
## LastAuthorFemale1   0.01267    0.01886    0.67    0.502
## Year1997            0.03786    0.04842    0.78    0.434
## Year1998            0.09992    0.04645    2.15    0.032 *
## Year1999            0.02747    0.04578    0.60    0.548
## Year2000           -0.00092    0.04489   -0.02    0.984
## Year2001            0.11472    0.04625    2.48    0.013 *
## Year2002            0.08153    0.04884    1.67    0.095 .
## Year2003            0.05064    0.04735    1.07    0.285
## Year2004           -0.00460    0.04526   -0.10    0.919
## Year2005            0.01482    0.04576    0.32    0.746
```

```
## Year2006            0.06636    0.04229    1.57    0.117
## Year2007           -0.01800    0.04016   -0.45    0.654
## Year2008           -0.03527    0.03997   -0.88    0.377
## Year2009           -0.02459    0.03899   -0.63    0.528
## Year2010           -0.04611    0.03899   -1.18    0.237
## Year2011           -0.05085    0.03983   -1.28    0.202
## Year2012           -0.04787    0.03888   -1.23    0.218
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.597
## Multiple R-squared:  0.0069, Adjusted R-squared:  0.00509
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
##  836 weights are ~= 1. The remaining 9101 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0481  0.8660  0.9510  0.9060  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.01e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##                   GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.03  1           1.015
## Year              1.03 16           1.001
```




```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.2866 -0.3996  0.0217  0.4047  2.4706
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)           1.1567     0.0339   34.12   <2e-16 ***
## FirstAuthorFemale1    0.0151     0.0170    0.89    0.374
## Year1997              0.0376     0.0484    0.78    0.438
## Year1998              0.0999     0.0464    2.15    0.031 *
## Year1999              0.0275     0.0458    0.60    0.549
## Year2000             -0.0011     0.0449   -0.02    0.980
## Year2001              0.1148     0.0462    2.48    0.013 *
## Year2002              0.0817     0.0488    1.67    0.094 .
## Year2003              0.0509     0.0473    1.08    0.282
## Year2004             -0.0044     0.0453   -0.10    0.923
## Year2005              0.0151     0.0458    0.33    0.742
## Year2006              0.0666     0.0423    1.57    0.115
```

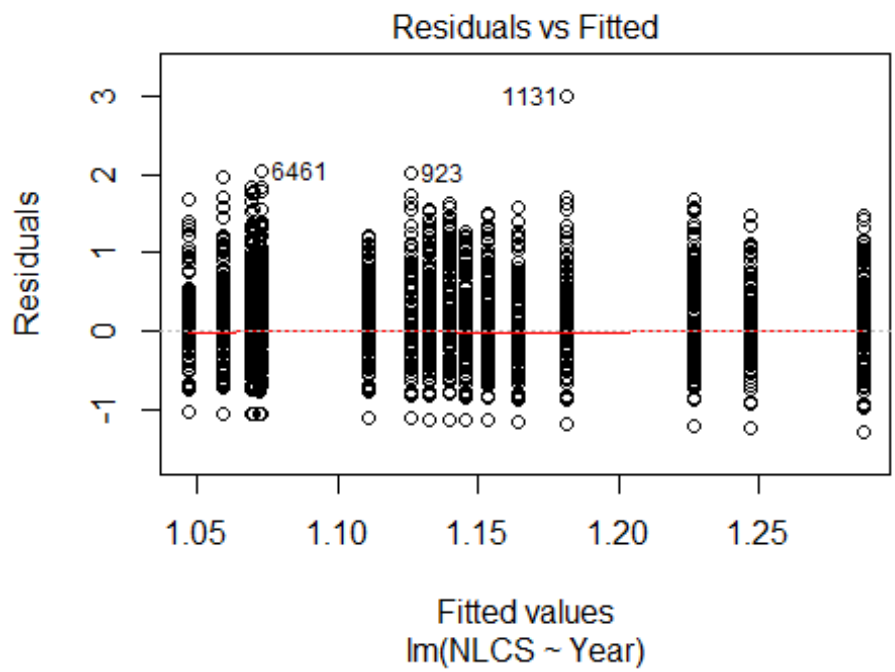
```
## Year2007            -0.0175     0.0401   -0.44    0.663
## Year2008            -0.0353     0.0400   -0.88    0.377
## Year2009            -0.0245     0.0390   -0.63    0.530
## Year2010            -0.0456     0.0390   -1.17    0.242
## Year2011            -0.0504     0.0398   -1.26    0.206
## Year2012            -0.0469     0.0388   -1.21    0.227
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.597
## Multiple R-squared:  0.00685,    Adjusted R-squared:  0.00515
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  837 weights are ~= 1. The remaining 9100 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0485  0.8660  0.9510  0.9060  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.01e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##                   GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.019  1           1.009
## Year             1.019 16           1.001
```



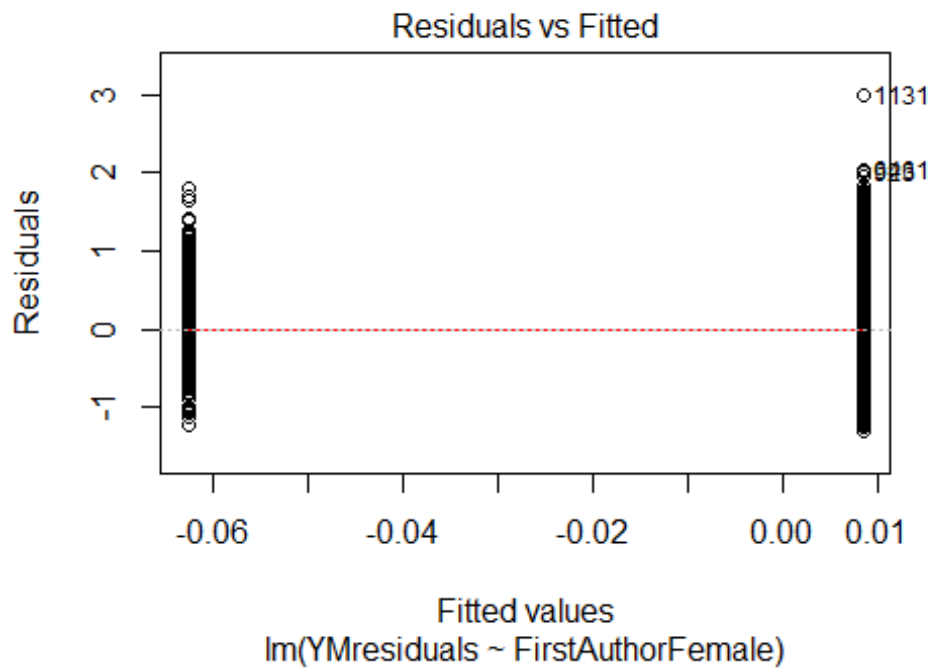

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.2880 -0.3996  0.0222  0.4044  2.4706
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.15681    0.03389   34.14   <2e-16 ***
## LastAuthorFemale1   0.01635    0.01845    0.89    0.376
## Year1997            0.03786    0.04841    0.78    0.434
## Year1998            0.09982    0.04645    2.15    0.032 *
## Year1999            0.02734    0.04576    0.60    0.550
## Year2000           -0.00121    0.04490   -0.03    0.978
## Year2001            0.11480    0.04626    2.48    0.013 *
## Year2002            0.08164    0.04885    1.67    0.095 .
## Year2003            0.05052    0.04735    1.07    0.286
## Year2004           -0.00474    0.04526   -0.10    0.917
## Year2005            0.01498    0.04578    0.33    0.744
## Year2006            0.06663    0.04231    1.57    0.115
```

```
## Year2007          -0.01766    0.04017   -0.44    0.660
## Year2008          -0.03461    0.03996   -0.87    0.387
## Year2009          -0.02383    0.03899   -0.61    0.541
## Year2010          -0.04551    0.03901   -1.17    0.243
## Year2011          -0.04994    0.03980   -1.25    0.210
## Year2012          -0.04686    0.03887   -1.21    0.228
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.597
## Multiple R-squared:  0.00685,    Adjusted R-squared:  0.00514
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
##  838 weights are ~= 1. The remaining 9099 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0483  0.8660  0.9510  0.9060  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.01e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Sample size for the above analysis:  9937"
## [1] ""
## [1] ""
## [1] "####################################"
## [1] "Analysis of AJSC 2612"
## [1] "####################################"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  326  325  281  274  337  266  272  264  312  318  343  367  381  382  350
## 2011 2012
##  303  436
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  165  156  126  137  150  123  131  128  151  162  186  190  206  222  185
## 2011 2012
```

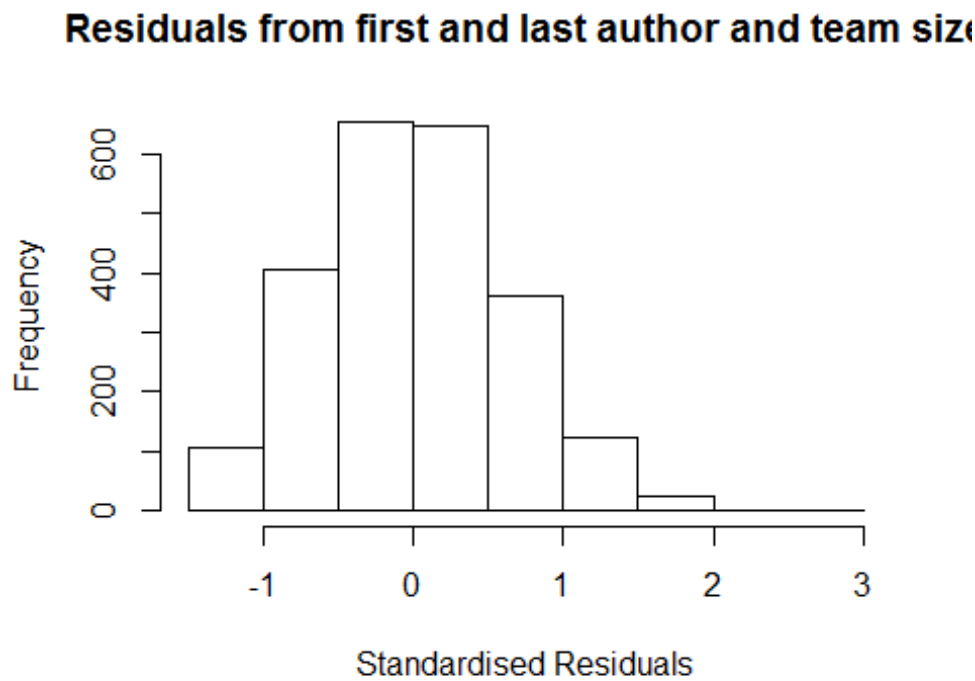
```
##  165  252
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  132  128  105  107  118  104  108  103  124  133  156  159  178  185  146
## 2011 2012
##  137  198
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
##  Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 31, df = 16, p-value = 0.01
```




```
##
##  Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 0.53, df = 1, p-value = 0.5
```

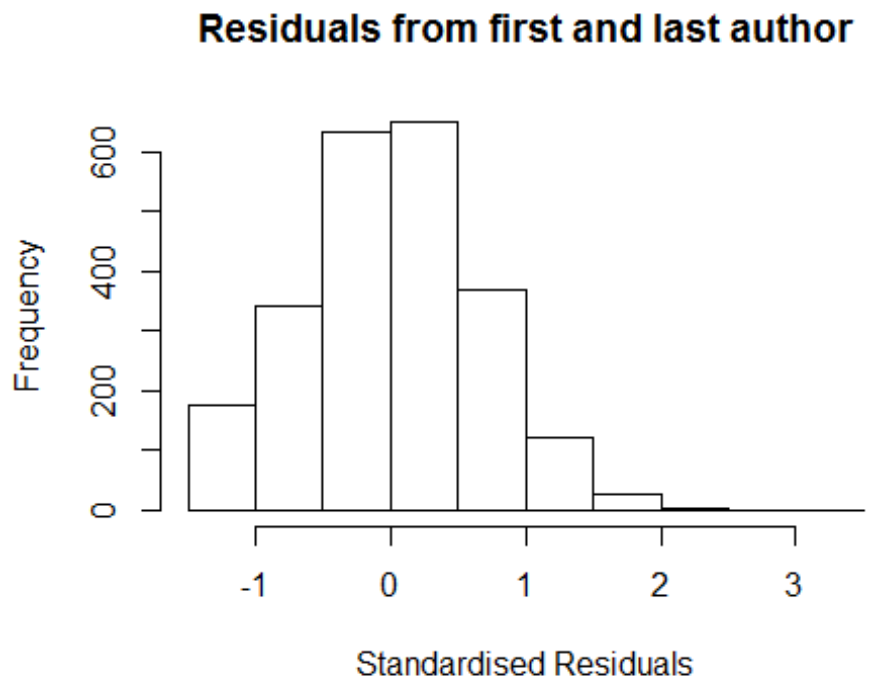



```
## [1] "Female first author team size 2018 geometric mean: 1.90470460043357"
## [1] "Male first author team size 2018 geometric mean: 1.65645468719078"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 1200, p-value = 0.3
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 1.72394564249396"
## [1] "Male last author team size 2018 geometric mean: 1.68586344561181"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 850, p-value = 0.9
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##                     GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.246  1           1.116
## LastAuthorFemale  1.251  1           1.118
## UniqueAuthors     1.154  4           1.018
## Year              1.168 16           1.005
```



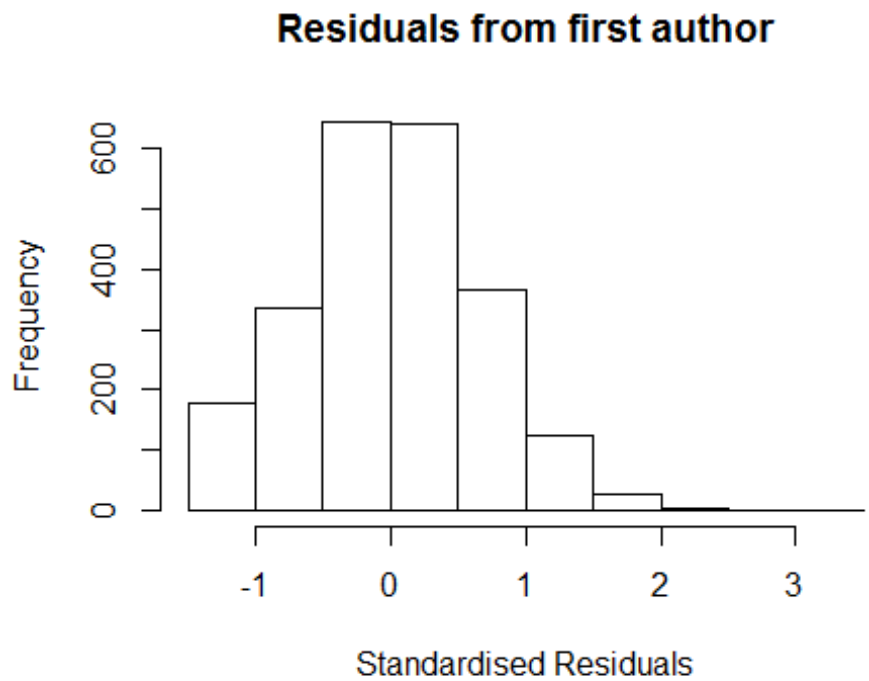

```
## [1] "List of  1 outliers with residuals above  2.5"
##        ScopusId  NLCS Year OneField Fields residuals
## 1131 0033349534 4.175 1999     2200      3     2.848
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
UniqueAuthors +
##     Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##     k.max = 1000))
##  \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.33132 -0.42005 -0.00245  0.42164  2.84769
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)          1.1186     0.0637   17.56  < 2e-16 ***
## FirstAuthorFemale1  -0.0517     0.0454   -1.14   0.2549
## LastAuthorFemale1   -0.1241     0.0498   -2.49   0.0128 *
## UniqueAuthors2       0.2128     0.0300    7.10  1.7e-12 ***
## UniqueAuthors3       0.3764     0.0409    9.20  < 2e-16 ***
## UniqueAuthors4       0.2329     0.0849    2.74   0.0061 **
## UniqueAuthors5       0.4021     0.1034    3.89   0.0001 ***
## Year1997            -0.1694     0.0909   -1.86   0.0625 .
## Year1998            -0.1424     0.0855   -1.67   0.0957 .
## Year1999            -0.1677     0.0913   -1.84   0.0664 .
```

```
## Year2000            -0.0624     0.0848   -0.74   0.4623
## Year2001            -0.2047     0.0882   -2.32   0.0204 *
## Year2002            -0.0400     0.0858   -0.47   0.6414
## Year2003            -0.1005     0.0858   -1.17   0.2417
## Year2004            -0.1173     0.0842   -1.39   0.1638
## Year2005            -0.1350     0.0791   -1.71   0.0880 .
## Year2006            -0.0895     0.0780   -1.15   0.2515
## Year2007            -0.2157     0.0825   -2.61   0.0090 **
## Year2008            -0.1715     0.0743   -2.31   0.0212 *
## Year2009            -0.2221     0.0750   -2.96   0.0031 **
## Year2010            -0.1505     0.0838   -1.80   0.0726 .
## Year2011            -0.2861     0.0832   -3.44   0.0006 ***
## Year2012            -0.2060     0.0812   -2.54   0.0113 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.63
## Multiple R-squared:  0.06,   Adjusted R-squared:  0.051
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  205 weights are ~= 1. The remaining 2116 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0047  0.8720  0.9500  0.9090  0.9860  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          4.31e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.221  1           1.105
## LastAuthorFemale  1.229  1           1.109
## Year              1.036 16           1.001
```



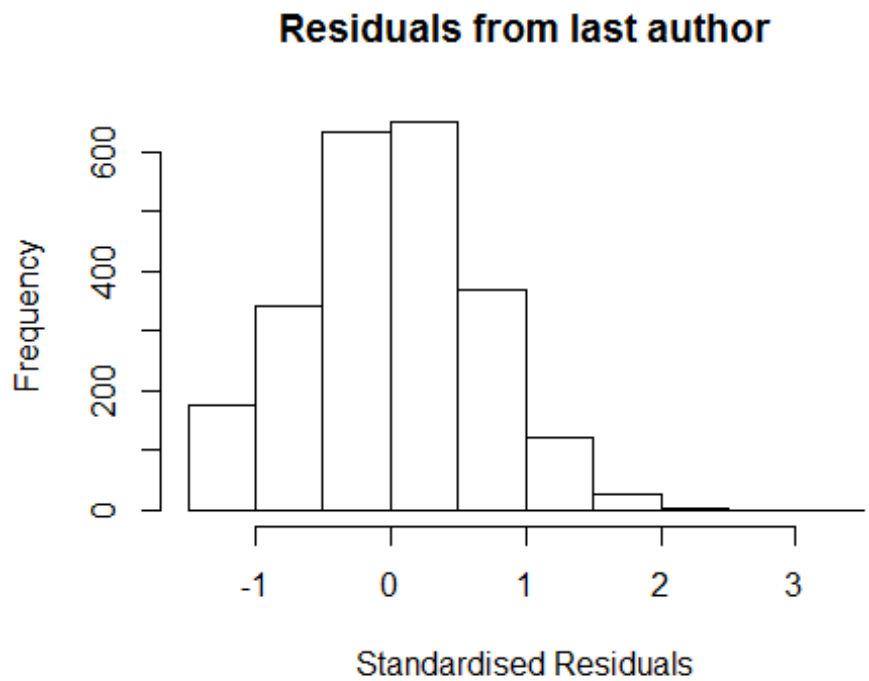

```
## [1] "List of  1 outliers with residuals above  2.5"
##        ScopusId  NLCS Year OneField Fields residuals
## 1131 0033349534 4.175 1999     2200      3     3.089
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.23419 -0.43111  0.00689  0.43263  3.08867
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)           1.2342     0.0624   19.79   <2e-16 ***
## FirstAuthorFemale1   -0.0135     0.0459   -0.29   0.7691
## LastAuthorFemale1    -0.1406     0.0509   -2.76   0.0058 **
## Year1997             -0.1684     0.0926   -1.82   0.0691 .
## Year1998             -0.1195     0.0883   -1.35   0.1762
## Year1999             -0.1479     0.0933   -1.59   0.1130
## Year2000             -0.0471     0.0863   -0.55   0.5854
## Year2001             -0.1965     0.0892   -2.20   0.0278 *
## Year2002             -0.0298     0.0867   -0.34   0.7309
## Year2003             -0.0985     0.0871   -1.13   0.2586
## Year2004             -0.0944     0.0870   -1.09   0.2780
## Year2005             -0.1011     0.0795   -1.27   0.2036
```

```
## Year2006            -0.0653     0.0797   -0.82   0.4126
## Year2007            -0.1972     0.0830   -2.38   0.0176 *
## Year2008            -0.1388     0.0751   -1.85   0.0645 .
## Year2009            -0.1904     0.0763   -2.50   0.0126 *
## Year2010            -0.1229     0.0854   -1.44   0.1502
## Year2011            -0.2318     0.0856   -2.71   0.0068 **
## Year2012            -0.1639     0.0823   -1.99   0.0467 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.648
## Multiple R-squared:  0.0133, Adjusted R-squared:  0.00561
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  observation 378 is an outlier with |weight| = 0 ( < 4.3e-05);
##  186 weights are ~= 1. The remaining 2134 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.302   0.873   0.951   0.911   0.986   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          4.31e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.015  1           1.008
## Year              1.015 16           1.000
```




```
## [1] "List of  1 outliers with residuals above  2.5"
##        ScopusId  NLCS Year OneField Fields residuals
## 1131 0033349534 4.175 1999     2200      3     3.089
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##       Min        1Q    Median        3Q       Max
## -1.23e+00 -4.35e-01  9.25e-05  4.35e-01  3.10e+00
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)           1.2330     0.0621   19.85   <2e-16 ***
## FirstAuthorFemale1   -0.0684     0.0427   -1.60   0.1097
## Year1997             -0.1713     0.0927   -1.85   0.0647 .
## Year1998             -0.1288     0.0881   -1.46   0.1442
## Year1999             -0.1583     0.0930   -1.70   0.0889 .
## Year2000             -0.0512     0.0863   -0.59   0.5529
## Year2001             -0.2054     0.0890   -2.31   0.0210 *
## Year2002             -0.0326     0.0865   -0.38   0.7060
## Year2003             -0.1073     0.0870   -1.23   0.2175
## Year2004             -0.1036     0.0868   -1.19   0.2328
## Year2005             -0.1061     0.0789   -1.34   0.1790
## Year2006             -0.0753     0.0793   -0.95   0.3423
```

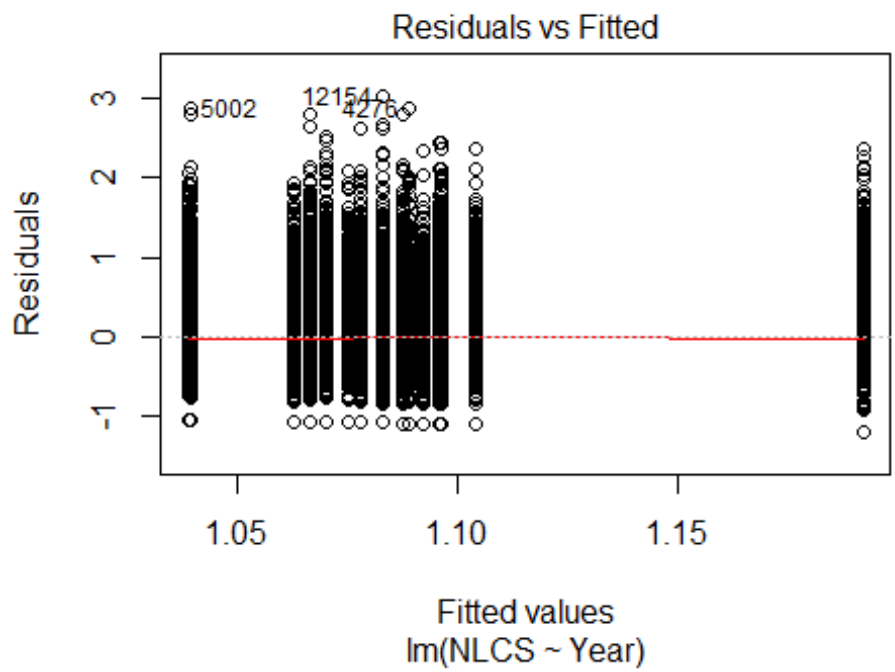
```
## Year2007            -0.2018     0.0829   -2.43   0.0151 *
## Year2008            -0.1427     0.0748   -1.91   0.0567 .
## Year2009            -0.1925     0.0763   -2.52   0.0117 *
## Year2010            -0.1251     0.0850   -1.47   0.1414
## Year2011            -0.2323     0.0857   -2.71   0.0067 **
## Year2012            -0.1684     0.0822   -2.05   0.0407 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.648
## Multiple R-squared:  0.0102, Adjusted R-squared:  0.00292
## Convergence in 10 IRWLS iterations
##
## Robustness weights:
##  observation 378 is an outlier with |weight| = 0 ( < 4.3e-05);
##  187 weights are ~= 1. The remaining 2133 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.299   0.873   0.950   0.911   0.986   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          4.31e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##                  GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.02  1           1.010
## Year             1.02 16           1.001
```



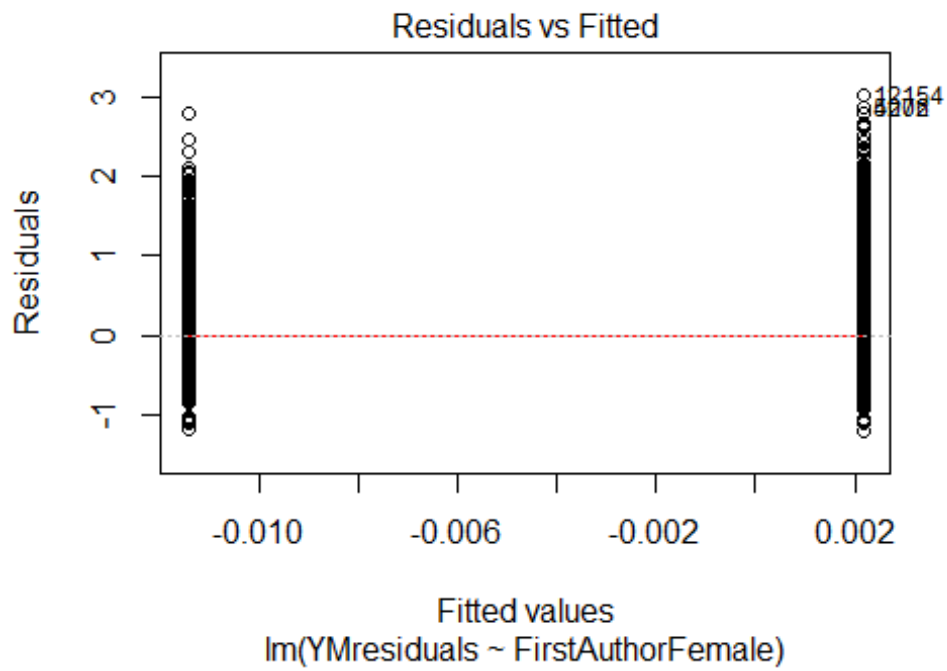

```
## [1] "List of  1 outliers with residuals above  2.5"
##        ScopusId  NLCS Year OneField Fields residuals
## 1131 0033349534 4.175 1999     2200      3     3.089
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.2331 -0.4302  0.0071  0.4337  3.0887
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.2331     0.0622   19.82   <2e-16 ***
## LastAuthorFemale1  -0.1473     0.0465   -3.17   0.0015 **
## Year1997           -0.1677     0.0925   -1.81   0.0699 .
## Year1998           -0.1196     0.0883   -1.35   0.1760
## Year1999           -0.1469     0.0932   -1.58   0.1153
## Year2000           -0.0470     0.0863   -0.54   0.5862
## Year2001           -0.1962     0.0893   -2.20   0.0281 *
## Year2002           -0.0302     0.0868   -0.35   0.7279
## Year2003           -0.0982     0.0871   -1.13   0.2601
## Year2004           -0.0938     0.0869   -1.08   0.2808
## Year2005           -0.1013     0.0796   -1.27   0.2032
## Year2006           -0.0650     0.0797   -0.82   0.4151
```

```
## Year2007           -0.1972     0.0831   -2.37   0.0177 *
## Year2008           -0.1393     0.0752   -1.85   0.0640 .
## Year2009           -0.1906     0.0763   -2.50   0.0126 *
## Year2010           -0.1230     0.0854   -1.44   0.1500
## Year2011           -0.2320     0.0857   -2.71   0.0068 **
## Year2012           -0.1638     0.0824   -1.99   0.0468 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.648
## Multiple R-squared:  0.0133, Adjusted R-squared:  0.006
## Convergence in 10 IRWLS iterations
##
## Robustness weights:
##  observation 378 is an outlier with |weight| = 0 ( < 4.3e-05);
##  184 weights are ~= 1. The remaining 2136 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.301   0.874   0.951   0.911   0.986   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          4.31e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                    psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Sample size for the above analysis:  2321"
## [1] ""
## [1] ""
## [1] "####################################"
## [1] "Analysis of AJSC 2613"
## [1] "####################################"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 1391 1340 1219 1288 1266 1671 1559 1383 1650 2009 1597 2367 2482 2743 2650
## 2011 2012
## 2793 2765
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  927  850  763  820  769  888  963  854  958 1135  909 1354 1448 1577 1508
```

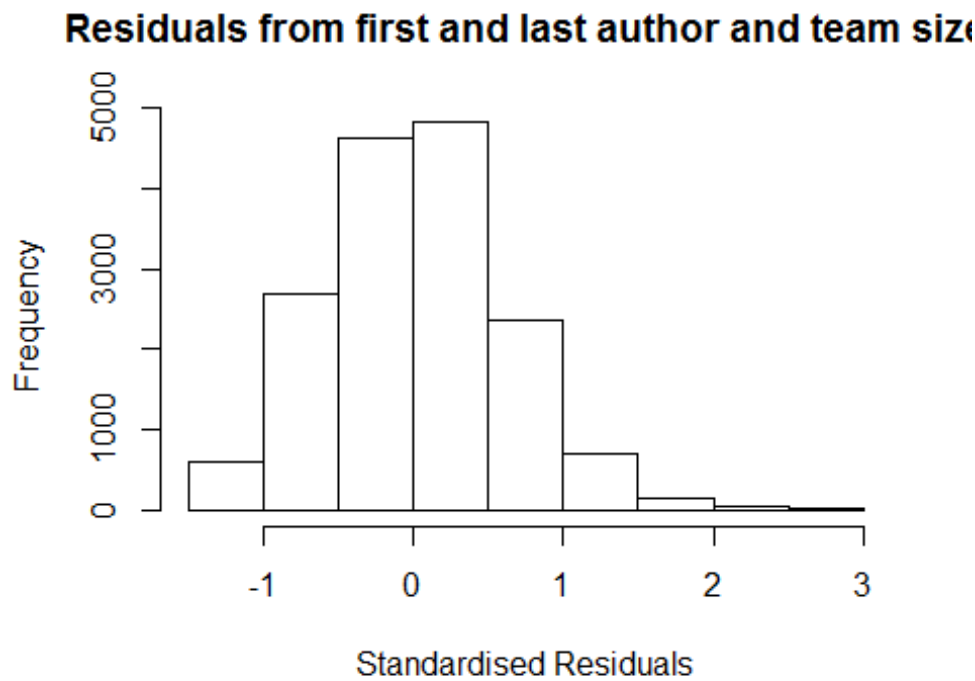
```
## 2011 2012
## 1656 1599
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  816  730  673  729  689  751  837  739  811  941  762 1133 1207 1316 1207
## 2011 2012
## 1338 1318
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
##  Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 110, df = 16, p-value = 3e-15
```




```
##
##  Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 17, df = 1, p-value = 4e-05
```



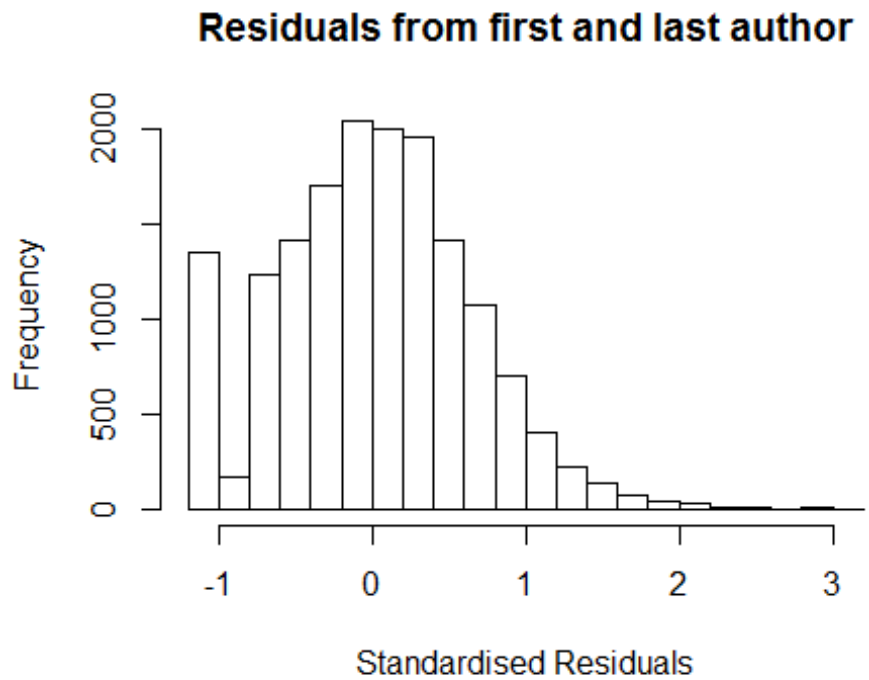

```
## [1] "Female first author team size 2018 geometric mean: 2.84920487301146"
## [1] "Male first author team size 2018 geometric mean: 2.18658994650781"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 140000, p-value = 5e-09
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 2.68580503803789"
## [1] "Male last author team size 2018 geometric mean: 2.23896695656134"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 120000, p-value = 2e-04
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.129  1           1.063
## LastAuthorFemale  1.115  1           1.056
## UniqueAuthors     1.062  4           1.008
## Year              1.049 16           1.001
```




```
## [1] "List of  11 outliers with residuals above  2.5"
##          ScopusId  NLCS Year OneField Fields residuals
## 3053   0031361662 3.465 1997     2613      1     2.528
## 4276   0031773680 3.959 1998     1303      6     2.807
## 5002   1542532754 3.921 1999     1804      2     2.827
## 5587   0001259111 3.825 1999     1804      3     2.649
## 7855   0035470889 3.537 2001     1804      2     2.633
## 9697   0036020892 3.891 2002     2600      2     2.989
## 12154  0141607824 4.104 2003     1702      4     2.868
## 19707 34548599794 3.702 2007     2613      2     2.830
## 26873 65449136284 3.759 2009     1303      7     2.552
## 27745 58449116301 3.389 2009     1804      4     2.519
## 31296 77950537175 3.862 2010     1712      3     2.670
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
UniqueAuthors +
##     Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##     k.max = 1000))
##  \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.37704 -0.41006  0.00786  0.40327  2.98910
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
```

```
## (Intercept)          1.0400     0.0259   40.22  < 2e-16 ***
## FirstAuthorFemale1  -0.0454     0.0136   -3.34  0.00084 ***
## LastAuthorFemale1   -0.0284     0.0147   -1.93  0.05390 .
## UniqueAuthors2       0.2252     0.0119   18.93  < 2e-16 ***
## UniqueAuthors3       0.3370     0.0142   23.70  < 2e-16 ***
## UniqueAuthors4       0.3532     0.0203   17.43  < 2e-16 ***
## UniqueAuthors5       0.4386     0.0216   20.28  < 2e-16 ***
## Year1997            -0.1031     0.0363   -2.84  0.00451 **
## Year1998            -0.1136     0.0369   -3.08  0.00206 **
## Year1999            -0.1717     0.0353   -4.87  1.1e-06 ***
## Year2000            -0.1030     0.0359   -2.87  0.00408 **
## Year2001            -0.1362     0.0347   -3.92  8.8e-05 ***
## Year2002            -0.1381     0.0326   -4.24  2.3e-05 ***
## Year2003            -0.1407     0.0330   -4.26  2.0e-05 ***
## Year2004            -0.1294     0.0336   -3.85  0.00012 ***
## Year2005            -0.1572     0.0321   -4.89  1.0e-06 ***
## Year2006            -0.1395     0.0334   -4.17  3.0e-05 ***
## Year2007            -0.1682     0.0307   -5.48  4.4e-08 ***
## Year2008            -0.1465     0.0303   -4.83  1.4e-06 ***
## Year2009            -0.1701     0.0301   -5.66  1.5e-08 ***
## Year2010            -0.1845     0.0308   -5.99  2.1e-09 ***
## Year2011            -0.1923     0.0304   -6.32  2.7e-10 ***
## Year2012            -0.2201     0.0312   -7.05  1.9e-12 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.605
## Multiple R-squared:  0.0579, Adjusted R-squared:  0.0566
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  2 observations c(4498,5589) are outliers with |weight| = 0 ( < 6.3e-06);
##  1337 weights are ~= 1. The remaining 14658 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.000   0.864   0.951   0.907   0.986   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          6.25e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
```

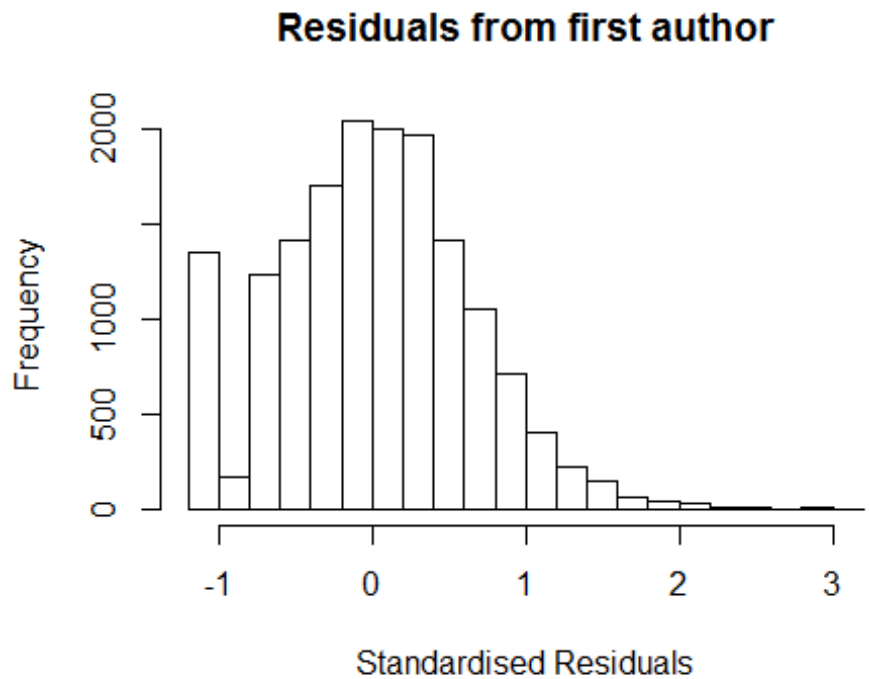
```
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##                     GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.097  1           1.047
## LastAuthorFemale  1.096  1           1.047
## Year              1.006 16           1.000
```




```
## [1] "List of  10 outliers with residuals above  2.5"
##          ScopusId  NLCS Year OneField Fields residuals
## 4276   0031773680 3.959 1998     1303      6     2.903
## 5002   1542532754 3.921 1999     1804      2     2.923
## 5587   0001259111 3.825 1999     1804      3     2.828
## 9697   0036020892 3.891 2002     2600      2     2.829
## 12154  0141607824 4.104 2003     1702      4     3.040
## 12632  0142121516 3.710 2003     1804      2     2.646
## 19707 34548599794 3.702 2007     2613      2     2.659
## 26873 65449136284 3.759 2009     1303      7     2.701
## 31296 77950537175 3.862 2010     1712      3     2.821
## 32754 80053558787 3.593 2011     1702      4     2.552
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
```

```
##      Min       1Q   Median       3Q      Max
## -1.15618 -0.42306  0.00765  0.41324  3.04024
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.15618    0.02611   44.28  < 2e-16 ***
## FirstAuthorFemale1 -0.00105    0.01384   -0.08  0.93959
## LastAuthorFemale1  -0.01268    0.01509   -0.84  0.40078
## Year1997           -0.09078    0.03736   -2.43  0.01512 *
## Year1998           -0.09973    0.03808   -2.62  0.00883 **
## Year1999           -0.15838    0.03600   -4.40  1.1e-05 ***
## Year2000           -0.10112    0.03686   -2.74  0.00608 **
## Year2001           -0.09447    0.03587   -2.63  0.00844 **
## Year2002           -0.09434    0.03336   -2.83  0.00470 **
## Year2003           -0.09242    0.03382   -2.73  0.00628 **
## Year2004           -0.06691    0.03431   -1.95  0.05118 .
## Year2005           -0.09961    0.03310   -3.01  0.00262 **
## Year2006           -0.08076    0.03422   -2.36  0.01827 *
## Year2007           -0.11346    0.03152   -3.60  0.00032 ***
## Year2008           -0.07969    0.03117   -2.56  0.01059 *
## Year2009           -0.09829    0.03077   -3.19  0.00140 **
## Year2010           -0.11488    0.03157   -3.64  0.00027 ***
## Year2011           -0.11475    0.03127   -3.67  0.00024 ***
## Year2012           -0.13379    0.03193   -4.19  2.8e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.625
## Multiple R-squared:  0.00239,    Adjusted R-squared:  0.00127
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  2 observations c(2438,5589)
##   are outliers with |weight| <= 6.1e-06 ( < 6.3e-06);
##  1363 weights are ~= 1. The remaining 14632 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0003  0.8730  0.9510  0.9080  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          6.25e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
```

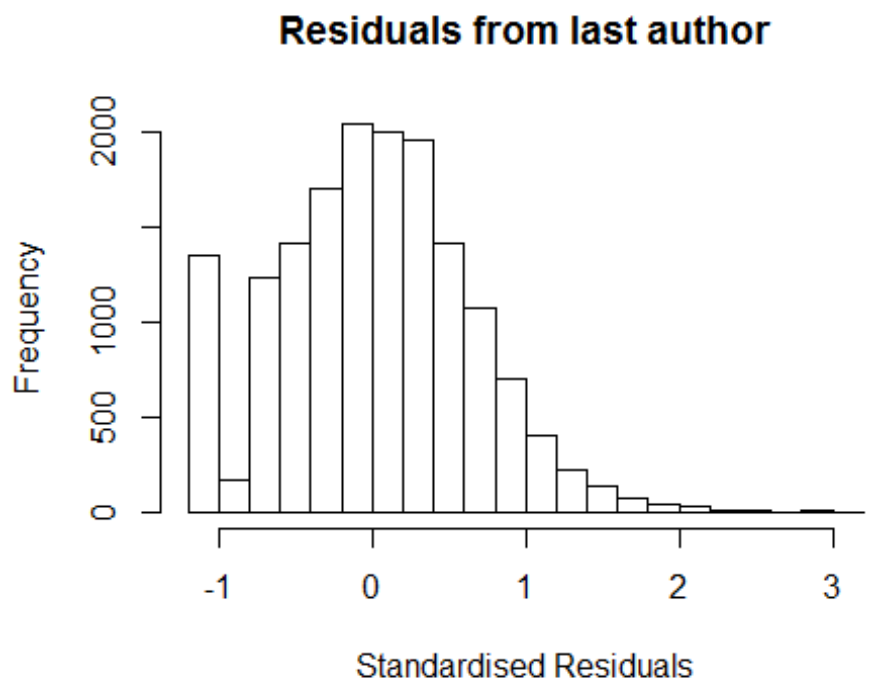
```
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.003  1           1.002
## Year              1.003 16           1.000
```




```
## [1] "List of  10 outliers with residuals above  2.5"
##          ScopusId  NLCS Year OneField Fields residuals
## 4276   0031773680 3.959 1998     1303      6     2.903
## 5002   1542532754 3.921 1999     1804      2     2.923
## 5587   0001259111 3.825 1999     1804      3     2.828
## 9697   0036020892 3.891 2002     2600      2     2.829
## 12154  0141607824 4.104 2003     1702      4     3.040
## 12632  0142121516 3.710 2003     1804      2     2.646
## 19707 34548599794 3.702 2007     2613      2     2.659
## 26873 65449136284 3.759 2009     1303      7     2.701
## 31296 77950537175 3.862 2010     1712      3     2.821
## 32754 80053558787 3.593 2011     1702      4     2.552
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
```

```
##      Min       1Q   Median       3Q      Max
## -1.15523 -0.42199  0.00757  0.41357  3.04136
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.15523    0.02609   44.28  < 2e-16 ***
## FirstAuthorFemale1 -0.00496    0.01327   -0.37  0.70835
## Year1997           -0.09084    0.03736   -2.43  0.01506 *
## Year1998           -0.09994    0.03808   -2.62  0.00870 **
## Year1999           -0.15841    0.03600   -4.40  1.1e-05 ***
## Year2000           -0.10134    0.03686   -2.75  0.00598 **
## Year2001           -0.09466    0.03586   -2.64  0.00831 **
## Year2002           -0.09424    0.03337   -2.82  0.00475 **
## Year2003           -0.09259    0.03382   -2.74  0.00619 **
## Year2004           -0.06687    0.03431   -1.95  0.05135 .
## Year2005           -0.09985    0.03309   -3.02  0.00256 **
## Year2006           -0.08095    0.03421   -2.37  0.01798 *
## Year2007           -0.11379    0.03152   -3.61  0.00031 ***
## Year2008           -0.07988    0.03118   -2.56  0.01041 *
## Year2009           -0.09828    0.03078   -3.19  0.00141 **
## Year2010           -0.11511    0.03157   -3.65  0.00027 ***
## Year2011           -0.11513    0.03127   -3.68  0.00023 ***
## Year2012           -0.13388    0.03193   -4.19  2.8e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.625
## Multiple R-squared:  0.00235,    Adjusted R-squared:  0.00129
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  2 observations c(2438,5589)
##   are outliers with |weight| <= 2.8e-06 ( < 6.3e-06);
##  1343 weights are ~= 1. The remaining 14652 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0002  0.8730  0.9510  0.9080  0.9860  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          6.25e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
```

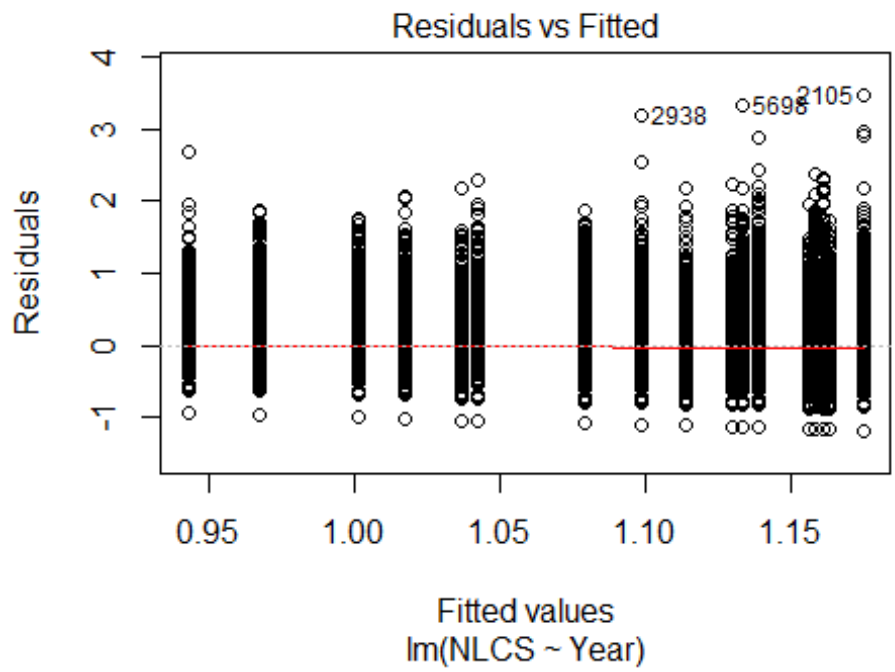
```
##                    "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##                     GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.002  1           1.001
## Year             1.002 16           1.000
```



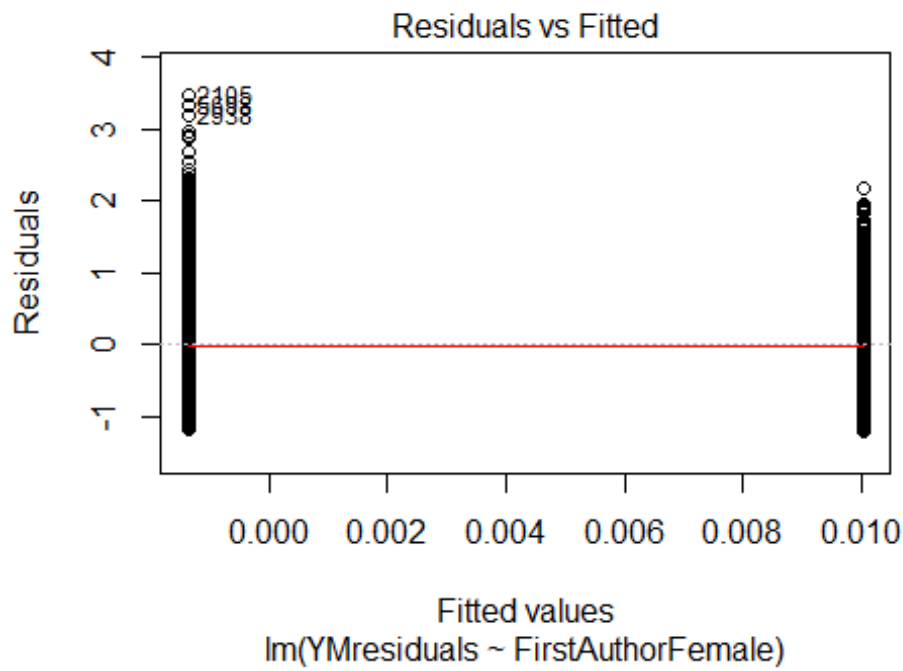

```
## [1] "List of  10 outliers with residuals above  2.5"
##          ScopusId  NLCS Year OneField Fields residuals
## 4276   0031773680 3.959 1998     1303      6     2.903
## 5002   1542532754 3.921 1999     1804      2     2.923
## 5587   0001259111 3.825 1999     1804      3     2.828
## 9697   0036020892 3.891 2002     2600      2     2.829
## 12154  0141607824 4.104 2003     1702      4     3.040
## 12632  0142121516 3.710 2003     1804      2     2.646
## 19707 34548599794 3.702 2007     2613      2     2.659
## 26873 65449136284 3.759 2009     1303      7     2.701
## 31296 77950537175 3.862 2010     1712      3     2.821
## 32754 80053558787 3.593 2011     1702      4     2.552
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
```

```
## -1.1561 -0.4230  0.0077  0.4134  3.0404
##
## Coefficients:
##                   Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.1561     0.0261   44.34  < 2e-16 ***
## LastAuthorFemale1  -0.0131     0.0144   -0.90  0.36607
## Year1997           -0.0908     0.0374   -2.43  0.01511 *
## Year1998           -0.0998     0.0381   -2.62  0.00880 **
## Year1999           -0.1584     0.0360   -4.40  1.1e-05 ***
## Year2000           -0.1011     0.0369   -2.74  0.00608 **
## Year2001           -0.0945     0.0359   -2.64  0.00841 **
## Year2002           -0.0944     0.0334   -2.83  0.00468 **
## Year2003           -0.0925     0.0338   -2.73  0.00626 **
## Year2004           -0.0670     0.0343   -1.95  0.05101 .
## Year2005           -0.0996     0.0331   -3.01  0.00261 **
## Year2006           -0.0808     0.0342   -2.36  0.01825 *
## Year2007           -0.1135     0.0315   -3.60  0.00032 ***
## Year2008           -0.0797     0.0312   -2.56  0.01054 *
## Year2009           -0.0983     0.0308   -3.20  0.00140 **
## Year2010           -0.1149     0.0316   -3.64  0.00027 ***
## Year2011           -0.1148     0.0313   -3.67  0.00024 ***
## Year2012           -0.1338     0.0319   -4.19  2.8e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.625
## Multiple R-squared:  0.00239,    Adjusted R-squared:  0.00133
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  2 observations c(2438,5589)
##   are outliers with |weight| <= 5.3e-06 ( < 6.3e-06);
##  1364 weights are ~= 1. The remaining 14631 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0003  0.8730  0.9510  0.9080  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          6.25e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
```

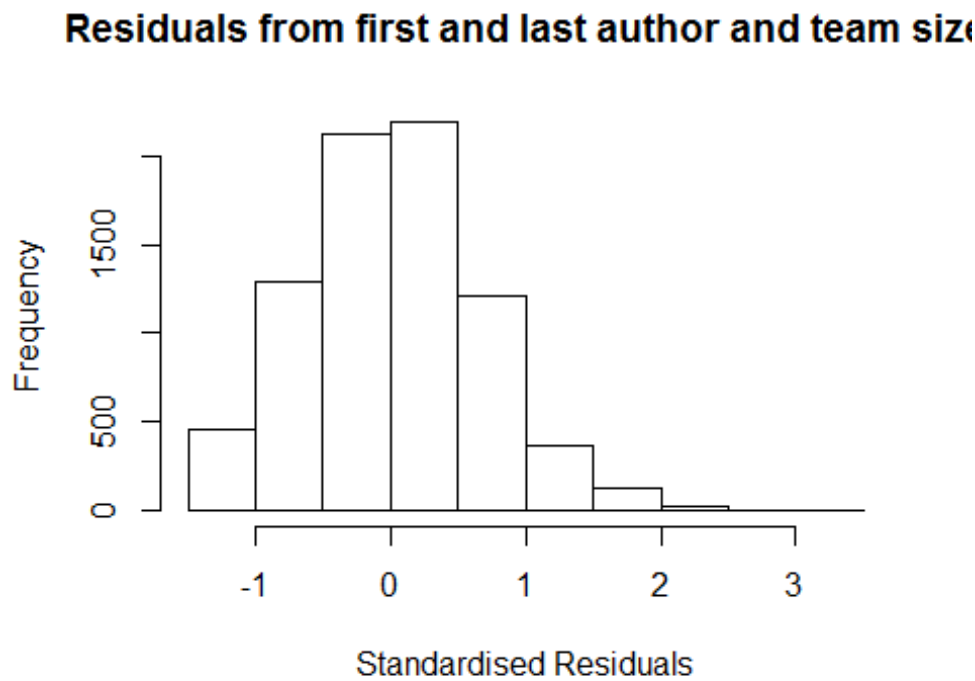
```
## seed : int(0)
## [1] "Sample size for the above analysis:  15997"
## [1] ""
## [1] ""
## [1] "####################################"
## [1] "Analysis of AJSC 2614"
## [1] "####################################"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 1110  943  809  788  975  921  927  634  688  728  792  901  879  928  743
## 2011 2012
##  768  786
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  710  547  486  508  622  535  614  396  464  457  526  616  596  618  491
## 2011 2012
##  503  534
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  606  451  427  426  544  460  505  319  405  388  439  525  492  539  406
## 2011 2012
##  419  441
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
##  Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 130, df = 16, p-value <2e-16
```




```
##
##  Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 0.18, df = 1, p-value = 0.7
```



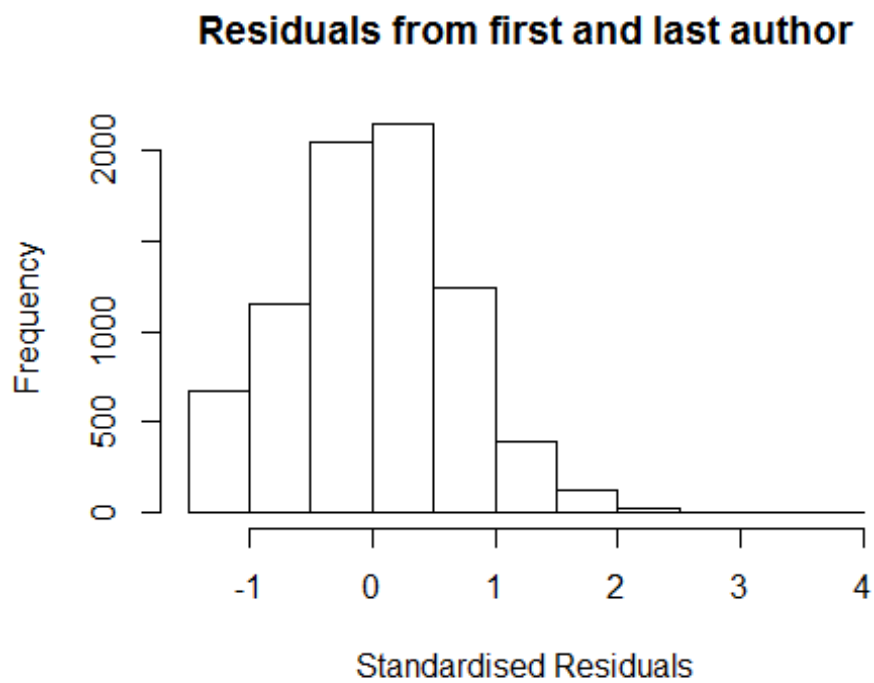

```
## [1] "Female first author team size 2018 geometric mean: 2.56607246992228"
## [1] "Male first author team size 2018 geometric mean: 1.98215006634008"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 12000, p-value = 0.002
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 2.10352926056085"
## [1] "Male last author team size 2018 geometric mean: 2.05883357903293"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 7200, p-value = 1
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.169  1           1.081
## LastAuthorFemale  1.171  1           1.082
## UniqueAuthors     1.054  4           1.007
## Year              1.058 16           1.002
```




```
## [1] "List of  7 outliers with residuals above  2.5"
##          ScopusId  NLCS Year OneField Fields residuals
## 1357   0031118203 4.089 1997     1703      3     2.917
## 2068   0031153727 4.130 1997     1708      3     2.958
## 2105   0031211090 4.640 1997     1703      4     3.468
## 2916   0002806618 3.648 1998     1702      5     2.518
## 2938   0032251894 4.293 1998     1712      2     3.072
## 5698   0035455653 4.472 2001     1708      3     3.275
## 14679 78549288866 3.635 2010     2604      3     2.592
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
UniqueAuthors +
##     Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##     k.max = 1000))
##  \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.35355 -0.43254  0.00727  0.44019  3.46837
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)          1.00189    0.03221   31.10  < 2e-16 ***
## FirstAuthorFemale1   0.02913    0.02481    1.17  0.24044
## LastAuthorFemale1   -0.02881    0.02435   -1.18  0.23676
## UniqueAuthors2       0.16825    0.01777    9.47  < 2e-16 ***
```

```
## UniqueAuthors3      0.24941    0.02343   10.64  < 2e-16 ***
## UniqueAuthors4      0.25945    0.03153    8.23  < 2e-16 ***
## UniqueAuthors5      0.33841    0.03566    9.49  < 2e-16 ***
## Year1997            0.00149    0.04833    0.03  0.97539
## Year1998           -0.04063    0.04484   -0.91  0.36488
## Year1999           -0.01588    0.04786   -0.33  0.74012
## Year2000           -0.05544    0.04360   -1.27  0.20352
## Year2001           -0.05467    0.04661   -1.17  0.24086
## Year2002           -0.01336    0.04270   -0.31  0.75432
## Year2003            0.00360    0.04623    0.08  0.93796
## Year2004           -0.04659    0.04345   -1.07  0.28358
## Year2005            0.00990    0.04471    0.22  0.82473
## Year2006           -0.02376    0.04304   -0.55  0.58086
## Year2007           -0.11902    0.03971   -3.00  0.00273 **
## Year2008           -0.11933    0.04379   -2.72  0.00645 **
## Year2009           -0.14372    0.04052   -3.55  0.00039 ***
## Year2010           -0.20825    0.04359   -4.78  1.8e-06 ***
## Year2011           -0.13836    0.04463   -3.10  0.00194 **
## Year2012           -0.22553    0.04307   -5.24  1.7e-07 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.657
## Multiple R-squared:  0.0386, Adjusted R-squared:  0.0358
## Convergence in 10 IRWLS iterations
##
## Robustness weights:
##  3 observations c(945,1326,2578)
##   are outliers with |weight| <= 6.2e-06 ( < 1.3e-05);
##  667 weights are ~= 1. The remaining 7122 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0056  0.8690  0.9510  0.9100  0.9860  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.28e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
```

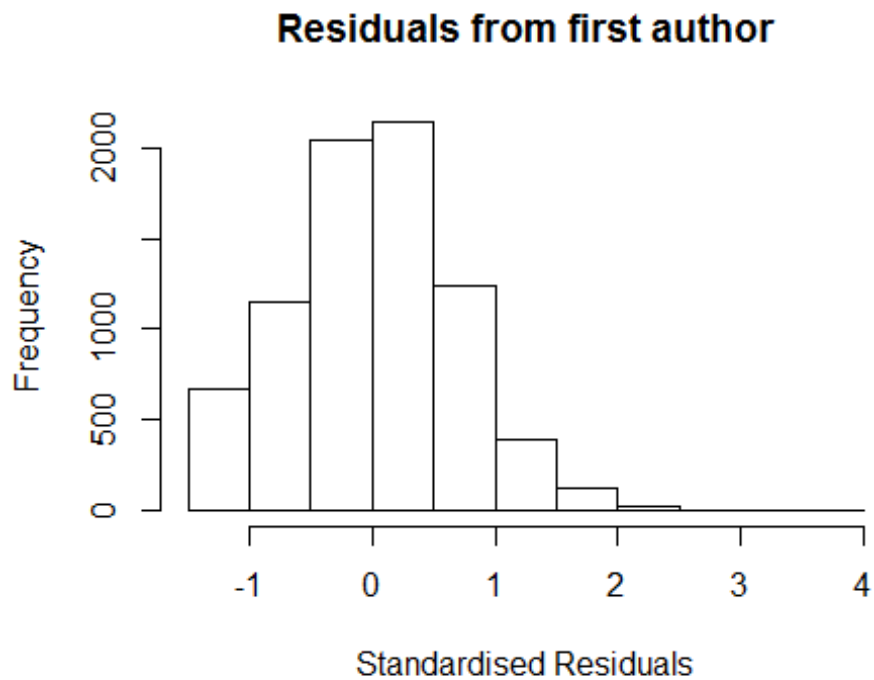
```
##                     GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.154  1           1.074
## LastAuthorFemale  1.155  1           1.075
## Year              1.017 16           1.001
```




```
## [1] "List of  7 outliers with residuals above  2.5"
##          ScopusId  NLCS Year OneField Fields residuals
## 1357   0031118203 4.089 1997     1703      3     2.982
## 2068   0031153727 4.130 1997     1708      3     3.023
## 2105   0031211090 4.640 1997     1703      4     3.533
## 2916   0002806618 3.648 1998     1702      5     2.580
## 2938   0032251894 4.293 1998     1712      2     3.225
## 5698   0035455653 4.472 2001     1708      3     3.394
## 14679 78549288866 3.635 2010     2604      3     2.720
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.1786 -0.4546  0.0056  0.4541  3.5334
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)          1.10040    0.03219   34.18  < 2e-16 ***
```

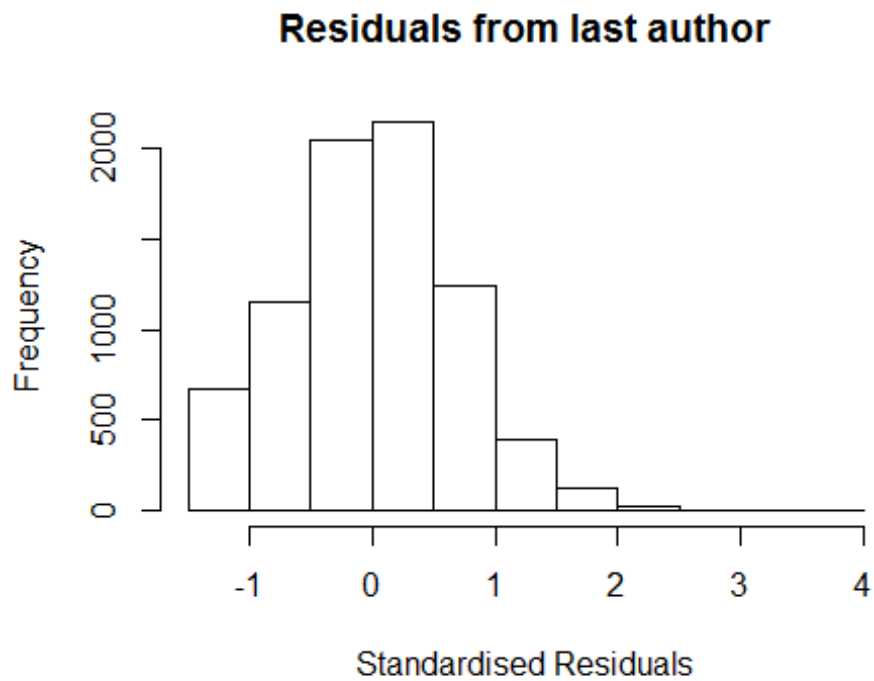
```
## FirstAuthorFemale1  0.03759    0.02523    1.49   0.1364
## LastAuthorFemale1  -0.00946    0.02473   -0.38   0.7022
## Year1997            0.00625    0.04918    0.13   0.8989
## Year1998           -0.03205    0.04584   -0.70   0.4844
## Year1999           -0.01083    0.04908   -0.22   0.8254
## Year2000           -0.03730    0.04450   -0.84   0.4019
## Year2001           -0.02252    0.04741   -0.47   0.6349
## Year2002            0.01627    0.04372    0.37   0.7098
## Year2003            0.03638    0.04714    0.77   0.4404
## Year2004           -0.03316    0.04466   -0.74   0.4578
## Year2005            0.04058    0.04550    0.89   0.3725
## Year2006            0.00641    0.04388    0.15   0.8838
## Year2007           -0.08750    0.04057   -2.16   0.0311 *
## Year2008           -0.09181    0.04519   -2.03   0.0422 *
## Year2009           -0.11470    0.04148   -2.77   0.0057 **
## Year2010           -0.18585    0.04455   -4.17  3.1e-05 ***
## Year2011           -0.09885    0.04550   -2.17   0.0299 *
## Year2012           -0.18441    0.04412   -4.18  3.0e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.669
## Multiple R-squared:  0.0101, Adjusted R-squared:  0.00783
## Convergence in 10 IRWLS iterations
##
## Robustness weights:
##  3 observations c(945,1326,2578) are outliers with |weight| = 0 ( < 1.3e-
05);
##  680 weights are ~= 1. The remaining 7109 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.005   0.874   0.949   0.910   0.985   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.28e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                    psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##                     GVIF Df GVIF^(1/(2*Df))
```

```
## FirstAuthorFemale 1.007  1           1.003
## Year              1.007 16           1.000
```




```
## [1] "List of  7 outliers with residuals above  2.5"
##          ScopusId  NLCS Year OneField Fields residuals
## 1357   0031118203 4.089 1997     1703      3     2.982
## 2068   0031153727 4.130 1997     1708      3     3.023
## 2105   0031211090 4.640 1997     1703      4     3.533
## 2916   0002806618 3.648 1998     1702      5     2.580
## 2938   0032251894 4.293 1998     1712      2     3.225
## 5698   0035455653 4.472 2001     1708      3     3.394
## 14679 78549288866 3.635 2010     2604      3     2.720
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.17372 -0.45604  0.00595  0.45438  3.53431
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.09978    0.03214   34.22  < 2e-16 ***
## FirstAuthorFemale1  0.03378    0.02361    1.43   0.1526
## Year1997            0.00591    0.04914    0.12   0.9043
```

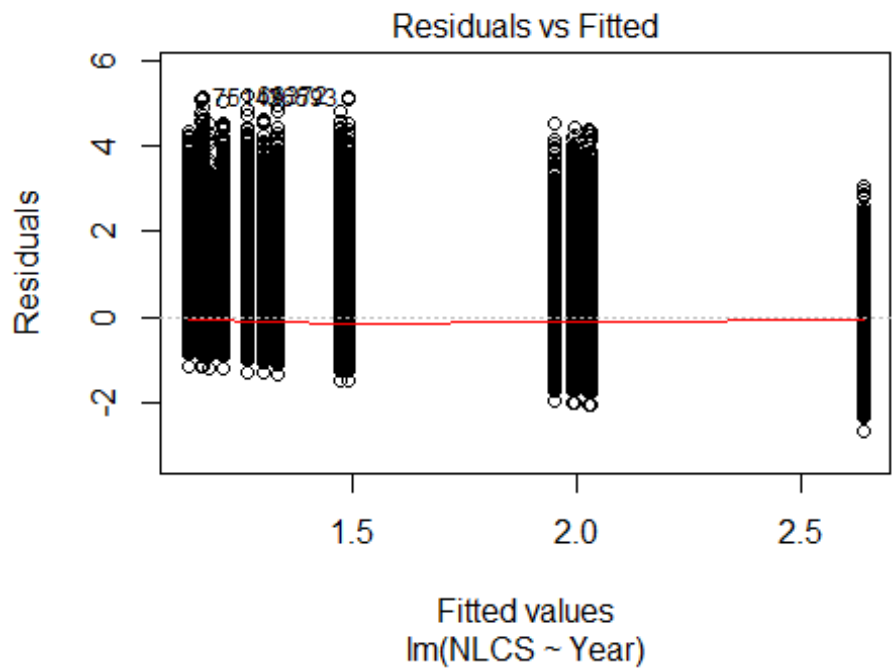
```
## Year1998           -0.03215    0.04582   -0.70   0.4830
## Year1999           -0.01104    0.04908   -0.22   0.8220
## Year2000           -0.03740    0.04450   -0.84   0.4007
## Year2001           -0.02251    0.04741   -0.47   0.6349
## Year2002            0.01604    0.04371    0.37   0.7137
## Year2003            0.03653    0.04714    0.77   0.4384
## Year2004           -0.03321    0.04464   -0.74   0.4570
## Year2005            0.04016    0.04547    0.88   0.3772
## Year2006            0.00653    0.04386    0.15   0.8817
## Year2007           -0.08762    0.04056   -2.16   0.0308 *
## Year2008           -0.09173    0.04518   -2.03   0.0424 *
## Year2009           -0.11482    0.04148   -2.77   0.0057 **
## Year2010           -0.18582    0.04454   -4.17  3.1e-05 ***
## Year2011           -0.09900    0.04549   -2.18   0.0296 *
## Year2012           -0.18444    0.04411   -4.18  2.9e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.67
## Multiple R-squared:  0.0101, Adjusted R-squared:  0.00793
## Convergence in 10 IRWLS iterations
##
## Robustness weights:
##  3 observations c(945,1326,2578) are outliers with |weight| = 0 ( < 1.3e-
05);
##  676 weights are ~= 1. The remaining 7113 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.005   0.874   0.949   0.910   0.985   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.28e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##                   GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.008  1           1.004
## Year             1.008 16           1.000
```



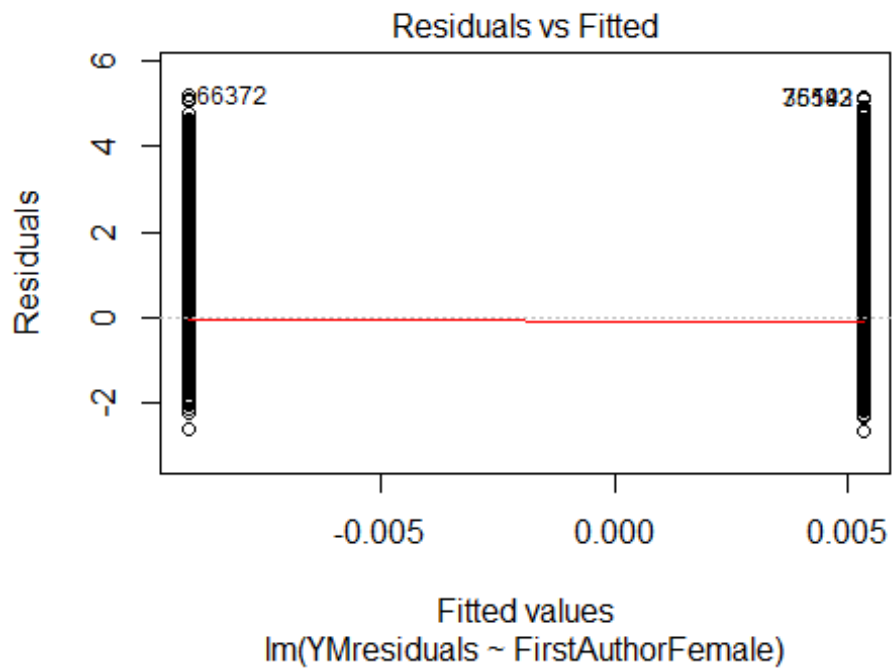

```
## [1] "List of  7 outliers with residuals above  2.5"
##          ScopusId  NLCS Year OneField Fields residuals
## 1357   0031118203 4.089 1997     1703      3     2.982
## 2068   0031153727 4.130 1997     1708      3     3.023
## 2105   0031211090 4.640 1997     1703      4     3.533
## 2916   0002806618 3.648 1998     1702      5     2.580
## 2938   0032251894 4.293 1998     1712      2     3.225
## 5698   0035455653 4.472 2001     1708      3     3.394
## 14679 78549288866 3.635 2010     2604      3     2.720
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.14855 -0.45624  0.00754  0.45130  3.53151
##
## Coefficients:
##                   Estimate Std. Error t value Pr(>|t|)
## (Intercept)        1.10353    0.03204   34.44  < 2e-16 ***
## LastAuthorFemale1  0.00540    0.02304    0.23   0.8146
## Year1997           0.00496    0.04912    0.10   0.9195
## Year1998          -0.03218    0.04582   -0.70   0.4824
## Year1999          -0.01230    0.04901   -0.25   0.8018
## Year2000          -0.03806    0.04449   -0.86   0.3922
```

```
## Year2001          -0.02348    0.04741   -0.50   0.6204
## Year2002           0.01628    0.04373    0.37   0.7097
## Year2003           0.03606    0.04711    0.77   0.4440
## Year2004          -0.03380    0.04463   -0.76   0.4489
## Year2005           0.03962    0.04546    0.87   0.3835
## Year2006           0.00636    0.04391    0.14   0.8848
## Year2007          -0.08796    0.04056   -2.17   0.0302 *
## Year2008          -0.09134    0.04520   -2.02   0.0434 *
## Year2009          -0.11440    0.04149   -2.76   0.0058 **
## Year2010          -0.18666    0.04455   -4.19  2.8e-05 ***
## Year2011          -0.09953    0.04550   -2.19   0.0287 *
## Year2012          -0.18431    0.04409   -4.18  2.9e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.67
## Multiple R-squared:  0.00983,    Adjusted R-squared:  0.00767
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  3 observations c(945,1326,2578) are outliers with |weight| = 0 ( < 1.3e-
05);
##  677 weights are ~= 1. The remaining 7112 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0053  0.8740  0.9490  0.9100  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.28e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Sample size for the above analysis:  7792"
## [1] ""
## [1] ""
## [1] "##################################"
## [1] "Analysis of AJSC 2700"
## [1] "##################################"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
```

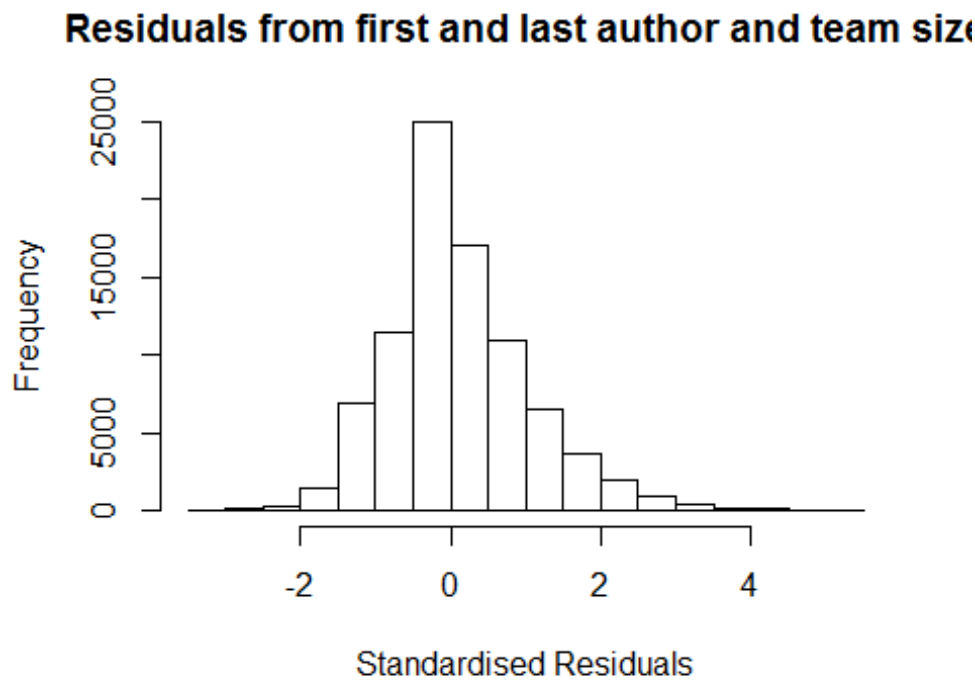
```
##  1996  1997  1998  1999  2000  2001  2002  2003  2004  2005  2006  2007
##  6842  6158  5688  5456   750  6839  5706  5141  5605  6367  8424 10786
##  2008  2009  2010  2011  2012
## 11805 12334 14041 10431 12018
##
##  1996  1997  1998  1999  2000  2001  2002  2003  2004  2005  2006  2007
##  2954  3045  2737  2603   363  2383  4548  4025  4433  5019  6734  8773
##  2008  2009  2010  2011  2012
##  9486  9843 10901  8307  9399
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 2660 2767 2509 2381  334 2159 4148 3674 4047 4520 6102 8028 8597 8845 9849
## 2011 2012
## 7578 8460
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
##  Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 3900, df = 16, p-value <2e-16
```




```
##
##  Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 170, df = 1, p-value <2e-16
```




```
## [1] "Female first author team size 2018 geometric mean: 3.75074271030995"
## [1] "Male first author team size 2018 geometric mean: 3.57061884135357"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 1100000, p-value = 0.05
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 3.47121146633238"
## [1] "Male last author team size 2018 geometric mean: 3.76875385592956"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 920000, p-value = 0.01
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.532  1           1.238
## LastAuthorFemale  1.549  1           1.244
## UniqueAuthors     1.111  4           1.013
## Year              1.088 16           1.003
```




```
## [1] "List of  1596 outliers with residuals above  2.5"
##           ScopusId  NLCS Year OneField Fields residuals
## 28      0029855881 4.801 1996     2700      1     2.587
## 31      0141469033 5.783 1996     2700      1     3.156
## 85      0029902517 4.191 1996     2700      1     3.068
## 87     10544235697 5.572 1996     2700      1     4.397
## 95      0029911547 5.746 1996     2700      1     3.583
## 113     0030445619 5.158 1996     2700      1     2.532
## 144     0029680124 4.121 1996     2700      1     2.946
## 153     0029683640 4.469 1996     2700      1     3.294
## 179     0029844930 4.941 1996     2700      1     2.727
## 217     0030090438 3.718 1996     2700      1     2.543
## 1348    0029956467 0.000 1996     2700      2    -2.627
## 1349    0029992586 0.000 1996     2700      2    -2.577
## 1355    0030590745 4.426 1996     2700      1     3.251
## 1397    0029919459 5.204 1996     2700      1     3.194
## 1424    0029907556 5.146 1996     2700      1     4.023
## 1425    0029908290 4.742 1996     2700      1     3.567
## 1610    0029830202 0.000 1996     1306      3    -2.627
## 1728    0029861579 4.669 1996     2700      1     2.862
## 1764    0007410805 5.440 1996     2700      1     2.813
## 1766   10144244674 5.218 1996     2700      1     2.591
## 1772    0029758899 4.682 1996     2700      1     2.520
## 1794    0029743742 3.747 1996     2700      1     2.572
## 1979    0029842830 5.287 1996     1300      2     2.660
## 2023    0030604909 4.869 1996     2700      1     2.859
## 2071    0029740950 5.202 1996     2700      1     3.191
```

```
## 2072    0029741921 5.309 1996     2700      1     2.682
## 2074    9544228424 5.880 1996     2700      1     3.253
## 2108    0029835392 5.259 1996     2700      1     4.084
## 2110    0242435873 5.375 1996     2700      1     3.617
## 2115    0029771618 4.842 1996     2700      1     2.680
## 2131    0029822091 4.840 1996     2700      1     2.626
## 2162    0029739970 5.663 1996     2700      1     3.036
## 2370    1842614368 0.000 1996     2700      1    -2.627
## 2477    0029767709 0.000 1996     2700      1    -2.627
## 2485    0029741822 5.190 1996     2700      1     2.977
## 2503    0029783235 4.667 1996     2700      1     3.492
## 2513    0029759025 5.617 1996     2700      1     2.991
## 2543   15844420661 5.335 1996     2700      1     2.708
## 2582    0029862164 4.517 1996     2700      1     2.558
## 2729    0029949917 4.390 1996     2700      1     2.582
## 2741    0030594830 5.438 1996     2700      1     2.811
## 2743    0030000230 4.826 1996     2700      1     3.018
## 2744    8944228913 5.408 1996     2700      1     2.782
## 2748    0030056083 5.607 1996     2700      1     2.980
## 2750    0030057106 5.671 1996     2700      1     3.044
## 2797    8944233864 5.465 1996     2700      1     2.838
## 2985    0039568286 0.000 1996     2700      1    -2.577
## 3058    0029900294 5.422 1996     2700      1     2.795
## 3059    0029942003 5.780 1996     2700      1     3.566
## 3063    0029886872 4.256 1996     2700      1     3.081
## 3085    0030594604 4.947 1996     2700      1     2.936
## 3109    0029948212 5.514 1996     2700      1     2.887
## 3144    0029941246 3.794 1996     2700      1     2.671
## 3164    0029972135 3.689 1996     2700      1     2.514
## 3462    0029892519 4.994 1996     2700      1     2.983
## 3473    0029881395 4.142 1996     2700      1     2.967
## 3474    0029897909 5.220 1996     2700      1     2.594
## 3479    0029886630 5.409 1996     2700      1     4.234
## 3498    0029870954 4.842 1996     2700      1     3.667
## 3502    0029873301 3.953 1996     2700      1     2.778
## 3521    0029925567 5.575 1996     2700      1     3.564
## 3537    0029871858 4.847 1996     2700      1     2.634
## 3539    0030008301 4.452 1996     2700      1     2.644
## 3552    9244262406 6.270 1996     2700      1     3.643
## 3560    0029863684 5.262 1996     2700      1     2.635
## 3572    0029913381 4.769 1996     2700      1     2.556
## 3578    0029929617 5.256 1996     2700      1     2.679
## 3579    0029929618 4.964 1996     2700      1     3.003
## 3795    0029875770 5.139 1996     2700      1     2.512
## 3866    0029881352 4.449 1996     2700      1     2.641
## 3877    0029869817 4.507 1996     2700      1     2.546
## 4105    0029970342 5.515 1996     2700      1     2.888
## 4111    0029917496 5.333 1996     2700      1     3.526
## 4137    0029928131 0.000 1996     2700      1    -2.627
## 4178    9044236527 5.250 1996     2700      1     2.673
```

```
## 4181    0029963060 3.740 1996     2700      1     2.565
## 4192    0030050806 4.345 1996     2700      1     3.170
## 4509    0942276272 5.608 1996     2700      1     2.981
## 4535    0030066467 5.150 1996     2700      1     2.523
## 4536    0030070793 4.738 1996     2700      1     2.980
## 4552    0030020630 4.999 1996     2700      1     2.988
## 4567    0030058515 5.136 1996     2700      1     2.509
## 4591    0030042790 5.170 1996     2700      1     2.543
## 4622    0030032444 5.228 1996     2700      1     2.601
## 4868    0030579588 4.838 1996     2700      1     3.030
## 4869    0030034463 5.009 1996     2700      1     3.202
## 4883    0030034592 5.638 1996     2700      1     3.011
## 4884    0030054309 6.200 1996     2700      1     3.573
## 4890    0030024585 4.558 1996     2700      1     2.750
## 4897    0000677401 5.404 1996     2700      1     4.281
## 4904    0001196648 5.048 1996     2700      1     3.873
## 4908    0001854685 4.563 1996     2700      1     3.440
## 4915    0003064116 3.844 1996     2700      1     2.669
## 4916    0004851872 6.351 1996     2700      1     3.774
## 5173    0029793997 4.081 1996     2700      1     2.958
## 5240    0029835297 3.760 1996     2700      1     2.585
## 5250    0029838529 4.175 1996     2700      1     3.000
## 5359    0029890872 0.000 1996     2700      1    -2.577
## 5633    0030029842 4.377 1996     2700      1     2.569
## 5635    0030030302 3.697 1996     2700      1     2.522
## 5653    0030034465 5.226 1996     2700      1     2.599
## 5703    0030053625 3.826 1996     2700      1     2.651
## 5733    0030061122 4.405 1996     2700      1     3.230
## 5737    0030061707 4.711 1996     2700      1     2.903
## 5748    0030067568 4.010 1996     2700      1     2.835
## 5772    0030077445 3.760 1996     2700      1     2.585
## 5847    0030183205 4.634 1996     2700      1     3.511
## 5941    0030306046 4.049 1996     2700      1     2.874
## 6074    0030576183 4.548 1996     2700      1     2.537
## 6075    0030576192 3.851 1996     2700      1     2.676
## 6086    0040419081 4.738 1996     2700      1     2.980
## 6089    0242437957 4.044 1996     2700      1     2.869
## 6155   85023808915 0.000 1996     2700      1    -2.627
## 6523   33749433783 0.000 1996     1300      2    -2.575
## 6549   33749435765 0.000 1996     1300      2    -2.627
## 6606   33749442532 0.000 1996     1300      2    -2.626
## 6636   33749446226 0.000 1996     1300      2    -2.577
## 6674   33749449781 0.000 1996     1300      2    -2.627
## 6777   33749545565 0.000 1996     1300      2    -2.627
## 6788   33749547000 0.000 1996     1300      2    -2.577
## 6816   33749552020 0.000 1996     1300      2    -2.626
## 6867   33749555874 0.000 1996     1300      2    -2.627
## 6925   33749561284 0.000 1996     1300      2    -2.575
## 7329    0031435838 4.794 1997     2700      1     2.851
## 7330    2642611953 4.758 1997     2700      1     3.652
```

```
## 7367    0031437624 4.545 1997     2700      1     2.552
## 7379    2642597076 5.193 1997     2700      1     2.997
## 7481    0030954396 3.869 1997     2700      1     2.763
## 7489    0031007965 4.037 1997     2700      1     2.879
## 7534    0031136510 4.780 1997     2700      1     2.787
## 7571    0031263974 4.374 1997     2700      1     3.216
## 7673    0031304421 0.000 1997     2700      1    -2.608
## 7802    0031554081 3.850 1997     2700      1     2.692
## 7822    0346580106 0.000 1997     2700      1    -2.610
## 8307    0030665744 4.564 1997     2700      1     2.571
## 8332    0030692782 5.405 1997     2700      1     2.795
## 8339    0030666228 4.564 1997     2700      1     3.458
## 8340    0030671045 4.349 1997     2700      1     2.609
## 8378    0030695138 5.236 1997     2700      1     2.626
## 8408    0030732965 4.209 1997     2700      1     3.103
## 8424    0030716498 5.322 1997     2700      1     4.164
## 8442    0030661687 5.440 1997     2700      1     2.830
## 8586    0030731486 5.178 1997     2700      1     3.439
## 8607    0030667982 4.315 1997     2700      1     2.524
## 8608    0030720886 5.370 1997     2700      1     2.811
## 8610    0030759072 5.217 1997     2700      1     2.607
## 8625    0030744945 5.294 1997     2700      1     2.684
## 8631    0030879325 4.935 1997     2700      1     3.777
## 8641    0030726989 4.299 1997     2700      1     2.508
## 8713    0030779037 5.224 1997     2700      1     2.616
## 8850    0030954873 5.207 1997     2700      1     3.264
## 8851    0030967953 6.065 1997     2700      1     3.455
## 8980    0030869269 5.624 1997     2700      1     3.014
## 9007    0030763532 5.703 1997     2700      1     3.508
## 9014    0030928107 4.550 1997     2700      1     2.810
## 9239    0030868293 5.188 1997     2700      1     2.578
## 9317    0030837662 4.285 1997     2700      1     3.127
## 9330    0030853307 4.301 1997     2700      1     3.143
## 9498    0030739904 4.155 1997     2700      1     2.997
## 9515    0030756101 4.907 1997     2700      1     3.168
## 9539    0030876805 5.124 1997     2700      1     2.514
## 9575    0030754066 4.117 1997     2700      1     2.959
## 9597    0030610461 5.331 1997     2700      1     2.723
## 9607    0030957310 5.247 1997     2700      1     2.637
## 9760    0031005933 5.176 1997     2700      1     2.566
## 9763    1842410169 4.665 1997     2700      1     2.673
## 9767    0030916401 5.117 1997     2700      1     2.507
## 9816    0030908055 5.171 1997     2700      1     4.013
## 9820    0030976173 4.495 1997     2700      1     2.755
## 9823    0031009275 4.344 1997     2700      1     2.605
## 9833    0343157354 5.144 1997     2700      1     2.534
## 10056   0030910022 5.719 1997     2700      1     3.160
## 10066   0031003334 5.309 1997     2700      1     2.699
## 10102   0030919511 5.119 1997     2700      1     3.961
## 10149   0030913316 4.821 1997     2700      1     2.828
```

```
## 10156   0030961921 5.444 1997     2700      1     2.885
## 10203   0030982247 3.862 1997     1000      2     2.704
## 10222   0030979720 4.965 1997     2700      1     3.174
## 10347   0030967165 5.834 1997     2700      1     3.226
## 10353   0030955080 4.722 1997     2700      1     2.576
## 10378   0012444519 6.220 1997     2700      1     3.610
## 10427   0030896996 4.738 1997     2700      1     2.795
## 10454   0030956673 6.413 1997     2700      1     3.803
## 10504   0031110088 0.000 1997     2700      1    -2.610
## 10535   0031127072 5.869 1997     2700      1     3.311
## 10598   0030976067 5.207 1997     2700      1     2.599
## 10671   0031054674 5.776 1997     2700      1     3.166
## 10694   0031047361 4.710 1997     2700      1     3.552
## 10700   8044240301 4.056 1997     2700      1     2.898
## 10767   0031086622 0.000 1997     2700      1    -2.610
## 10895   0031052369 4.993 1997     2700      1     3.835
## 10919   0031045652 5.274 1997     2700      1     2.664
## 10933   0031052862 5.331 1997     2700      1     2.723
## 10941   0030889653 4.190 1997     2700      1     3.084
## 10987   0031028804 5.501 1997     2700      1     2.891
## 11136   0031012726 5.498 1997     2700      1     2.888
## 11137   0031030450 4.891 1997     2700      1     2.898
## 11138   0031032055 5.515 1997     2700      1     2.905
## 11153   0031030620 5.167 1997     2700      1     2.557
## 11162   0031012532 4.615 1997     2700      1     3.457
## 11486   0030679798 3.716 1997     2700      1     2.610
## 11773   0030914380 3.929 1997     2700      1     2.771
## 11955   0031011193 4.851 1997     2700      1     2.655
## 13373   0031472453 4.648 1997     2700      1     2.857
## 13417   0031566681 4.037 1997     2700      1     2.931
## 13487   1842295778 5.282 1997     2700      1     2.672
## 13961   0032583492 5.604 1998     2700      1     3.009
## 14018   0032506614 5.304 1998     2700      1     3.478
## 14040   0032477294 5.316 1998     2700      1     2.721
## 14041   0032477305 5.311 1998     2700      1     2.716
## 14052   0002724443 4.318 1998     2700      1     3.125
## 14056   0007436490 3.696 1998     2700      1     2.503
## 14500   0032420039 0.000 1998     1300      2    -2.644
## 14800   0032567110 3.908 1998     2700      1     2.715
## 14801   0032567115 5.271 1998     2700      1     3.039
## 14816   0032547938 6.002 1998     2700      1     3.407
## 14866   0032508952 6.390 1998     2700      1     3.745
## 14888   0032483685 5.248 1998     2700      1     2.603
## 15052   0032576177 3.796 1998     2700      1     2.603
## 15070   0031719941 0.000 1998     2700      4    -2.644
## 15080   0032556180 5.427 1998     2700      1     2.832
## 15186   0032189897 5.522 1998     2700      1     2.877
## 15357   0032541672 4.883 1998     2700      1     2.854
## 15367   0032529862 0.000 1998     2700      1    -2.644
## 15530   0032169485 5.282 1998     2700      1     2.637
```

```
## 15666   0032547326 6.369 1998     2700      1     3.724
## 15700   0008926519 5.466 1998     2700      1     4.273
## 15791   0032146983 3.942 1998     2700      1     2.749
## 15910   0032527149 0.000 1998     2700      1    -2.595
## 15914   0032527531 3.889 1998     2700      1     2.748
## 15935   0032528180 5.239 1998     2700      1     3.007
## 15947   0032508297 5.453 1998     2700      1     4.260
## 16180   0031807705 0.000 1998     2700      2    -2.645
## 16251   0032482329 5.835 1998     2700      1     3.240
## 16283   0031847902 0.000 1998     2700      1    -2.645
## 16483   0032572043 5.088 1998     2700      1     3.312
## 16487   0032572086 6.274 1998     2700      1     3.629
## 16504   0032550626 5.681 1998     2700      1     4.488
## 16535   0032516250 3.822 1998     2700      1     2.629
## 16539   0032516296 5.565 1998     2700      1     2.920
## 16788  18144441631 3.731 1998     2700      1     2.538
## 16800   0032557174 4.358 1998     2700      1     3.217
## 16802   2642713359 4.651 1998     2700      1     3.458
## 16816   0032537068 4.638 1998     2700      1     2.609
## 17163   0032485350 5.150 1998     2700      1     2.505
## 17289   0032032575 4.574 1998     2700      1     2.545
## 17386   0032564902 5.256 1998     2700      1     2.612
## 17429   0032519925 5.694 1998     2700      1     3.050
## 17444   0032510076 5.618 1998     2700      1     2.974
## 17498   0031915909 0.000 1998     2700      1    -2.645
## 17595   0032573850 4.468 1998     2700      1     3.275
## 17627   0032518022 6.088 1998     2700      1     4.895
## 17696   0009440989 4.352 1998     2700      1     2.576
## 17954   0031694575 0.000 1998     2700      1    -2.645
## 18335   0031940044 3.864 1998     2700      1     2.671
## 18344   0031945285 5.058 1998     2700      1     3.865
## 18436   0031975335 5.254 1998     2700      1     2.661
## 18437   0031975401 4.758 1998     2700      1     2.526
## 18454   0031979959 0.000 1998     2700      1    -2.645
## 20189   0033619959 5.403 1999     2700      1     2.837
## 20242   0033572972 5.596 1999     2700      1     3.797
## 20277   0033518192 4.365 1999     2700      1     2.568
## 20279   0033518231 5.873 1999     2700      1     3.255
## 20319   0032803699 4.811 1999     2700      1     3.012
## 20463   0033391677 0.000 1999     2700      1    -2.618
## 20688  33750645471 0.000 1999     2700      2    -2.618
## 20693  33750829720 0.000 1999     2700      2    -2.618
## 20874   0032748385 5.646 1999     2700      1     3.030
## 20925   0033544340 6.198 1999     2700      1     4.247
## 20967   0032589826 4.410 1999     2700      1     3.244
## 21115   0032694263 5.196 1999     2700      1     2.578
## 21120   0032695482 5.349 1999     2700      1     2.731
## 21124   0032742418 5.935 1999     2700      1     3.319
## 21126   0032743289 5.343 1999     2700      1     4.177
## 21134   0033598598 5.651 1999     2700      1     3.701
```

```
## 21152   0032696770 5.142 1999     2700      1     2.524
## 21407   0033619146 3.942 1999     2700      1     2.776
## 21427   0033598326 4.355 1999     2700      1     3.189
## 21442   0033575990 3.844 1999     2700      1     2.678
## 21500   0033536483 4.408 1999     2700      1     2.660
## 21501   0033536501 4.406 1999     2700      1     2.658
## 21656   0033199833 4.945 1999     2700      1     3.779
## 21775   0342961310 4.661 1999     2700      1     2.507
## 21785   0033584440 5.727 1999     2700      1     3.111
## 21792   0033581218 4.318 1999     2700      1     2.519
## 21826   0033546647 5.256 1999     2700      1     2.640
## 22042   0033542870 5.579 1999     2700      1     2.963
## 22186   0033162159 0.000 1999     2700      2    -2.566
## 22286   0033542393 4.685 1999     2700      1     2.734
## 22491   0033606238 4.793 1999     2700      1     2.791
## 22563   0033526309 5.723 1999     2700      1     3.105
## 22620   0033124317 0.000 1999     2700      1    -2.618
## 22779   0033594375 4.872 1999     2700      1     2.668
## 22832   0033553188 4.783 1999     2700      1     2.832
## 22855   0033515827 5.255 1999     1000      2     2.637
## 22909   0033119011 4.749 1999     2700      1     3.635
## 23056   0033545541 5.796 1999     2700      1     3.997
## 23057   0033545542 4.631 1999     2700      1     3.465
## 23076   0033522146 5.479 1999     2700      1     2.863
## 23077   0033522206 5.337 1999     2700      1     2.719
## 23271   0033602049 5.262 1999     2700      1     2.644
## 23358   0033540646 5.996 1999     2700      1     3.994
## 23360   0033540714 4.973 1999     2700      1     3.859
## 23404   0032913899 0.000 1999     2700      1    -2.567
## 23441   0033072447 3.953 1999     2700      1     2.787
## 23455   0033082906 3.824 1999     2700      1     2.658
## 23605   0033537343 4.731 1999     2700      1     2.527
## 23882   0032713075 5.527 1999     2700      1     2.909
## 23974   0032742992 5.265 1999     2700      1     3.061
## 24083   0032824349 4.787 1999     2700      1     3.621
## 24424   0033013684 4.775 1999     2700      1     2.572
## 24669   0033090874 3.721 1999     2700      1     2.607
## 24827   0033125302 4.256 1999     2700      1     3.090
## 25647   0033428375 0.000 1999     2700      1    -2.566
## 25693   0033494450 4.153 1999     2700      1     2.987
## 25901  85023880658 0.000 1999     2700      1    -2.567
## 26303   0034570594 0.494 2000     2700      1    -2.950
## 26307   0034578262 0.494 2000     2700      1    -2.952
## 26313  20244377663 0.000 2000     2700      1    -2.779
## 26314  24044533275 0.000 2000     1306      3    -2.626
## 26319   0005981935 0.418 2000     1312      3    -3.026
## 26344   1542493233 0.759 2000     1109      3    -2.636
## 26349  22144462988 0.783 2000     1702      5    -2.663
## 26353   0034735826 4.991 2000     2700      1     2.997
## 26365   0033754163 0.436 2000     1306      3    -3.010
```

```
## 26368   0034626988 5.215 2000     2700      1     3.221
## 26410   0034305633 0.000 2000     2700      2    -3.033
## 26414   0034727059 5.099 2000     2700      1     3.105
## 26509  33746530366 0.000 2000     2700      1    -3.446
## 26578   0033756474 4.527 2000     2700      1     2.533
## 26606   0033803048 4.759 2000     2700      1     2.765
## 26651   0033932709 4.578 2000     2700      1     2.584
## 26692   0034048842 0.000 2000     2700      1    -2.830
## 26703   0034090148 4.539 2000     2700      1     2.545
## 26792   0034302210 0.000 2000     2700      1    -2.576
## 26910  12944252960 0.000 2000     2700      1    -2.830
## 26915   0033950423 4.593 2000     2700      1     2.599
## 26949  85047694984 0.000 2000     2700      1    -2.576
## 26964   0034732201 5.487 2000     2700      1     3.493
## 26978   0034607347 4.705 2000     2700      1     2.711
## 27018   0034102286 0.000 2000     2700      1    -2.627
## 27025   0010838455 5.034 2000     2700      1     3.040
## 27050   0034210885 0.759 2000     1303      8    -2.685
## 27110   0035956483 5.056 2001     2700      1     2.531
## 27142   0035924765 5.181 2001     2700      1     2.656
## 27211   0035852037 5.008 2001     2700      1     3.300
## 27503   0035708764 0.000 2001     2700      1    -2.525
## 27591   0442328578 0.000 2001     2700      1    -2.527
## 28027   0035969487 3.523 2001     2700      1     2.500
## 28120   0035892018 5.798 2001     2700      1     3.271
## 28179   0035829842 6.472 2001     2700      1     3.945
## 28353   0035522330 5.300 2001     2700      1     2.773
## 28832   0035913589 5.497 2001     2700      1     2.970
## 28849   0035856010 4.264 2001     2700      1     3.189
## 28851   0035856017 5.495 2001     2700      1     2.968
## 28898   0035818039 3.656 2001     2700      1     2.633
## 28900   0035818048 5.736 2001     2700      1     3.209
## 29160   0035940040 4.567 2001     2700      1     2.859
## 29299   0035413919 4.889 2001     2700      1     2.826
## 29417   0035954660 4.486 2001     2700      1     3.411
## 29418   0035954670 4.624 2001     2700      1     2.916
## 29428   0035948630 4.447 2001     2700      1     2.739
## 29524   0035850010 5.106 2001     2700      1     2.581
## 29532   0034945516 5.330 2001     2700      1     3.217
## 29705   0035963494 4.396 2001     2700      1     2.688
## 29708   0035963529 5.289 2001     2700      1     2.762
## 29767   0035897888 4.789 2001     2700      1     3.714
## 29787   0035854054 5.981 2001     2700      1     3.505
## 29811   0035821989 3.751 2001     2700      1     2.676
## 29826   0035816007 5.037 2001     2700      1     2.510
## 29832   0035816032 5.067 2001     2700      1     2.540
## 30017   0035938676 3.839 2001     2700      1     2.764
## 30038   0035902194 4.674 2001     2700      1     3.599
## 30047   0035897688 5.037 2001     2700      1     3.126
## 30084   0035832261 6.117 2001     2700      1     3.590
```

```
## 30111   0035799768 4.118 2001     2700      1     3.043
## 30189   0035347549 3.710 2001     2700      1     2.635
## 30872   0035946664 5.610 2001     2700      1     3.085
## 31007   0035804852 4.734 2001     2700      1     2.621
## 31257   0035826089 5.495 2001     2700      1     2.968
## 31264   0035819910 4.759 2001     2700      1     2.850
## 31434   0035961566 5.156 2001     2700      1     2.629
## 31485   0035901051 5.395 2001     2700      1     3.484
## 31492   0035864977 4.470 2001     2700      1     2.559
## 31541   0035825726 5.310 2001     2700      1     2.783
## 31555   0035819462 4.142 2001     2700      1     3.067
## 31569   0035819481 4.496 2001     2700      1     2.587
## 31734   0035915688 4.647 2001     2700      1     2.535
## 31773   0035843243 5.197 2001     2700      1     4.174
## 31782   0035835027 4.362 2001     2700      1     2.503
## 31799   0035804277 5.895 2001     2700      1     3.368
## 31939   0034800418 3.751 2001     2700      1     2.676
## 32222   0034992555 3.603 2001     2700      1     2.528
## 32248   0035013499 4.468 2001     2700      1     2.559
## 33892   0035960866 5.082 2001     2700      1     2.970
## 34255  84984780131 0.000 2001     1311      2    -2.527
## 34423   0037132575 4.351 2002     2700      1     2.633
## 34438   0037079352 4.660 2002     2700      1     2.530
## 34454   0037065306 5.195 2002     2700      1     3.064
## 34467   0037028001 4.145 2002     2700      1     2.630
## 34491   0036369590 3.133 2002     2700      1     2.506
## 34977   0037184416 4.294 2002     2700      1     2.576
## 34979   0037184418 4.896 2002     2700      1     3.584
## 35005   0037153042 5.145 2002     2700      1     3.066
## 35013   0037145856 5.083 2002     2700      1     2.952
## 35014   0037145863 5.578 2002     2700      1     3.447
## 35032   0037078969 5.588 2002     2700      1     3.508
## 35039   0037073256 4.972 2002     2700      1     2.893
## 35079   0037032401 4.648 2002     2700      1     2.517
## 35155   0036850836 3.575 2002     2700      1     2.896
## 35358   0037206364 5.141 2002     2700      1     3.010
## 35359   0037206368 4.285 2002     2700      1     2.770
## 35377   0037167982 4.361 2002     2700      1     3.682
## 35386   0037164054 5.560 2002     2700      1     3.430
## 35393   0037136930 3.570 2002     2700      1     2.891
## 35397   0037126201 4.294 2002     2700      1     2.982
## 35398   0037126332 4.518 2002     2700      1     2.802
## 35400   0037126341 5.773 2002     2700      1     3.642
## 35408   0037120864 4.463 2002     2700      1     2.948
## 35411   0037120907 4.813 2002     2700      1     2.733
## 35448   0037048669 6.012 2002     2700      1     4.296
## 35451   0037048695 5.664 2002     2700      1     3.948
## 35461   0037015541 3.490 2002     2700      1     2.863
## 35672   0036810576 3.330 2002     1704      3     2.651
## 35699   0037179578 4.011 2002     2700      1     3.332
```

```
## 35704   0037179705 3.405 2002     2700      1     2.726
## 35709   0037174346 4.951 2002     2700      1     2.820
## 35759   0037063403 4.739 2002     2700      1     2.660
## 35777   0037026697 3.399 2002     2700      1     2.720
## 35778  18644385210 4.777 2002     2700      1     2.647
## 36123   0037190086 4.360 2002     2700      1     2.845
## 36181   0037158614 4.762 2002     2700      1     2.631
## 36254   0037103420 4.673 2002     2700      1     2.955
## 36269   0037077489 4.481 2002     2700      1     2.968
## 36270   0037077490 4.958 2002     2700      1     2.878
## 36292   0037043652 4.862 2002     2700      1     2.732
## 36306   0037036787 4.703 2002     2700      1     2.572
## 36365   0036681988 5.232 2002     2700      1     3.101
## 36367   0036682163 4.454 2002     2700      1     3.142
## 36369   0036682267 4.857 2002     2700      1     2.726
## 36554   0037166985 4.791 2002     2700      1     2.660
## 36555   0037166986 4.893 2002     2700      1     2.762
## 36557   0037167018 3.189 2002     2700      1     2.510
## 36583   0037130113 4.121 2002     2700      1     2.658
## 36593   0037125379 6.631 2002     2700      1     4.551
## 36658   0037019322 4.383 2002     2700      1     3.071
## 36666   0037014584 5.135 2002     2700      1     3.056
## 36764  34247610837 3.715 2002     2700      1     3.036
## 36860   0037134824 4.216 2002     2700      1     2.703
## 36864   0037134872 4.163 2002     2700      1     2.852
## 36904   0037071808 5.799 2002     2700      1     3.668
## 36956   0037024313 4.125 2002     2700      1     2.661
## 37144   0037198420 4.789 2002     2700      1     2.658
## 37146   0037198432 5.119 2002     2700      1     3.039
## 37181   0037162115 4.877 2002     2700      1     2.746
## 37182   0037162116 5.260 2002     2700      1     3.129
## 37231   0037118288 3.241 2002     2700      1     2.562
## 37246   0037093012 4.590 2002     2700      1     2.510
## 37280   0037046655 3.529 2002     2700      1     2.850
## 37286   0037042231 4.892 2002     2700      1     4.213
## 37311   0037007683 4.588 2002     2700      1     3.075
## 37362   0036570099 4.672 2002     2700      1     2.542
## 37363   0036570101 3.740 2002     2700      1     3.061
## 37461   0037171695 3.465 2002     2700      1     2.786
## 37476   0037165703 4.734 2002     2700      1     2.603
## 37478   0037165728 4.888 2002     2700      1     2.757
## 37479   0037165740 5.429 2002     2700      1     3.350
## 37570   0037061912 5.703 2002     2700      1     3.572
## 37582   0037051975 4.690 2002     2700      1     2.610
## 37626   0037018496 4.615 2002     2700      1     2.949
## 37638   0037012465 4.786 2002     2700      1     2.655
## 37752   0037187907 5.013 2002     2700      1     2.882
## 37754   0037187951 3.295 2002     2700      1     2.616
## 37759   0037181511 4.841 2002     2700      1     4.162
## 37760   0037181515 4.728 2002     2700      1     2.597
```

```
## 37868   0037076025 5.024 2002     2700      1     2.945
## 37875   0037070525 4.224 2002     2700      1     2.709
## 37900   0037035121 4.647 2002     2700      1     2.568
## 37919   0037022835 3.622 2002     2700      1     2.995
## 38074   0037186922 3.807 2002     2700      1     3.128
## 38083   0037181148 4.810 2002     2700      1     2.679
## 38109   0037148928 4.712 2002     2700      1     2.582
## 38143   0037116832 5.558 2002     2700      1     3.427
## 38145   0037116836 4.256 2002     2700      1     2.540
## 38161   0037075257 5.488 2002     2700      1     3.358
## 38163   0037075268 3.927 2002     2700      1     3.300
## 38172   0037070259 4.711 2002     2700      1     2.580
## 38191   0036155193 3.358 2002     2700      1     2.679
## 38208   0037034257 6.602 2002     2700      1     4.471
## 38220   0037028769 4.250 2002     2700      1     2.532
## 38328   0037204196 4.718 2002     2700      1     4.039
## 38375   0037122950 4.667 2002     2700      1     2.588
## 38377   0037122987 3.981 2002     2700      1     3.302
## 38381   0037116641 6.031 2002     2700      1     4.516
## 38382   0037116642 5.093 2002     2700      1     2.962
## 38398   0037050352 5.876 2002     2700      1     3.746
## 38421   0037012134 4.436 2002     2700      1     2.921
## 38432   0037006117 3.976 2002     2700      1     2.664
## 38553   0036106054 4.887 2002     2700      1     2.807
## 38606   0036143018 4.750 2002     2700      1     3.237
## 38755   0036210052 4.381 2002     2700      1     3.121
## 38789   0036239507 4.691 2002     2700      1     2.561
## 40605   0346030568 5.028 2003     2700      1     2.986
## 40608   0347882750 5.233 2003     2700      1     3.139
## 40610   0348011697 4.079 2003     2700      1     3.436
## 40646   0344235454 4.984 2003     2700      1     2.940
## 40659   0345168925 4.208 2003     2700      1     2.527
## 40660   0345492466 4.735 2003     2700      1     2.641
## 40682   0345690179 5.350 2003     2700      1     3.256
## 40771   0348230958 6.275 2003     2700      1     4.181
## 41122   0345714860 5.499 2003     2700      1     3.406
## 41128   0344861847 4.621 2003     2700      1     2.527
## 41148   0344926414 5.006 2003     2700      1     2.962
## 41151   0642278662 3.806 2003     2700      1     3.215
## 41156   0344943245 4.973 2003     2700      1     2.879
## 41157   0345374595 4.903 2003     2700      1     2.809
## 41159   0642272544 4.675 2003     2700      1     2.582
## 41161   0242551979 3.904 2003     2700      1     2.628
## 41183   0242410368 4.973 2003     2700      1     2.879
## 41237   0344444858 4.520 2003     2700      1     2.840
## 41258   0346969978 5.002 2003     2700      1     2.909
## 41460   0142089171 4.888 2003     2700      1     2.795
## 41463   0142123411 4.222 2003     2700      1     3.579
## 41468   0142087597 4.833 2003     2700      1     4.190
## 41484   0142009533 5.306 2003     2700      1     3.213
```

```
## 41532   0141863200 3.490 2003     2700      1     2.847
## 41575   0141653838 4.222 2003     2700      1     2.541
## 41587   0141705375 4.884 2003     2700      1     2.791
## 41593   0141816759 4.768 2003     2700      1     2.675
## 41772   0141484564 4.167 2003     2700      1     3.524
## 41778   0141573545 5.008 2003     2700      1     2.915
## 41783   0141796739 4.693 2003     2700      1     2.599
## 41810   0141723557 3.783 2003     2700      1     2.507
## 41829   0042387879 4.992 2003     2700      1     2.899
## 41850   0042360213 5.273 2003     2700      1     3.179
## 41858   0042884162 3.758 2003     2700      1     2.533
## 42124   0041327804 5.015 2003     2700      1     2.921
## 42140   0042422040 4.594 2003     2700      1     2.550
## 42171   0042195833 4.746 2003     2700      1     2.652
## 42175   0042697063 4.958 2003     2700      1     2.864
## 42203   0042661252 4.135 2003     2700      1     2.861
## 42216   0042093742 5.388 2003     2700      1     3.346
## 42228   0042125511 4.751 2003     2700      1     2.658
## 42267   0042888943 4.611 2003     2700      1     2.517
## 42398   0042343801 5.646 2003     2700      1     3.552
## 42438   0038825532 3.870 2003     2700      1     3.227
## 42444   0038601952 4.724 2003     2700      1     3.296
## 42449   0038304776 5.520 2003     2700      1     4.877
## 42466   0038690424 5.414 2003     2700      1     4.138
## 42467   0038690437 4.549 2003     2700      1     2.507
## 42690   0037973279 5.594 2003     2700      1     3.500
## 42691   0038482206 5.915 2003     2700      1     3.873
## 42713   0038130715 5.029 2003     2700      1     2.935
## 42719   0037493499 4.842 2003     2700      1     2.748
## 42736   0037840165 4.892 2003     2700      1     2.798
## 42746   0038242968 5.038 2003     2700      1     2.944
## 42756   0038142845 4.115 2003     2700      1     2.891
## 42772   0037986208 4.721 2003     2700      1     2.627
## 43074   0037638884 5.206 2003     2700      1     3.112
## 43077   0038724280 4.848 2003     2700      1     2.804
## 43085   0037737900 5.165 2003     2700      1     3.071
## 43088   0037947320 3.982 2003     2700      1     3.339
## 43089   0038075468 5.272 2003     2700      1     3.178
## 43091   0038751994 4.929 2003     2700      1     2.885
## 43129   0038316599 4.797 2003     2700      1     2.704
## 43132   0037906573 4.466 2003     2700      1     2.989
## 43138   0012868624 4.671 2003     2700      1     2.629
## 43146   0038824056 4.162 2003     2700      1     2.531
## 43187   0037865316 3.141 2003     2700      1     2.550
## 43287   0242684416 5.438 2003     2700      1     3.396
## 43416   0037464536 4.084 2003     2700      1     2.860
## 43428   0038528237 4.980 2003     2700      1     4.337
## 43435   0037451905 5.245 2003     2700      1     3.151
## 43438   0037451929 6.034 2003     2700      1     4.353
## 43479   0037502809 5.082 2003     2700      1     3.038
```

```
## 43481   0037840394 4.140 2003     2700      1     2.509
## 43483   0037986313 5.056 2003     2700      1     3.578
## 43485   0038516861 4.525 2003     2700      1     3.099
## 43498   0037417219 4.629 2003     2700      1     2.587
## 43514   0037235543 3.306 2003     2700      1     2.715
## 43561   0037399439 5.125 2003     2700      1     3.031
## 43726   0037471795 3.465 2003     2700      1     2.822
## 43733   0037468682 3.829 2003     2700      1     2.604
## 43767   0037454282 4.138 2003     2700      1     2.661
## 43808   0037420274 5.142 2003     2700      1     3.048
## 44104   0037466917 4.108 2003     2700      1     2.631
## 44105   0037466924 3.525 2003     2700      1     2.934
## 44121   0037456358 4.402 2003     2700      1     2.924
## 44150   0037434515 4.739 2003     2700      1     4.096
## 44151   0037434553 3.249 2003     2700      1     2.658
## 44182   0037289753 4.297 2003     2700      1     2.819
## 44192   0037308404 5.103 2003     2700      1     3.009
## 44201   0037311377 4.989 2003     2700      1     2.896
## 44223   0037316824 3.170 2003     2700      1     2.527
## 44261   0037330518 3.985 2003     2700      1     2.507
## 44384   0037472875 3.112 2003     2700      1     2.521
## 44428   0037438792 4.376 2003     2700      1     2.898
## 44454   0037425535 5.533 2003     2700      1     3.440
## 44456   0037425564 5.704 2003     2700      1     3.610
## 44457   0037425578 5.336 2003     2700      1     3.242
## 44461   0037417523 5.839 2003     2700      1     4.413
## 44469   0037413492 5.094 2003     2700      1     4.503
## 44472   0037413628 3.962 2003     1000      2     2.688
## 44490   0037216353 5.144 2003     2700      1     3.050
## 44501   0037218814 4.737 2003     2700      1     2.643
## 44529   0037241112 4.447 2003     2700      1     2.766
## 44659   0037284255 3.631 2003     2700      1     2.988
## 44662   0037286981 3.091 2003     2700      1     2.500
## 44720   0037382807 4.266 2003     2700      1     2.585
## 44745   0037498595 4.430 2003     2700      1     2.801
## 44788   0037662718 3.343 2003     2700      1     2.700
## 44793   0037685168 4.190 2003     2700      1     2.712
## 44851   0037979236 3.290 2003     2700      1     2.699
## 45110   0042379780 3.884 2003     2700      1     2.610
## 45365   0345374590 4.847 2003     2700      1     4.204
## 45366   0345374591 4.435 2003     2700      1     3.210
## 46010   6344263674 3.170 2003     2700      1     2.579
## 46148   0042737443 4.087 2003     2700      1     2.813
## 46153   0141499228 4.645 2003     2700      1     4.054
## 46157   0242363254 3.616 2003     2700      1     2.973
## 46333  11144239923 5.214 2004     2700      1     3.232
## 46334  19944395825 4.649 2004     2700      1     2.667
## 46367  10344231441 4.676 2004     2700      1     2.695
## 46368  10344239881 4.585 2004     2700      1     2.603
## 46376  10344229444 4.159 2004     2700      1     2.642
```

```
## 46574  14544301908 3.200 2004     2700      1     2.670
## 47019   9244247612 4.578 2004     2700      1     2.646
## 47044   8544252449 4.680 2004     2700      1     2.750
## 47045   8544258102 4.526 2004     2700      1     2.596
## 47062  19644400578 4.870 2004     2700      1     2.889
## 47069   7744231805 4.691 2004     2700      1     2.709
## 47072   7744239901 4.468 2004     2700      1     2.538
## 47074   7244254356 3.066 2004     2700      2     2.536
## 47147  13944251605 4.572 2004     2700      1     2.590
## 47393   6944244875 4.867 2004     2700      1     2.885
## 47399   6944229468 4.497 2004     2700      1     3.967
## 47458  19544385747 5.348 2004     2700      1     3.366
## 47490  13244262642 3.694 2004     2700      1     3.164
## 47601   7444240833 3.982 2004     2700      1     3.504
## 47688   4544260101 3.421 2004     1704      3     2.891
## 47760   4644367114 3.723 2004     2700      1     3.193
## 47790   4544279029 4.188 2004     2700      1     2.619
## 47806   4544332903 5.079 2004     2700      1     3.097
## 47825   4444358501 4.553 2004     2700      1     2.571
## 47967   4644226115 3.064 2004     2700      1     2.586
## 47995   4944239035 4.273 2004     2700      1     2.907
## 48164   4344661328 3.400 2004     1702      5     2.870
## 48207   4143067005 4.774 2004     2700      1     2.792
## 48228   4143094988 5.453 2004     2700      1     3.471
## 48236   4043076922 3.999 2004     2700      1     2.633
## 48238   4043153049 4.489 2004     2700      1     2.507
## 48239   4043156247 4.890 2004     2700      1     4.360
## 48367   4043055316 4.632 2004     2700      1     2.650
## 48478   3843102623 3.655 2004     1704      3     3.125
## 48512   3342892905 4.641 2004     2700      1     2.659
## 48553   3142735110 4.217 2004     2700      1     2.901
## 48554   3142745348 4.790 2004     2700      1     2.809
## 48665   3042642128 5.510 2004     2700      1     3.528
## 48674   3042781155 3.884 2004     2700      1     3.354
## 48680   3142748335 3.450 2004     2700      1     2.920
## 48952   2942554887 5.618 2004     2700      1     3.688
## 49006   2542548063 5.360 2004     2700      1     3.428
## 49074   2942720933 4.617 2004     2700      1     2.636
## 49281   2442665224 4.280 2004     2700      1     2.914
## 49283   2442715038 5.135 2004     2700      1     3.153
## 49290   2442696436 4.547 2004     2700      1     2.565
## 49306   2342471392 6.269 2004     2700      1     4.287
## 49307   2442479695 5.758 2004     2700      1     3.776
## 49309   2442572117 5.004 2004     2700      1     3.022
## 49313   2442590642 4.673 2004     2700      1     2.691
## 49369   2342517421 5.147 2004     2700      1     3.215
## 49373   3843094224 4.764 2004     2700      1     3.398
## 49502  34548102940 3.358 2004     2700      1     2.828
## 49649   2142649221 4.212 2004     2700      1     2.694
## 49664   1942436013 4.103 2004     2700      1     2.737
```

```
## 49710   3042762336 4.314 2004     2700      1     2.948
## 49722   1842864234 5.018 2004     2700      1     3.036
## 49739   1842815777 4.502 2004     2700      1     2.521
## 49745   1842433746 3.730 2004     2700      1     2.568
## 49967  12144288049 4.590 2004     2700      1     2.608
## 49969   1642363460 3.700 2004     2700      1     2.537
## 49977   1642296706 3.739 2004     2700      1     2.576
## 50062  10744220250 4.551 2004     2700      1     2.570
## 50063  10744225301 5.224 2004     2700      1     3.242
## 50136   2342434822 3.104 2004     2700      1     2.574
## 50332  10744233940 4.677 2004     2700      1     2.695
## 50335   1342331884 3.850 2004     2700      1     3.320
## 50413  10744223871 4.045 2004     2700      1     3.515
## 50414   1342288777 4.142 2004     2700      1     2.624
## 50613   1642499234 4.500 2004     2700      1     2.518
## 50643   0345872128 4.849 2004     2700      1     2.867
## 50674   0346599193 4.602 2004     2700      1     3.236
## 50678   0348047527 4.767 2004     2700      1     2.786
## 50712   0346787782 3.666 2004     2700      1     3.188
## 50713   0346787909 3.820 2004     2700      1     2.657
## 50864  13844322241 3.302 2004     2700      1     2.772
## 50998   1542571327 3.817 2004     2700      1     2.654
## 51821   3242733807 3.931 2004     2700      1     2.565
## 51895   3442899889 4.734 2004     2700      1     4.204
## 52090   4944231251 3.072 2004     2700      1     2.594
## 52387   3142514201 3.624 2004     1300      2     2.511
## 52521  29544447206 3.996 2005     2700      1     2.852
## 52543  29144490030 3.866 2005     2700      1     2.519
## 52545  29144533834 4.497 2005     2700      1     2.532
## 52559  28944442441 3.671 2005     2700      1     3.158
## 52561  28944445445 4.451 2005     2700      1     2.538
## 52569  28944447646 4.919 2005     2700      1     2.956
## 52607  28944437578 4.673 2005     2700      1     2.759
## 52767  30144444279 4.733 2005     2700      1     2.768
## 53762  28144433147 4.679 2005     2700      1     2.716
## 53772  28144451163 4.332 2005     2700      1     2.984
## 53795  27744494434 4.618 2005     2700      1     2.704
## 53807  27744496592 4.514 2005     2700      1     3.368
## 53820  27744606737 4.954 2005     2700      1     2.989
## 53845  27644443333 4.575 2005     2700      1     3.277
## 53883  27644568878 4.218 2005     2700      1     2.920
## 54155  27244440305 4.455 2005     2700      1     2.542
## 54170  26844536978 5.830 2005     2700      1     3.865
## 54171  26844552488 5.412 2005     2700      1     3.447
## 54178  26844494491 4.633 2005     2700      1     3.337
## 54179  26844538114 4.966 2005     2700      1     3.618
## 54208  26444452073 5.212 2005     2700      1     3.249
## 54217  26444574824 4.419 2005     2700      1     2.869
## 54246  25844509466 4.220 2005     2700      1     2.873
## 54319  26444598506 4.478 2005     2700      1     2.513
```

```
## 54402  34548099673 3.256 2005     2700      1     2.743
## 54562  25144470719 3.920 2005     2700      1     3.407
## 54575  25144456112 5.883 2005     2700      1     3.918
## 54578  25144455560 3.277 2005     2700      1     2.764
## 54580  25144502974 3.881 2005     2700      1     2.735
## 54581  25144505285 3.742 2005     2700      1     3.229
## 54582  25144518364 4.074 2005     2700      1     2.523
## 54632  24644439190 4.674 2005     2700      1     3.123
## 54633  24644443217 4.833 2005     2700      1     2.868
## 54634  24644460574 3.447 2005     2700      1     2.986
## 54742  24744470476 4.035 2005     2700      1     2.687
## 54827  28844477526 3.557 2005     2700      1     3.096
## 55093  23944469073 3.391 2005     2700      1     2.878
## 55102  23944448617 3.379 2005     2700      1     2.866
## 55113  23844539714 4.976 2005     2700      1     3.011
## 55164  23744459272 3.820 2005     2700      1     3.307
## 55238  23644439061 4.516 2005     2700      1     2.551
## 55428  22844445230 4.774 2005     2700      1     2.811
## 55430  22844447051 4.585 2005     2700      1     2.620
## 55438  22844434877 4.780 2005     2700      1     2.815
## 55442  22844448780 4.222 2005     2700      1     2.874
## 55470  22344457517 3.205 2005     2700      1     2.692
## 55471  22344457679 4.119 2005     2700      1     2.568
## 55485  22244446183 4.420 2005     2700      1     2.506
## 55571  21444450527 4.501 2005     2700      1     2.588
## 55887  20444481448 3.631 2005     2700      1     3.118
## 55888  20444501831 4.934 2005     2700      1     2.969
## 55914  20544462061 4.452 2005     2700      1     2.539
## 55957  21144448879 4.965 2005     2700      1     3.000
## 55970  18844374582 3.884 2005     2700      1     2.586
## 55989  19644391940 4.484 2005     2700      1     2.934
## 56002  20444457518 4.596 2005     2700      1     3.298
## 56318  18944367562 4.915 2005     2700      1     2.952
## 56337  18644370841 4.807 2005     2700      1     2.893
## 56349  18244371651 5.531 2005     2700      1     3.566
## 56378  18744406314 3.417 2005     1000      2     2.904
## 56397  18744406319 3.706 2005     2700      1     2.560
## 56407  17944362664 4.751 2005     2700      1     2.788
## 56491  20944436157 4.708 2005     2700      1     2.743
## 56656  17644392071 4.823 2005     2700      1     2.860
## 56684  17844402965 4.271 2005     2700      1     3.125
## 56691  17244370719 3.361 2005     2700      1     2.848
## 56693  17244379193 4.136 2005     2700      1     3.623
## 56702  17144385790 5.319 2005     2700      1     3.769
## 56705  17144430418 4.592 2005     2700      1     2.627
## 56707  21544449115 4.018 2005     2700      2     2.671
## 56712  17144398838 3.739 2005     2700      1     2.593
## 56748  15744372810 5.001 2005     2700      1     3.036
## 56756  16444379380 3.627 2005     2700      1     3.114
## 56757  20144386837 4.994 2005     2700      1     3.080
```

```
## 56965  15744378579 5.147 2005     2700      1     3.233
## 56967  20144366696 4.939 2005     2700      1     2.974
## 57041  14744275842 4.576 2005     2700      1     2.611
## 57048  14544304095 5.144 2005     2700      1     3.179
## 57417  13444256401 4.619 2005     2700      1     2.654
## 57420  13444293219 5.349 2005     2700      1     3.384
## 57426  13444309949 4.491 2005     2700      1     3.978
## 57437  13444263619 4.684 2005     2700      1     2.719
## 57678  12544253745 4.628 2005     2700      1     4.115
## 57680  19944431791 3.425 2005     2700      1     2.912
## 57683  12544253835 4.548 2005     2700      1     3.200
## 57710  12244277647 3.277 2005     2700      1     2.764
## 57732  11844268070 4.117 2005     2700      1     2.566
## 57734  13744249546 3.742 2005     2700      1     3.281
## 57784  19944429519 4.900 2005     2700      1     2.935
## 57788  11244331742 4.906 2005     2700      1     2.941
## 57831  12744263350 4.085 2005     2700      1     2.584
## 57836  12744278427 3.829 2005     2700      1     2.685
## 57874  13444293166 3.304 2005     2700      1     2.791
## 58080  16244415880 3.909 2005     2700      1     2.611
## 58156  17844391245 3.929 2005     2700      1     3.416
## 58586  25144445294 3.355 2005     2700      1     2.842
## 59279  33846642542 3.639 2005     2700      1     3.126
## 59588  33845880922 5.714 2006     2700      1     3.812
## 59601  33845890326 4.968 2006     2700      1     3.016
## 59636  33845707784 4.606 2006     2700      1     2.654
## 59643  33845717310 4.910 2006     2700      1     3.010
## 59679  33845490014 6.122 2006     2700      1     4.170
## 59681  33845500545 4.941 2006     2700      1     2.989
## 59686  33845478495 4.306 2006     2700      1     3.174
## 59687  33845485717 3.595 2006     2700      1     2.514
## 59689  33947325340 4.051 2006     2700      1     3.603
## 59705  33845449051 4.316 2006     2700      1     3.183
## 59714  33845450660 4.586 2006     2700      1     3.251
## 59746  33748560432 4.653 2006     2700      1     2.701
## 59798  33845354407 3.759 2006     2700      1     2.626
## 60010  33847190697 3.751 2006     2700      1     3.251
## 60017  33847400559 3.416 2006     2700      1     2.916
## 60134  34247572414 3.275 2006     2700      1     2.827
## 61363  33751502226 3.551 2006     2700      1     3.103
## 61367  33751529506 4.659 2006     2700      1     2.708
## 61370  33751545838 5.100 2006     2700      1     3.149
## 61389  33751185526 3.780 2006     2700      1     3.332
## 61412  33751206860 4.493 2006     2700      1     2.541
## 61413  33751217149 4.658 2006     2700      1     2.706
## 61441  33750983605 5.495 2006     2700      1     3.543
## 61499  33750726707 4.632 2006     2700      1     2.732
## 61500  33750731595 4.693 2006     2700      1     2.741
## 61532  33750500326 4.054 2006     2700      1     3.554
## 61570  33750532312 5.148 2006     2700      1     3.246
```

```
## 61574  33750584214 5.548 2006     2700      1     3.597
## 62170  33750346251 4.432 2006     2700      1     2.532
## 62213  33750049670 3.545 2006     2700      1     3.045
## 62216  33750081777 3.983 2006     2700      1     2.851
## 62223  33750106228 4.641 2006     2700      1     2.689
## 62224  33750121615 4.499 2006     2700      1     2.547
## 62260  33749597936 4.523 2006     2700      1     2.572
## 62261  33749599695 5.460 2006     2700      1     3.558
## 62263  33749618085 4.879 2006     2700      1     2.927
## 62273  33749613119 5.139 2006     2700      1     3.237
## 62296  33749440307 3.374 2006     2700      1     2.874
## 62297  33749445317 5.971 2006     2700      1     4.019
## 62331  33749441325 4.657 2006     2700      1     2.705
## 62430  33750895725 3.224 2006     2700      1     2.724
## 62539  39049174296 3.751 2006     2700      1     3.251
## 62724  43749107283 4.870 2006     2700      1     2.918
## 62896  33748976348 3.449 2006     2700      1     3.001
## 62995  33748696778 3.130 2006     2700      1     2.630
## 63010  33748354709 3.584 2006     2700      1     2.501
## 63033  33748433966 3.099 2006     2700      1     2.599
## 63048  33748438134 4.637 2006     2700      1     2.685
## 63132  33748312093 4.065 2006     2700      1     2.526
## 63833  33748161812 4.051 2006     2700      1     2.715
## 63860  33747834620 3.716 2006     2700      1     2.583
## 63864  33747870163 5.283 2006     2700      1     3.331
## 63887  33747135682 4.969 2006     2700      1     3.017
## 63904  33747343208 5.434 2006     2700      1     3.482
## 63948  33746875641 4.520 2006     2700      1     2.568
## 63950  33746897394 3.283 2006     2700      1     2.783
## 64006  33746437130 4.877 2006     2700      1     2.926
## 64083  33748046126 3.747 2006     2700      1     3.247
## 64377  33746430561 3.742 2006     2700      1     2.609
## 64421  33746075560 4.642 2006     2700      1     2.690
## 64422  33746088001 5.656 2006     2700      1     3.754
## 64482  33745890309 4.593 2006     2700      1     2.642
## 64484  33745913326 3.707 2006     2700      1     3.259
## 64514  33745698685 4.688 2006     2700      1     2.788
## 64520  33745611449 5.100 2006     2700      1     3.198
## 64522  33745614361 5.082 2006     2700      1     3.180
## 64537  33745851479 4.385 2006     2700      1     2.846
## 64543  33745940542 3.067 2006     2700      1     2.567
## 64563  33745712365 3.631 2006     2700      1     3.131
## 64968  33745591698 3.436 2006     2700      1     2.936
## 64979  33745661633 3.776 2006     2700      1     3.328
## 65008  33745227399 4.917 2006     2700      1     2.965
## 65009  33745227751 3.734 2006     2700      1     2.601
## 65010  33745255382 4.457 2006     2700      1     2.557
## 65030  33646360865 4.352 2006     2700      1     3.852
## 65051  33745081608 5.495 2006     2700      1     3.593
## 65054  33745102555 5.215 2006     2700      1     3.263
```

```
## 65093  33744718154 3.600 2006     2700      1     2.517
## 65096  33744735681 3.099 2006     2700      1     2.599
## 65106  33744966030 4.777 2006     2700      1     2.825
## 65135  33646825385 4.208 2006     2700      1     3.708
## 65136  33646835423 3.449 2006     2700      1     2.949
## 65145  33646948881 3.534 2006     2700      1     3.034
## 65147  33744455686 3.034 2006     2700      1     2.534
## 65365  33846406448 3.854 2006     2700      1     2.721
## 65669  33646859471 4.708 2006     2700      1     2.806
## 65716  33646676272 4.484 2006     2700      1     3.984
## 65765  33646488054 4.825 2006     2700      1     2.874
## 65772  33646450281 3.851 2006     2700      1     2.516
## 65773  33646452922 4.459 2006     2700      1     2.507
## 65816  30044438368 5.433 2006     2700      1     3.533
## 65833  33646138925 3.089 2006     2700      1     2.589
## 65838  33646178951 5.085 2006     2700      1     3.133
## 65975  33745489614 3.646 2006     2700      1     3.198
## 66271  33646148492 3.436 2006     2700      1     2.936
## 66300  33645865234 4.803 2006     2700      1     2.851
## 66319  33646052556 5.164 2006     2700      1     3.212
## 66344  33645729203 4.579 2006     2700      1     2.627
## 66372  33645523067 6.444 2006     2700      1     4.544
## 66394  29944442726 4.794 2006     2700      1     2.842
## 66423  33645530745 4.561 2006     2700      1     2.609
## 66489  33646412049 3.258 2006     2700      1     2.810
## 66490  33646420626 3.788 2006     2700      1     3.288
## 66492  33646432226 4.312 2006     2700      1     3.028
## 66752  33645213907 3.442 2006     2700      1     2.942
## 66778  33645399959 4.643 2006     2700      1     2.691
## 66806  33645102811 3.897 2006     2700      1     3.449
## 66807  33645103550 5.432 2006     2700      1     3.944
## 66813  33645068471 4.040 2006     2700      1     2.502
## 66814  33645082502 3.442 2006     2700      1     2.994
## 66846  33644982043 3.275 2006     2700      1     2.775
## 66880  33644833147 4.678 2006     2700      1     2.726
## 66883  33644807437 5.185 2006     2700      1     3.234
## 66925  33644584065 4.544 2006     2700      1     2.593
## 66926  33644605043 5.148 2006     2700      1     3.812
## 66958  33644792884 3.120 2006     2700      1     2.620
## 67034  33645838177 2.975 2006     2700      1     2.527
## 67354  33344467770 4.351 2006     2700      1     3.851
## 67363  33644536465 4.565 2006     2700      1     2.613
## 67394  32644457434 4.090 2006     2700      1     2.957
## 67395  32644467389 5.178 2006     2700      1     3.227
## 67432  32144436098 4.713 2006     2700      1     2.762
## 67446  32144441493 4.605 2006     2700      1     2.703
## 67447  32144443648 4.776 2006     2700      1     2.876
## 67510  32444443951 4.627 2006     2700      1     2.676
## 67786  31344467254 4.933 2006     2700      1     3.647
## 67787  31344470705 5.219 2006     2700      1     3.268
```

```
## 67815  30944457531 4.478 2006     2700      1     2.578
## 67854  30444460269 3.615 2006     2700      1     3.115
## 67876  32244433703 3.636 2006     2700      1     2.553
## 67938  31044436630 4.688 2006     2700      1     2.736
## 67942  31044446194 4.154 2006     2700      1     2.666
## 67950  31044456529 5.039 2006     2700      1     3.087
## 67972  32044434020 3.847 2006     2700      1     2.511
## 67974  32044435429 4.349 2006     2700      1     3.014
## 68834  37349119004 2.993 2007     2700      1     2.553
## 68923  37249041477 3.789 2007     2700      1     3.349
## 68924  37249061691 4.883 2007     2700      1     2.991
## 68931  37149028276 4.444 2007     2700      1     2.603
## 68952  38449085341 3.654 2007     2700      1     3.214
## 68962  37249065591 3.858 2007     2700      1     2.635
## 69060  34249853586 2.941 2007     2700      1     2.501
## 69120  35748981241 2.976 2007     2700      1     2.536
## 69140  36549022943 2.959 2007     2700      1     2.519
## 69155  36849011221 4.476 2007     2700      1     2.584
## 69239  38049053061 4.415 2007     2700      1     2.523
## 69241  38049087576 4.711 2007     2700      1     3.233
## 69579  41649108328 3.041 2007     2700      1     2.601
## 69604  42649084025 3.041 2007     2700      1     2.601
## 69806  54449102231 3.220 2007     2700      1     2.780
## 70668  36549006511 3.668 2007     2700      1     3.228
## 70728  36349035495 4.062 2007     2700      1     2.585
## 70766  36148981435 4.360 2007     2700      1     3.920
## 70798  36148951701 3.556 2007     2700      1     3.168
## 70800  36149000149 5.380 2007     2700      1     4.105
## 70829  35848938907 3.182 2007     2700      1     2.742
## 70832  35848963822 4.711 2007     2700      1     2.819
## 70833  35848964343 4.234 2007     2700      1     3.846
## 70835  35848969753 4.853 2007     2700      1     3.578
## 70842  35848935201 5.343 2007     2700      1     3.866
## 70844  35848964571 4.402 2007     2700      1     3.381
## 70845  35848968871 6.256 2007     2700      1     4.364
## 70888  35248896228 3.992 2007     2700      1     2.718
## 70910  35748974550 4.877 2007     2700      1     4.437
## 70935  36048968753 4.402 2007     2700      1     2.510
## 71265  41149161956 4.099 2007     2700      1     2.673
## 71308  48849113878 4.849 2007     2700      1     2.957
## 71622  35348961960 3.335 2007     2700      1     2.895
## 71624  35348997893 4.319 2007     2700      1     3.297
## 71653  35148831175 4.279 2007     2700      1     3.257
## 71668  35248820655 2.959 2007     2700      1     2.519
## 71669  35248829599 4.645 2007     2700      1     2.805
## 71672  35248893940 3.233 2007     2700      1     2.793
## 71673  35248895346 4.762 2007     2700      1     2.870
## 71714  34948849235 3.498 2007     2700      1     3.110
## 71717  34948891641 4.829 2007     2700      1     2.937
## 71753  34548827323 4.070 2007     2700      1     2.997
```

```
## 71755  34548835744 4.466 2007     2700      1     2.574
## 71864  35649003880 4.876 2007     2700      1     3.602
## 71959  38449098511 3.009 2007     2700      1     2.621
## 72327  34648832073 4.784 2007     2700      1     3.509
## 72400  34548849816 4.178 2007     2700      1     3.738
## 72423  34548551610 2.976 2007     2700      1     2.536
## 72454  34548776934 3.753 2007     2700      1     3.313
## 72465  34548570632 4.484 2007     2700      1     2.592
## 72494  34547426346 3.681 2007     2700      1     2.608
## 72546  34548418912 4.594 2007     2700      1     2.702
## 72548  34548433987 4.078 2007     2700      1     3.638
## 72555  34548426480 4.019 2007     2700      1     2.745
## 72606  34547502731 3.531 2007     2700      1     2.510
## 72609  34547545269 4.719 2007     2700      1     3.496
## 72613  34548040666 3.695 2007     2700      1     3.255
## 72614  34548050674 3.041 2007     2700      1     2.601
## 72623  34548262715 4.493 2007     2700      1     2.603
## 72673  34648836403 4.569 2007     2700      1     4.181
## 72822  36248973631 3.841 2007     2700      1     2.768
## 73331  38849137287 3.410 2007     2700      2     3.022
## 73480  34548014259 2.941 2007     2700      1     2.501
## 73492  34548066571 3.759 2007     2700      1     3.319
## 73493  34548075820 5.741 2007     2700      1     3.849
## 73500  34548092397 4.740 2007     2700      1     3.465
## 73516  34547872324 3.208 2007     2700      1     2.768
## 73550  34547942630 3.548 2007     2700      1     3.108
## 73551  34547950500 4.540 2007     2700      1     3.317
## 73558  34547876797 4.649 2007     2700      1     2.757
## 73559  34547906405 4.663 2007     2700      1     2.771
## 73594  34547757915 5.372 2007     2700      1     3.482
## 73683  34547630852 3.983 2007     2700      1     2.557
## 73698  34547683917 3.195 2007     2700      1     2.755
## 74193  34547148985 4.266 2007     2700      1     3.826
## 74194  34547170053 6.085 2007     2700      1     5.012
## 74226  34447521458 4.639 2007     2700      1     2.749
## 74241  34447544253 3.830 2007     2700      1     3.390
## 74246  34447300446 4.669 2007     2700      1     3.191
## 74247  34447512756 4.710 2007     2700      1     2.869
## 74248  34447520135 3.406 2007     2700      1     2.966
## 74249  34447520136 5.342 2007     2700      1     3.450
## 74250  34447558140 5.428 2007     2700      1     4.153
## 74312  34447340945 3.992 2007     2700      1     2.919
## 74385  34447137331 4.461 2007     2700      1     2.571
## 74426  34447514261 3.056 2007     2700      1     2.668
## 74704  35248881664 4.126 2007     2700      1     2.851
## 74716  35348930399 3.169 2007     2700      1     2.781
## 75040  34548567516 3.824 2007     2700      1     2.599
## 75044  34347210370 4.041 2007     2700      1     3.601
## 75047  34347255039 4.665 2007     2700      1     2.775
## 75085  34250696907 4.873 2007     2700      1     3.033
```

```
## 75089  34250779802 4.066 2007     2700      1     3.626
## 75142  34250212715 6.307 2007     2700      1     5.285
## 75144  34250309164 4.292 2007     2700      1     3.219
## 75145  34250330354 3.291 2007     2700      1     2.851
## 75168  34249899377 4.920 2007     2700      1     3.030
## 75203  34250027957 3.983 2007     2700      1     2.506
## 75236  34248149838 3.765 2007     2700      1     3.325
## 75239  34248169298 3.795 2007     2700      1     2.722
## 75241  34248179572 2.993 2007     2700      1     2.553
## 75242  34248201151 4.476 2007     2700      1     4.036
## 75260  34249662628 3.925 2007     2700      1     3.485
## 75274  34249913494 4.647 2007     2700      1     2.757
## 75275  34249915334 4.224 2007     2700      1     3.151
## 75325  34347261062 3.498 2007     2700      1     3.058
## 75384  34547098109 2.941 2007     2700      1     2.501
## 75506  34548404625 3.233 2007     2700      1     2.793
## 75793  34547309949 3.854 2007     2700      2     3.414
## 75961  34249830160 4.754 2007     2700      1     2.913
## 75990  34249657825 4.383 2007     2700      1     3.943
## 75993  34249673868 4.807 2007     2700      1     3.532
## 76036  34249047454 4.895 2007     2700      1     3.822
## 76043  34248559834 4.710 2007     2700      1     3.232
## 76060  34248183628 3.515 2007     2700      1     3.075
## 76121  34247615981 3.835 2007     2700      1     2.560
## 76133  34247862190 5.281 2007     2700      1     3.389
## 76835  34247863975 5.154 2007     2700      1     3.264
## 76849  34247476744 5.028 2007     2700      1     3.136
## 76851  34247517444 3.453 2007     2700      1     3.013
## 76865  34247498668 4.800 2007     2700      1     2.910
## 76883  34247167109 3.675 2007     2700      1     2.604
## 76903  34247241630 5.117 2007     2700      1     3.225
## 76904  34247256786 3.471 2007     2700      1     3.031
## 76946  34247144499 4.185 2007     2700      1     2.707
## 76947  34247149303 4.041 2007     2700      1     3.601
## 76950  34247523621 4.713 2007     2700      1     3.640
## 76959  34047096257 4.791 2007     2700      1     2.950
## 76970  34047243811 4.396 2007     2700      1     2.555
## 76972  34047268898 3.900 2007     2700      1     3.460
## 76981  34047237367 5.526 2007     2700      1     3.685
## 76988  34147096980 4.461 2007     2700      1     2.569
## 76993  34147124264 4.538 2007     2700      1     2.648
## 77024  33947623101 3.025 2007     2700      1     2.585
## 77049  34247148137 3.071 2007     2700      1     2.631
## 77052  34247180752 3.376 2007     2700      1     2.988
## 77570  34247282706 4.091 2007     2700      1     3.070
## 77574  34147113263 3.593 2007     1000      2     2.520
## 77582  34047185285 3.728 2007     2700      1     2.655
## 77604  33947210018 3.884 2007     2700      1     3.444
## 77607  33947595236 5.510 2007     2700      1     3.670
## 77613  33947491227 3.425 2007     2700      1     2.985
```

```
## 77641  33947176599 2.941 2007     2700      1     2.501
## 77661  33947196096 4.621 2007     2700      1     2.729
## 77673  33947243452 3.034 2007     1000      2     2.594
## 77685  33947271991 3.734 2007     2700      1     3.294
## 77712  33847705701 5.514 2007     2700      1     3.624
## 77720  33847723397 4.950 2007     2700      1     3.109
## 77761  33847253945 4.880 2007     2700      1     2.990
## 77772  33847368173 5.051 2007     2700      1     3.159
## 77773  33847369469 4.655 2007     2700      1     2.763
## 77776  33847381116 5.213 2007     2700      1     3.938
## 77779  33847404482 4.431 2007     2700      1     2.539
## 77786  33847413701 3.920 2007     2700      1     3.480
## 77790  33847675339 3.071 2007     2700      1     2.683
## 77839  33947710793 4.749 2007     2700      1     2.857
## 78101  34250364999 3.071 2007     2700      1     2.683
## 78607  33847353773 4.810 2007     2700      1     2.970
## 78634  33847040329 3.056 2007     2700      1     2.616
## 78651  33847165011 4.126 2007     2700      1     3.686
## 78654  33847103656 4.307 2007     2700      1     2.879
## 78678  33846940450 3.268 2007     2700      1     2.828
## 78689  33846944669 2.959 2007     2700      1     2.519
## 78695  33847766654 3.233 2007     2700      1     2.845
## 78707  33846996345 5.465 2007     2700      1     4.037
## 78725  33846826988 3.071 2007     2700      1     2.683
## 78737  33846866271 4.955 2007     2700      1     4.515
## 78738  33846875851 5.005 2007     2700      1     3.113
## 78746  33846863146 4.895 2007     2700      1     3.003
## 78812  33846673816 4.935 2007     2700      1     3.043
## 78813  33846688887 5.379 2007     2700      1     3.489
## 78822  33846794896 5.202 2007     2700      1     3.310
## 78832  33846817495 4.130 2007     2700      1     3.108
## 79061  34248581736 3.523 2007     2700      1     3.135
## 79072  34249676415 3.086 2007     2700      1     2.698
## 79356  33846440513 3.498 2007     2700      1     3.058
## 79377  33846294746 5.096 2007     2700      1     4.023
## 79416  33846126939 4.946 2007     2700      1     3.056
## 79420  33846175081 3.777 2007     2700      1     2.704
## 79427  33846125284 5.173 2007     2700      1     3.281
## 79434  33845921647 3.025 2007     2700      1     2.585
## 79489  33845970192 5.126 2007     2700      1     3.234
## 79495  33845994359 4.653 2007     2700      1     2.763
## 79510  33845984538 4.757 2007     2700      1     3.279
## 79512  33846026341 4.599 2007     2700      1     2.707
## 79537  33845783867 3.086 2007     2700      1     2.646
## 79540  33845870422 3.734 2007     2700      1     2.509
## 79704  33847694390 3.156 2007     2700      1     2.716
## 79797  34147163870 2.905 2007     2700      1     2.517
## 79957  34250207328 2.959 2007     2700      1     2.519
## 80013  34347343739 3.668 2007     2700      1     2.595
## 80106  35649016098 4.629 2007     2700      1     2.737
```

```
## 80112  35848954757 4.777 2007     2700      1     2.937
## 80217  42449096529 3.935 2007     2700      1     2.913
## 80636  58149083708 4.778 2008     2700      1     2.928
## 80649  57749186668 3.127 2008     1213      3     2.781
## 80683  58049213696 4.583 2008     2700      1     2.733
## 80744  57449112519 4.616 2008     2700      1     2.766
## 80796  56849101309 2.883 2008     2700      1     2.537
## 80797  57349093766 3.663 2008     2700      1     3.265
## 80880  47249111218 3.230 2008     2700      1     2.832
## 80918  55249083968 3.131 2008     2700      1     2.733
## 80942  55949137877 3.172 2008     2700      1     2.774
## 80980  57449084208 4.389 2008     2700      1     2.539
## 81062  58149387739 4.794 2008     2700      1     3.763
## 81205  61549115925 2.853 2008     2700      1     2.507
## 82702  57149091205 3.088 2008     2700      1     2.690
## 82785  56649112752 4.452 2008     2700      1     2.602
## 82845  56249128276 4.627 2008     2700      1     2.777
## 82846  56249141853 3.916 2008     2700      1     2.885
## 82950  54049103289 3.570 2008     2700      1     3.172
## 83609  55249119765 4.514 2008     2700      1     3.077
## 83648  55549139481 4.069 2008     2700      1     2.835
## 83714  54049121079 3.886 2008     2700      1     3.488
## 83794  53749095835 4.145 2008     2700      1     3.747
## 83849  54849146500 4.137 2008     2700      1     3.739
## 83854  53349178089 3.534 2008     1000      2     2.503
## 83859  54949141082 3.589 2008     1704      3     2.558
## 83888  52949098281 3.030 2008     2700      1     2.632
## 83891  52949103244 3.781 2008     2700      1     2.802
## 83897  52949144507 4.137 2008     2700      1     3.739
## 83898  52949144863 3.722 2008     2700      1     3.376
## 84213  65549087532 3.839 2008     2700      1     2.860
## 84585  52649120886 4.627 2008     2700      1     2.778
## 84607  51949095058 4.868 2008     2700      1     3.070
## 84728  50949089029 4.489 2008     2700      1     4.091
## 84734  51649088233 4.453 2008     2700      1     2.603
## 84831  52749095854 3.593 2008     2700      1     2.562
## 84832  52749096126 3.437 2008     2700      1     3.039
## 85277  65249106599 3.570 2008     2700      1     2.591
## 85341  80051646958 3.099 2008     2700      1     2.701
## 85698  50449093260 4.360 2008     2700      1     2.510
## 85709  50449095362 4.693 2008     2700      1     3.459
## 85743  50149121231 4.030 2008     1000      2     2.999
## 85752  49949087905 4.118 2008     2700      1     3.139
## 85753  49949091204 3.678 2008     2700      1     3.280
## 85827  49449100226 3.274 2008     2700      1     2.876
## 85831  49249087104 4.356 2008     2700      1     2.558
## 85883  49249092086 3.274 2008     2700      1     2.876
## 85886  49249104701 5.150 2008     2700      1     3.300
## 85962  48149090717 3.855 2008     2700      1     3.457
## 86505  47949092540 4.171 2008     2700      1     3.825
```

```
## 86507  47949114668 5.232 2008     2700      1     3.383
## 86565  47549105407 4.511 2008     2700      1     2.662
## 86680  46449113799 4.877 2008     2700      1     3.028
## 86729  45749146166 3.855 2008     2700      1     2.672
## 86742  46449098821 3.879 2008     2700      1     2.697
## 86755  46749156706 4.460 2008     2700      1     2.611
## 86795  48749129542 4.693 2008     2700      1     3.714
## 87540  45949092534 4.637 2008     2700      1     2.787
## 87543  45949107235 3.831 2008     2700      1     3.433
## 87544  46749131798 5.204 2008     2700      1     3.354
## 87602  45349084317 4.501 2008     2700      1     2.651
## 87603  45349103702 4.595 2008     2700      1     2.746
## 87663  45149127545 4.115 2008     2700      1     3.717
## 87691  44849103815 4.757 2008     2700      1     3.574
## 87696  44849115945 4.501 2008     2700      1     2.651
## 87739  45749114895 4.710 2008     2700      1     2.861
## 88004  54249162813 3.201 2008     2700      1     2.803
## 88491  44349145194 4.455 2008     2700      1     2.605
## 88501  44349165737 5.327 2008     2700      1     4.145
## 88546  44249122799 5.169 2008     2700      1     4.138
## 88548  44649131269 3.291 2008     2700      1     2.893
## 88558  43649093343 3.030 2008     2700      1     2.632
## 88637  43549110959 4.509 2008     2700      1     2.660
## 88680  43049170245 4.574 2008     2700      1     2.776
## 88682  43049174930 3.968 2008     2700      1     2.533
## 88683  43049176549 4.334 2008     2700      1     2.536
## 88751  42949153474 3.484 2008     2700      1     3.086
## 88760  43049095682 4.414 2008     2700      1     2.564
## 88998  49149105211 3.142 2008     2700      1     2.796
## 89489  42549101925 3.437 2008     2700      1     3.039
## 89498  42449106956 4.533 2008     2700      1     2.733
## 89501  44149123099 4.666 2008     2700      1     2.866
## 89567  42249098105 4.181 2008     2700      1     2.744
## 89601  42249083269 4.686 2008     2700      1     2.886
## 89602  42249092533 4.619 2008     2700      1     2.819
## 89703  40749147526 3.324 2008     2700      1     2.926
## 89729  41449101361 3.316 2008     2700      1     2.918
## 90277  43049134750 3.647 2008     2700      1     2.616
## 90287  41449112372 5.371 2008     2700      1     3.521
## 90289  41449113183 4.795 2008     2700      1     2.945
## 90338  40949100835 3.944 2008     2700      1     2.762
## 90346  40949098491 2.940 2008     2700      1     2.542
## 90347  40949102607 5.086 2008     2700      1     3.237
## 90361  40949127393 4.633 2008     2700      1     3.652
## 90407  40849124009 4.384 2008     2700      1     2.534
## 90409  40849142102 4.916 2008     2700      1     3.067
## 90415  40749141596 4.476 2008     2700      1     2.678
## 90432  41649090186 4.800 2008     2700      1     2.950
## 90458  40449141326 4.620 2008     2700      1     4.222
## 90521  40449086087 4.445 2008     2700      1     2.596
```

```
## 90930  50849106701 3.615 2008     2700      1     2.585
## 91383  39549100109 3.988 2008     2700      1     2.551
## 91384  39549102834 4.683 2008     2700      1     2.833
## 91386  39549111874 3.958 2008     2700      1     2.927
## 91415  39749102149 4.416 2008     2700      1     2.616
## 91490  38949188680 3.871 2008     2700      1     3.473
## 91565  38849174979 4.368 2008     2700      1     2.570
## 91993  38749111066 5.445 2008     2700      1     3.595
## 91995  38749127156 4.616 2008     2700      1     2.766
## 92033  38549088096 4.842 2008     2700      1     2.992
## 92034  38549110771 3.088 2008     2700      1     2.690
## 92040  38549177642 4.453 2008     2700      1     2.603
## 92046  38349164176 4.531 2008     2700      1     2.681
## 92068  38349049478 5.357 2008     2700      1     3.507
## 92111  38049146378 4.148 2008     2700      1     3.750
## 92113  38049169559 4.935 2008     2700      1     3.085
## 92134  38049062200 4.629 2008     2700      1     3.395
## 92136  38049077991 3.843 2008     2700      1     3.445
## 92140  38049082149 4.298 2008     2700      1     2.861
## 92161  38149141767 3.958 2008     2700      1     2.523
## 92193  37549009899 4.066 2008     2700      1     2.884
## 92199  37549027612 4.244 2008     2700      1     3.213
## 92219  38149074110 3.934 2008     2700      1     3.536
## 92220  38149075361 2.853 2008     2700      1     2.507
## 92922  57149102701 3.912 2008     2700      1     2.729
## 93198  42549116694 2.966 2008     2700      2     2.568
## 93527  38349164474 3.310 2008     1400      6     2.912
## 93540  73949085764 3.241 2009     2700      1     2.824
## 93592  73549083853 3.805 2009     2700      1     3.388
## 93633  70449636557 4.842 2009     2700      1     2.975
## 93654  70350738289 4.729 2009     2700      1     2.860
## 93760  73349094732 4.184 2009     2700      1     2.984
## 93775  84870055817 4.573 2009     2700      1     2.755
## 93886  67650735607 3.294 2009     2700      1     2.877
## 93887  67650745371 3.144 2009     2700      1     2.727
## 93889  67650760521 3.826 2009     2700      1     3.409
## 93892  67650766434 3.175 2009     2700      1     2.758
## 93961  70749106456 3.134 2009     2700      1     2.717
## 93962  70749114018 3.645 2009     2700      1     3.228
## 93964  70749145982 3.814 2009     2700      1     3.397
## 94224  77049105843 3.904 2009     2700      1     2.704
## 94263  77449156169 3.067 2009     2700      1     2.650
## 94281  77649213935 3.838 2009     2700      1     2.840
## 95604  72049129430 4.683 2009     2700      1     2.814
## 95681  70449346890 2.981 2009     2700      4     2.564
## 95723  70349873173 5.082 2009     2700      1     3.265
## 95726  70449636163 5.302 2009     2700      1     3.485
## 95743  77950669604 5.246 2009     2700      1     3.377
## 95820  70350436639 2.982 2009     2700      1     2.565
## 95842  70449448312 4.397 2009     2700      1     2.528
```

```
## 95900  70449726880 2.956 2009     2700      1     2.539
## 96014  73949153085 4.118 2009     2700      1     3.701
## 96050  74349123981 3.453 2009     2700      1     3.036
## 96572  84881264341 4.415 2009     2700      1     2.546
## 96673  70350064091 3.705 2009     2700      1     2.707
## 96695  70350451433 3.351 2009     2700      1     2.934
## 96696  70350458703 4.870 2009     2700      1     3.053
## 96716  70350212398 4.751 2009     2700      1     2.882
## 96719  70350236228 4.659 2009     2700      1     2.841
## 96753  70349506141 4.798 2009     2700      1     2.980
## 96870  69949144253 3.696 2009     2700      1     2.646
## 96895  70349481938 4.575 2009     2700      1     3.525
## 96897  70349610473 4.853 2009     2700      1     3.601
## 96899  70349625757 5.493 2009     2700      1     3.624
## 96900  70349629019 5.086 2009     2700      1     3.217
## 97145  77952988108 4.600 2009     2700      1     4.183
## 97251  77449128265 2.980 2009     2700      2     2.615
## 97527  70349467696 3.134 2009     2700      1     2.717
## 97550  70349335796 4.765 2009     2700      1     2.896
## 97633  70349238733 4.938 2009     2700      1     3.120
## 97634  70349244812 4.857 2009     2700      1     3.039
## 97636  70349254065 3.850 2009     2700      1     3.433
## 97666  66849121699 3.710 2009     2700      1     2.660
## 97668  66849128097 3.947 2009     2700      1     2.897
## 97740  70149093912 5.343 2009     2700      1     3.476
## 97741  70149098511 3.748 2009     2700      1     3.331
## 97742  70149115051 3.343 2009     2700      1     2.926
## 97845  68849107831 3.343 2009     2700      1     2.978
## 97850  68949148762 3.567 2009     2700      1     3.150
## 97851  68949148766 3.184 2009     2700      1     2.767
## 97852  68949150765 3.725 2009     2700      1     3.360
## 98067  73449108112 3.846 2009     2700      1     2.646
## 98073  73449123799 4.061 2009     2700      1     3.644
## 98210  75349114036 3.766 2009     2700      1     3.349
## 98756  69249219296 5.111 2009     2700      1     3.242
## 98813  68949090752 4.850 2009     2700      1     2.983
## 98814  68949092632 4.197 2009     2700      1     2.995
## 98930  68849119553 4.573 2009     2700      1     2.704
## 98931  68849123400 4.335 2009     2700      1     3.083
## 98992  68249154875 5.117 2009     2700      1     3.248
## 99081  67650635227 3.991 2009     2700      1     2.537
## 99104  67651160419 4.452 2009     2700      1     3.047
## 99146  69349094143 4.819 2009     2700      1     2.950
## 99564  68049123594 4.967 2009     2700      1     3.098
## 99608  67650604624 2.956 2009     2700      1     2.539
## 99616  68149144251 3.826 2009     2700      1     3.409
## 99644  67651218924 4.011 2009     2700      1     2.759
## 99677  67649867178 3.302 2009     2700      1     2.885
## 99679  67650318719 3.926 2009     2700      1     2.675
## 99684  67650876227 3.184 2009     2700      1     2.819
```

```
## 99874  67649967683 4.202 2009     2700      1     2.797
## 99913  68049122014 3.725 2009     2700      1     2.726
## 100692 67649390890 3.995 2009     2700      1     2.945
## 100761 67449110743 5.734 2009     2700      1     3.867
## 100762 67449152290 4.549 2009     2700      1     2.731
## 100793 66749147927 3.076 2009     1000      2     2.659
## 100831 67149142042 5.168 2009     2700      1     3.299
## 101054 67650269851 3.623 2009     2700      1     2.575
## 101788 65849187703 5.312 2009     2700      1     3.443
## 101794 66249084112 4.991 2009     2700      1     3.122
## 101808 66149152644 4.179 2009     2700      1     2.725
## 101852 66149113684 3.500 2009     2700      1     3.083
## 101904 66149123748 3.968 2009     1000      2     2.918
## 101994 65449136284 3.759 2009     1303      7     2.507
## 102250 68849130007 3.079 2009     2700      1     2.714
## 102491 64949130692 2.873 2009     2700      2     2.508
## 102867 65449152185 5.067 2009     2700      1     3.198
## 102988 64749093541 4.983 2009     2700      1     3.114
## 103008 64749114159 4.724 2009     2700      1     2.857
## 103806 64749084426 3.321 2009     2400      2     2.904
## 103846 63249124498 3.838 2009     2700      1     2.788
## 103903 62749097289 4.481 2009     2700      1     2.663
## 104012 62149120632 4.393 2009     2700      1     2.576
## 104068 61849118968 4.245 2009     2700      1     3.880
## 104069 61849123616 2.982 2009     2700      1     2.565
## 104079 61549093691 3.134 2009     2700      1     2.717
## 104122 60049085107 2.885 2009     2700      1     2.520
## 104126 60049098815 4.857 2009     2700      1     3.402
## 104132 61549094077 5.418 2009     2700      1     3.600
## 104216 64849088553 2.969 2009     2700      1     2.604
## 104936 66149118531 3.178 2009     2700      3     2.813
## 105045 60649118584 4.654 2009     2700      1     2.837
## 105057 61449094816 3.277 2009     2700      1     2.860
## 105061 61449183682 5.109 2009     2700      1     3.240
## 105118 60849115270 5.804 2009     2700      1     3.935
## 105119 60849137226 3.044 2009     2700      1     2.627
## 105211 60249099307 4.478 2009     2700      1     3.024
## 105240 59249084303 4.172 2009     2700      1     2.717
## 105242 59449101337 4.234 2009     2700      1     3.186
## 105251 59749096464 4.506 2009     2700      1     4.089
## 105271 58949085856 3.525 2009     2700      1     2.526
## 105274 58949102427 3.467 2009     2700      1     3.050
## 105884 58749094444 5.536 2009     2700      1     3.669
## 105947 58349091054 3.908 2009     2700      1     2.657
## 105990 58149385407 4.564 2009     2700      1     2.695
## 105992 58149389215 5.900 2009     2700      1     4.031
## 106003 58149385649 4.940 2009     2700      1     3.073
## 106063 57749092004 3.743 2009     2700      1     3.326
## 106066 57749093663 3.623 2009     2700      1     3.206
## 106107 58249087710 3.985 2009     2700      1     2.733
```

```
## 106110 58249093973 3.204 2009     2700      1     2.787
## 106144 59449111011 3.007 2009     2700      1     2.590
## 106406 64249145106 3.453 2009     2700      1     3.088
## 106907 75549083609 3.044 2009     2700      1     2.679
## 107395 79952376571 2.865 2009     2700      2     2.500
## 107838 78650858488 4.531 2010     2700      1     2.681
## 107918 78649906060 4.743 2010     2700      1     2.893
## 109519 78649429727 4.440 2010     2700      1     2.590
## 109591 78549264482 3.230 2010     2700      1     2.832
## 109892 78049515807 4.522 2010     2700      1     2.672
## 110805 78049351929 4.511 2010     2700      1     3.074
## 111025 77957944426 3.436 2010     2700      1     3.090
## 111043 77957916132 3.591 2010     2700      1     2.560
## 112043 77957341503 4.430 2010     2700      1     2.581
## 112173 77956929653 3.485 2010     2700      1     2.505
## 112558 77956279672 3.546 2010     2700      1     2.516
## 113861 77955877759 5.447 2010     2700      1     3.598
## 115221 77955066199 5.339 2010     2700      1     3.489
## 115295 77954825540 3.164 2010     2700      1     2.818
## 115455 77954421915 4.566 2010     2700      1     2.716
## 115552 77954183132 3.685 2010     2700      1     2.654
## 116866 77953956205 4.405 2010     2700      1     4.059
## 117018 77953522402 3.213 2010     2700      1     2.815
## 117020 77953577951 4.518 2010     2700      1     2.668
## 117104 77953017910 3.069 2010     2700      1     2.671
## 117177 77952766987 3.851 2010     2700      1     2.820
## 117245 77953360245 3.933 2010     2700      1     2.953
## 117571 77952732642 3.053 2010     2700      2     2.655
## 118169 77952760196 3.757 2010     2700      1     2.727
## 118307 77952301687 4.418 2010     2700      1     2.568
## 118351 77951868945 4.370 2010     2700      1     2.520
## 118352 77951874018 4.849 2010     2700      1     3.050
## 119308 77951462316 3.083 2010     2700      1     2.737
## 119348 77950857180 3.456 2010     2700      2     3.058
## 119499 77950486825 3.875 2010     2700      1     2.844
## 119583 77950478147 4.036 2010     2700      1     2.650
## 119585 77950494853 4.319 2010     2700      1     3.087
## 120217 77950246403 3.346 2010     2700      1     2.948
## 120324 77949376759 3.520 2010     2700      1     2.541
## 120445 77449146854 3.611 2010     2700      1     2.580
## 120453 77649329576 4.566 2010     2700      1     2.717
## 120521 77449089099 3.402 2010     2700      1     3.004
## 120522 77449094873 2.941 2010     2700      1     2.595
## 120523 77449140027 2.950 2010     2700      1     2.552
## 120619 77950355514 3.621 2010     2700      1     2.641
## 121099 77949392974 2.965 2010     2700      2     2.567
## 121527 77149136099 3.803 2010     2700      1     2.571
## 121596 76649086071 4.375 2010     2700      1     2.576
## 121682 76149092535 3.000 2010     2700      1     2.602
## 122328 74549140993 5.722 2010     2700      1     4.287
```

```
## 122329 74549159752 5.207 2010     2700      1     3.409
## 122451 74049098673 4.147 2010     2700      1     2.762
## 122453 74049114688 4.434 2010     2700      1     2.584
## 122867 77950651061 4.285 2010     2700      1     3.102
## 125455 84855370035 4.953 2011     2700      1     3.102
## 125458 84855412101 2.999 2011     2700      1     2.548
## 125459 84855425106 5.298 2011     2700      1     3.445
## 125541 84255194001 4.947 2011     2700      1     3.044
## 125681 83055177179 3.952 2011     1000      2     2.868
## 125944 82555168440 4.771 2011     2700      1     2.868
## 127205 82255175382 5.249 2011     2700      1     3.346
## 127216 81855218115 4.488 2011     2700      1     2.585
## 127217 81855227033 3.596 2011     2700      1     3.145
## 127277 81555213207 3.799 2011     2700      1     2.564
## 127287 81455154874 4.549 2011     2700      1     3.465
## 127339 81055139703 3.047 2011     1000      2     2.648
## 127362 80755130317 4.428 2011     2700      1     2.575
## 127364 80755188166 4.426 2011     2700      1     2.523
## 127437 80455162464 4.949 2011     2700      1     3.046
## 127457 80355135138 4.850 2011     2700      1     2.997
## 127547 80054967262 3.163 2011     2700      1     2.712
## 128312 80054769456 4.528 2011     2700      1     2.625
## 128493 80053305924 3.256 2011     2700      1     2.805
## 128577 80053426842 4.065 2011     2700      1     2.982
## 129173 84555215993 3.637 2011     2700      1     3.238
## 129220 80053069652 4.106 2011     2700      1     2.616
## 129326 80052834654 5.140 2011     2700      1     3.237
## 129501 80052462089 3.945 2011     2700      1     2.658
## 129589 80052173840 3.245 2011     2700      1     2.794
## 129619 80052436335 4.001 2011     2700      1     2.511
## 129626 80052475830 2.999 2011     2700      1     2.600
## 129667 80052790486 2.918 2011     2700      1     2.519
## 130507 80051720194 5.682 2011     2700      1     3.779
## 130603 84860389821 4.428 2011     2700      1     2.938
## 130622 80051723862 4.452 2011     2700      1     2.549
## 130629 84867338003 4.503 2011     2700      1     2.601
## 130688 80051571901 4.590 2011     2700      1     2.739
## 130690 80051775476 5.397 2011     2700      1     3.494
## 130779 79961108629 6.254 2011     2700      1     4.403
## 130847 79960629891 3.029 2011     2700      1     2.630
## 131346 79960855656 4.776 2011     2700      1     2.873
## 131363 79960704300 3.825 2011     2700      1     2.539
## 131435 79960575079 4.596 2011     2700      1     2.693
## 131505 79960169424 3.085 2011     2700      1     2.686
## 131652 79953303630 3.072 2011     2700      1     2.621
## 131727 79960015327 4.590 2011     2700      1     2.739
## 131728 79960015997 5.714 2011     2700      1     3.861
## 131731 79960027556 4.472 2011     2700      1     2.621
## 131763 79960293663 3.125 2011     2700      1     2.674
## 132089 84857065457 3.387 2011     2700      1     2.936
```

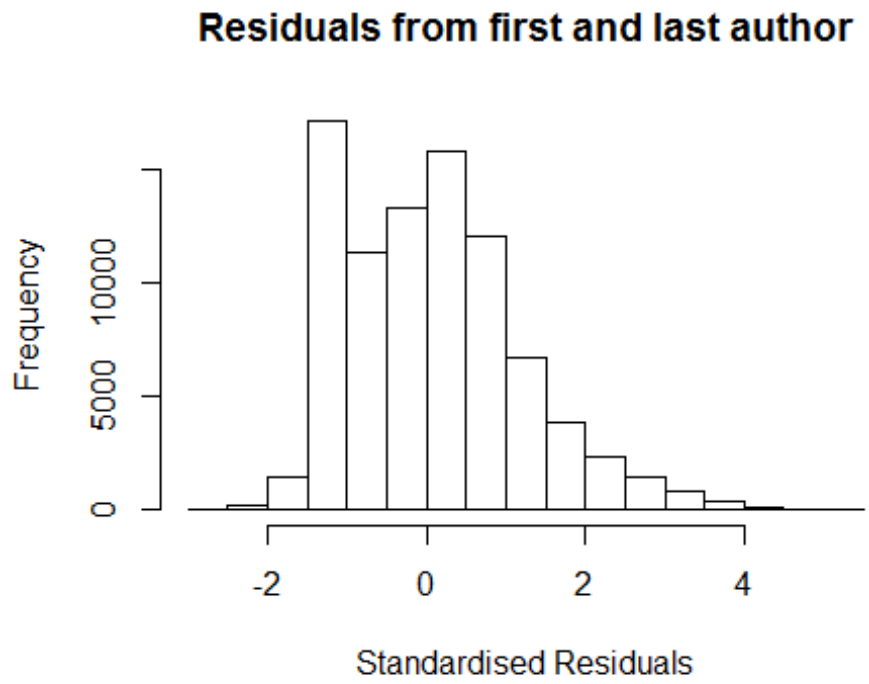
```
## 132437 79959794787 5.141 2011     2700      1     3.238
## 132474 79959772357 4.719 2011     2700      1     2.816
## 132530 79959478455 5.314 2011     2700      1     3.411
## 132531 79959478969 4.354 2011     2700      1     3.903
## 132546 79959487586 4.565 2011     2700      1     2.662
## 132615 79959297709 3.871 2011     2700      1     2.839
## 132662 79958724181 4.832 2011     2700      1     2.930
## 132736 79958195203 4.483 2011     2700      1     2.580
## 132738 79958226299 4.639 2011     2700      1     4.188
## 132745 79958066836 4.674 2011     2700      1     2.771
## 132790 79957795474 4.944 2011     2700      1     3.657
## 132796 79957880210 3.151 2011     2700      1     2.700
## 132905 79958782230 3.713 2011     2700      1     2.629
## 132933 79959798617 3.405 2011     2700      1     2.954
## 133601 79957444504 4.382 2011     2700      1     2.531
## 133602 79957445348 4.662 2011     2700      1     2.809
## 133604 79957497011 3.256 2011     2700      1     2.805
## 133644 79956204057 3.846 2011     2700      1     3.395
## 133690 79956218863 5.583 2011     2700      1     3.680
## 133759 79955909191 4.421 2011     2700      1     2.518
## 133897 79955624770 4.409 2011     2700      1     2.506
## 133939 79955517235 4.745 2011     2700      1     2.843
## 134759 79954561234 4.639 2011     2700      1     2.736
## 134775 79954487919 5.185 2011     2700      1     3.332
## 134778 79954553689 4.428 2011     2700      1     2.525
## 134883 79953748230 4.910 2011     2700      1     3.008
## 134968 79953307878 4.700 2011     2700      1     2.797
## 134979 79953331206 4.492 2011     2700      1     2.589
## 135559 79953225169 4.658 2011     2700      1     2.805
## 135629 79953232077 4.567 2011     2700      1     2.664
## 135637 79953118839 5.685 2011     2700      1     3.783
## 135705 79952795093 4.650 2011     2700      1     2.797
## 135762 79952376771 4.188 2011     2700      1     2.698
## 135793 79952476762 4.836 2011     2700      1     2.985
## 135911 79952363727 5.013 2011     2700      1     3.929
## 136036 79952570027 3.688 2011     2700      1     3.237
## 136140 79955701330 3.138 2011     2700      1     2.687
## 136594 79952026360 2.928 2011     2700      2     2.529
## 136941 79951961755 3.835 2011     2700      1     2.751
## 136973 79951850287 4.863 2011     2700      1     2.960
## 136988 79951821800 4.881 2011     2700      1     3.646
## 136990 79951939809 3.630 2011     2700      1     3.179
## 136997 79951699237 4.608 2011     2700      1     3.322
## 137056 79851493086 4.470 2011     2700      1     3.184
## 137064 79751522951 5.633 2011     2700      1     3.780
## 137776 79251525181 3.985 2011     2700      1     2.901
## 137800 78751610848 4.907 2011     2700      1     3.005
## 137805 79952148734 4.771 2011     2700      1     4.320
## 137860 78651378996 3.950 2011     2700      1     3.499
## 137861 78651379500 5.349 2011     2700      1     3.446
```

```
## 137865 78651393550 4.594 2011     2700      1     2.691
## 137883 78650913180 3.289 2011     2700      1     2.838
## 137909 78650949272 4.632 2011     2700      1     2.730
## 137995 78651245300 4.046 2011     2700      1     3.014
## 138009 79151481493 3.670 2011     2700      1     2.586
## 138037 79251527232 2.967 2011     2700      1     2.516
## 138816 83155182714 3.589 2011     2700      1     2.505
## 139857 84871293779 4.110 2012     2700      1     3.114
## 140050 84870887067 4.610 2012     2700      1     2.795
## 140140 84870549609 4.439 2012     2700      1     2.624
## 141706 84870012939 4.380 2012     2700      1     2.565
## 141913 84868697417 2.973 2012     2700      1     2.610
## 141917 84869069213 2.882 2012     1000      2     2.519
## 142148 84866788499 3.256 2012     2700      1     2.893
## 142464 84875771151 2.960 2012     2700      1     2.597
## 142976 84867828250 4.526 2012     2700      1     2.763
## 143055 84867584625 3.204 2012     2700      1     2.841
## 143075 84867506672 2.947 2012     2700      1     2.636
## 144167 84866419651 2.866 2012     2700      1     2.503
## 144331 84866098277 4.037 2012     2700      1     2.838
## 145628 84865294834 2.960 2012     2700      1     2.597
## 145641 84865320778 4.511 2012     2700      1     2.696
## 145642 84865323065 4.426 2012     2700      1     2.662
## 146664 84864242034 4.379 2012     2700      1     2.564
## 146785 84864008031 5.104 2012     2700      1     3.289
## 147015 84863524855 5.149 2012     2700      1     3.336
## 147039 84863527915 3.128 2012     2700      1     2.765
## 147118 84862840164 3.272 2012     2700      1     2.961
## 147811 84862859820 6.239 2012     2700      1     4.426
## 147813 84862868517 2.908 2012     2700      1     2.545
## 147814 84862903106 5.924 2012     2700      1     4.111
## 148146 84862260833 4.332 2012     2700      1     2.519
## 148149 84865580177 3.320 2012     2700      1     3.009
## 149011 84861394764 4.694 2012     2700      1     2.879
## 149012 84861409254 3.735 2012     2700      1     3.372
## 149013 84861409551 4.560 2012     2700      1     2.797
## 149127 84861022041 4.793 2012     2700      1     3.029
## 149215 84860914047 3.560 2012     2700      1     2.616
## 149238 84860741191 4.704 2012     2700      1     2.889
## 149259 84860776122 4.355 2012     2700      1     2.540
## 149389 84859481364 3.496 2012     2700      1     2.500
## 150155 84860136615 5.159 2012     2700      1     3.344
## 150503 84858665432 4.608 2012     2700      1     2.793
## 150508 84859544287 4.701 2012     2700      1     3.705
## 150550 84859092587 3.942 2012     2700      1     2.743
## 150563 84859718265 4.683 2012     2700      1     2.868
## 151344 84858638369 4.465 2012     2700      1     2.650
## 151347 84863337617 4.454 2012     2700      1     2.639
## 151408 84858331964 4.546 2012     1000      2     3.145
## 151434 84858265970 4.576 2012     2700      1     2.761
```

```
## 151641 84858019974 4.876 2012     2700      1     3.061
## 152008 84865286457 2.934 2012     2700      1     2.623
## 152292 84858199964 3.072 2012     1213      3     2.761
## 152564 84857427752 4.330 2012     2700      1     2.515
## 152684 84857065859 4.573 2012     2700      1     2.760
## 152856 84856433589 5.714 2012     2700      1     4.365
## 152866 84856487711 5.900 2012     2700      1     4.551
## 152908 84857643783 4.842 2012     2700      1     3.027
## 153390 84856245226 4.422 2012     2700      1     2.607
## 153392 84856249964 3.802 2012     2700      1     2.806
## 153404 84856133161 3.633 2012     2700      1     2.689
## 153473 84856159009 4.542 2012     2700      1     2.729
## 153532 84855843640 3.790 2012     2700      1     2.845
## 153533 84855854046 4.736 2012     2700      1     3.740
## 153672 84855459760 4.667 2012     2700      1     2.852
## 153731 84055193579 3.440 2012     2700      1     3.077
## 153732 84055199809 4.449 2012     2700      1     4.086
## 153733 84055199825 2.880 2012     2700      1     2.517
## 153976 84859259397 2.921 2012     2700      1     2.558
## 154252 84863718620 3.870 2012     2700      1     2.519
## 154329 84864861863 4.382 2012     2700      1     2.567
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
UniqueAuthors +
##     Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##     k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -3.4459 -0.4582 -0.0524  0.6117  5.2849
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.17540    0.03047   38.57  < 2e-16 ***
## FirstAuthorFemale1 -0.00144    0.00754   -0.19    0.849
## LastAuthorFemale1  -0.05059    0.00779   -6.49  8.5e-11 ***
## UniqueAuthors2      0.63301    0.00915   69.21  < 2e-16 ***
## UniqueAuthors3      0.83577    0.00959   87.12  < 2e-16 ***
## UniqueAuthors4      1.03864    0.01034  100.43  < 2e-16 ***
## UniqueAuthors5      1.45190    0.00915  158.72  < 2e-16 ***
## Year1997           -0.01772    0.04028   -0.44    0.660
## Year1998            0.01780    0.04040    0.44    0.660
## Year1999           -0.00955    0.04008   -0.24    0.812
## Year2000            0.81857    0.11327    7.23  5.0e-13 ***
## Year2001           -0.10060    0.04072   -2.47    0.014 *
## Year2002           -0.49631    0.03448  -14.39  < 2e-16 ***
## Year2003           -0.53281    0.03549  -15.01  < 2e-16 ***
## Year2004           -0.64497    0.03359  -19.20  < 2e-16 ***
```

```
## Year2005           -0.66269    0.03341  -19.84  < 2e-16 ***
## Year2006           -0.67502    0.03231  -20.89  < 2e-16 ***
## Year2007           -0.73569    0.03167  -23.23  < 2e-16 ***
## Year2008           -0.77714    0.03138  -24.77  < 2e-16 ***
## Year2009           -0.75868    0.03123  -24.29  < 2e-16 ***
## Year2010           -0.77734    0.03096  -25.11  < 2e-16 ***
## Year2011           -0.72407    0.03134  -23.10  < 2e-16 ***
## Year2012           -0.81262    0.03106  -26.17  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.756
## Multiple R-squared:  0.364,  Adjusted R-squared:  0.364
## Convergence in 20 IRWLS iterations
##
## Robustness weights:
##  189 observations
c(43,47,602,603,850,852,1168,1317,1325,1338,1354,1843,1852,1855,1862,2664,322
9,3346,3956,4113,4249,4288,4407,4503,5905,6313,6331,6455,6707,6709,7226,7524,
7959,8307,8398,8401,8608,9207,9278,9417,9419,9714,11130,11580,11692,11719,121
45,13188,13437,13455,13491,13492,14272,14454,14709,14903,15027,15323,15402,15
435,15442,17085,17513,17516,18069,18098,18109,18231,18674,18680,18711,19100,1
9285,19289,19294,19972,20513,21320,21845,22268,22486,22487,24250,25695,25702,
25951,27038,27265,27269,27702,27834,29201,29252,29261,30414,30470,30491,30962
,30986,32419,32801,32817,32861,33256,33665,33937,34003,34286,34555,36597,3661
9,36638,36640,36644,36647,36684,37312,37682,37769,37837,38417,38879,38880,389
16,39473,39503,39540,39602,40121,40143,40758,40759,40775,41221,41279,41319,41
932,41960,41977,42445,43610,45396,45430,45462,46021,47208,47382,48002,48016,4
8553,48572,49776,49833,50868,50941,53322,53875,53877,54047,54601,56296,58296,
58326,58998,59064,59454,59514,63959,65598,68714,73112,73247,73749,74155,74258
,74286,74793,75845,75956,76557,76587,77018,82361,82363,83771,85041,85049,8540
3,85489)
##   are outliers with |weight| <= 6.6e-07 ( < 1.2e-06);
##  5374 weights are ~= 1. The remaining 81095 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.000   0.833   0.954   0.875   0.979   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.15e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
```

```
##                    "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##                     GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.282  1           1.132
## LastAuthorFemale  1.277  1           1.130
## Year              1.008 16           1.000
```




```
## [1] "List of  2719 outliers with residuals above  2.5"
##           ScopusId  NLCS Year OneField Fields residuals
## 28      0029855881 4.801 1996     2700      1     2.881
## 30      0029906954 4.525 1996     2700      1     2.605
## 31      0141469033 5.783 1996     2700      1     3.863
## 47      0030453430 5.115 1996     2700      1     3.195
## 79     10544241545 4.738 1996     2700      1     2.719
## 87     10544235697 5.572 1996     2700      1     3.652
## 95      0029911547 5.746 1996     2700      1     4.031
## 113     0030445619 5.158 1996     2700      1     3.139
## 153     0029683640 4.469 1996     2700      1     2.549
## 179     0029844930 4.941 1996     2700      1     3.021
## 197     0029979690 4.533 1996     2700      1     2.514
## 1288    0029827650 4.623 1996     2700      1     2.703
## 1305    0030599533 4.661 1996     2700      1     2.642
## 1309    0029828983 4.411 1996     2700      1     2.596
## 1355    0030590745 4.426 1996     2700      1     2.506
## 1359    0029808003 5.038 1996     2700      1     3.019
```

```
## 1369    0029845013 4.799 1996     2700      1     2.879
## 1394    0029859887 4.512 1996     2700      1     2.797
## 1397    0029919459 5.204 1996     2700      1     3.185
## 1399    3042828081 4.528 1996     2700      1     2.509
## 1414   10544220023 4.705 1996     2700      1     2.785
## 1424    0029907556 5.146 1996     2700      1     3.331
## 1425    0029908290 4.742 1996     2700      1     2.822
## 1649    0029849462 4.874 1996     2700      1     2.954
## 1672    0030603334 4.697 1996     2700      1     2.678
## 1677    0029862114 4.329 1996     2700      1     2.614
## 1708    0029823520 4.701 1996     2700      1     2.781
## 1717   10144258656 4.587 1996     2700      1     2.667
## 1725    0029861097 4.833 1996     2700      1     3.118
## 1728    0029861579 4.669 1996     2700      1     2.650
## 1755    0030581586 5.042 1996     2700      1     3.327
## 1756    0030581587 4.964 1996     2700      1     2.945
## 1757    0030581590 4.755 1996     2700      1     2.736
## 1758    0242409954 4.538 1996     2700      1     2.618
## 1763    0005150171 4.797 1996     2700      1     2.877
## 1764    0007410805 5.440 1996     2700      1     3.520
## 1766   10144244674 5.218 1996     2700      1     3.298
## 1772    0029758899 4.682 1996     2700      1     2.867
## 1773    0029759549 4.361 1996     2700      1     2.646
## 1787    0010496839 5.103 1996     2700      1     3.183
## 1788    0029795224 5.088 1996     2700      1     3.069
## 1798    0029773862 4.819 1996     2700      1     2.899
## 1958   10144241022 5.066 1996     2700      1     3.047
## 1977    0029798819 4.729 1996     1300      2     2.809
## 1979    0029842830 5.287 1996     1300      2     3.367
## 2023    0030604909 4.869 1996     2700      1     2.850
## 2032    0029790755 4.602 1996     2700      1     2.583
## 2033    0029840653 5.036 1996     2700      1     3.116
## 2070    9544253883 4.485 1996     2700      1     2.770
## 2071    0029740950 5.202 1996     2700      1     3.282
## 2072    0029741921 5.309 1996     2700      1     3.389
## 2074    9544228424 5.880 1996     2700      1     3.960
## 2107    0029835263 5.098 1996     2700      1     3.079
## 2108    0029835392 5.259 1996     2700      1     3.339
## 2109    0029838499 4.691 1996     2700      1     2.672
## 2110    0242435873 5.375 1996     2700      1     3.660
## 2112    9544222721 4.784 1996     2700      1     3.069
## 2113    0029739543 4.452 1996     2700      1     2.532
## 2115    0029771618 4.842 1996     2700      1     3.027
## 2130    0029818344 4.305 1996     2700      1     2.590
## 2131    0029822091 4.840 1996     2700      1     2.920
## 2159    0029737324 4.674 1996     2700      1     2.754
## 2162    0029739970 5.663 1996     2700      1     3.743
## 2167    0029758146 4.466 1996     2700      1     2.546
## 2354    0030250166 4.259 1996     2700      1     2.544
## 2485    0029741822 5.190 1996     2700      1     3.171
```

```
## 2486    0029746491 4.535 1996     2700      1     2.615
## 2489    9444268678 5.105 1996     2700      1     3.185
## 2490    0029740582 4.658 1996     2700      1     2.738
## 2500    0030017720 4.990 1996     2700      1     3.070
## 2503    0029783235 4.667 1996     2700      1     2.747
## 2510    0029739053 4.836 1996     2700      1     2.916
## 2513    0029759025 5.617 1996     2700      1     3.598
## 2526    0029830177 4.802 1996     2700      1     2.987
## 2537    0029737310 4.288 1996     2700      1     2.573
## 2543   15844420661 5.335 1996     2700      1     3.415
## 2574    0029830551 4.585 1996     2700      1     2.665
## 2582    0029862164 4.517 1996     2700      1     2.702
## 2588    0029916051 4.499 1996     2700      1     2.684
## 2741    0030594830 5.438 1996     2700      1     3.518
## 2743    0030000230 4.826 1996     2700      1     2.906
## 2744    8944228913 5.408 1996     2700      1     3.389
## 2748    0030056083 5.607 1996     2700      1     3.687
## 2750    0030057106 5.671 1996     2700      1     3.751
## 2797    8944233864 5.465 1996     2700      1     3.545
## 2806    0029902170 4.900 1996     2700      1     3.085
## 2830    0030015522 4.848 1996     2700      1     2.928
## 3004    0029944290 4.689 1996     1300      2     2.769
## 3042    0029786164 4.533 1996     2700      1     2.613
## 3058    0029900294 5.422 1996     2700      1     3.502
## 3059    0029942003 5.780 1996     2700      1     3.860
## 3085    0030594604 4.947 1996     2700      1     3.027
## 3109    0029948212 5.514 1996     2700      1     3.594
## 3111    0029948967 4.647 1996     2700      1     2.727
## 3154    9344267138 4.432 1996     2700      1     2.717
## 3166   15844364006 4.499 1996     2700      1     2.579
## 3185    0029888972 4.889 1996     2700      1     3.174
## 3214    0030000088 4.916 1996     2700      1     2.996
## 3462    0029892519 4.994 1996     2700      1     3.074
## 3468    0029939573 4.570 1996     2700      1     2.650
## 3471    0029878679 4.693 1996     2700      1     2.773
## 3474    0029897909 5.220 1996     2700      1     3.201
## 3479    0029886630 5.409 1996     2700      1     3.489
## 3487    0029889718 4.606 1996     2700      1     2.686
## 3498    0029870954 4.842 1996     2700      1     2.922
## 3503    0029932125 4.845 1996     2700      1     2.826
## 3521    0029925567 5.575 1996     2700      1     3.655
## 3537    0029871858 4.847 1996     2700      1     2.828
## 3539    0030008301 4.452 1996     2700      1     2.532
## 3552    9244262406 6.270 1996     2700      1     4.350
## 3560    0029863684 5.262 1996     2700      1     3.342
## 3572    0029913381 4.769 1996     2700      1     2.750
## 3573    0029913643 4.471 1996     2700      1     2.551
## 3578    0029929617 5.256 1996     2700      1     3.541
## 3579    0029929618 4.964 1996     2700      1     3.249
## 3795    0029875770 5.139 1996     2700      1     3.219
```

```
## 3797    0029982081 4.757 1996     2700      1     2.837
## 3806    0029988316 4.987 1996     2700      1     3.067
## 3830    0029920903 4.358 1996     2700      1     2.643
## 3866    0029881352 4.449 1996     2700      1     2.529
## 3877    0029869817 4.507 1996     2700      1     2.792
## 3890    0029993565 5.060 1996     2700      1     3.140
## 3892   13344285352 4.993 1996     2700      1     2.974
## 4094    0029915834 4.443 1996     2700      1     2.523
## 4105    0029970342 5.515 1996     2700      1     3.595
## 4111    0029917496 5.333 1996     2700      1     3.314
## 4150    0029993531 4.882 1996     2700      1     2.962
## 4157    0030009316 4.530 1996     2700      1     2.610
## 4168    0029873844 4.915 1996     2700      1     2.896
## 4169    0029987128 4.757 1996     2700      1     2.738
## 4171    0029870311 4.393 1996     2700      1     2.678
## 4178    9044236527 5.250 1996     2700      1     3.535
## 4195    0030070405 4.558 1996     2700      1     2.743
## 4217    0029866721 4.724 1996     2700      1     2.705
## 4509    0942276272 5.608 1996     2700      1     3.688
## 4519   13344293705 4.423 1996     2700      1     2.608
## 4535    0030066467 5.150 1996     2700      1     3.230
## 4536    0030070793 4.738 1996     2700      1     3.023
## 4548    0030020590 4.565 1996     1300      2     2.546
## 4552    0030020630 4.999 1996     2700      1     3.079
## 4567    0030058515 5.136 1996     2700      1     3.216
## 4591    0030042790 5.170 1996     2700      1     3.250
## 4593   13344286314 4.575 1996     2700      1     2.655
## 4596    0030068231 4.602 1996     2700      1     2.682
## 4598    0030058666 4.263 1996     2700      1     2.548
## 4603    0030033367 4.824 1996     2700      1     2.904
## 4622    0030032444 5.228 1996     2700      1     3.308
## 4629    0030049026 4.570 1996     2700      1     2.650
## 4632    0030056373 4.773 1996     2700      1     2.853
## 4784    0030032258 4.399 1996     2700      1     2.684
## 4802    0030044553 4.676 1996     2700      1     2.961
## 4818    0030049823 4.429 1996     2700      1     2.614
## 4819   13344260688 4.602 1996     2700      1     2.682
## 4826    0030048691 4.811 1996     2700      1     3.096
## 4827    0030060343 4.437 1996     2700      1     2.517
## 4833    0030026954 4.778 1996     2700      1     2.963
## 4843    0030032378 4.408 1996     2700      1     2.693
## 4866    0030043663 4.884 1996     2700      1     2.964
## 4868    0030579588 4.838 1996     2700      1     2.918
## 4869    0030034463 5.009 1996     2700      1     2.990
## 4883    0030034592 5.638 1996     2700      1     3.718
## 4884    0030054309 6.200 1996     2700      1     4.280
## 4890    0030024585 4.558 1996     2700      1     2.638
## 4891    0030046074 4.501 1996     2700      1     2.786
## 4897    0000677401 5.404 1996     2700      1     3.589
## 4904    0001196648 5.048 1996     2700      1     3.128
```

```
## 4908    0001854685 4.563 1996     2700      1     2.748
## 4916    0004851872 6.351 1996     2700      1     4.636
## 5482    0029959975 4.583 1996     2700      1     2.564
## 5653    0030034465 5.226 1996     2700      1     3.306
## 5706    0030054203 4.894 1996     2700      1     2.974
## 5729    0030060710 4.744 1996     2700      1     2.725
## 5737    0030061707 4.711 1996     2700      1     2.791
## 5847    0030183205 4.634 1996     2700      1     2.819
## 6074    0030576183 4.548 1996     2700      1     2.628
## 6086    0040419081 4.738 1996     2700      1     3.023
## 6118   16144365807 4.546 1996     2700      1     2.527
## 6164    9244238642 4.854 1996     2700      1     2.934
## 7329    0031435838 4.794 1997     2700      1     3.086
## 7330    2642611953 4.758 1997     2700      1     2.951
## 7352    0031468254 4.620 1997     2700      1     2.708
## 7360    0031472414 5.042 1997     2700      1     3.130
## 7367    0031437624 4.545 1997     2700      1     2.633
## 7375    0031466866 4.796 1997     2700      1     2.884
## 7379    2642597076 5.193 1997     2700      1     3.281
## 7415    0030704598 4.531 1997     2700      1     2.619
## 7534    0031136510 4.780 1997     2700      1     2.868
## 7757    0031454125 4.470 1997     2700      1     2.558
## 7787    0031472452 4.431 1997     2700      1     2.723
## 8304    0030703240 4.771 1997     2700      1     2.759
## 8307    0030665744 4.564 1997     2700      1     2.652
## 8323    0031590631 4.483 1997     2700      1     2.571
## 8332    0030692782 5.405 1997     2700      1     3.493
## 8339    0030666228 4.564 1997     2700      1     2.757
## 8340    0030671045 4.349 1997     2700      1     2.641
## 8357    0030809817 4.318 1997     2700      1     2.610
## 8378    0030695138 5.236 1997     2700      1     3.324
## 8424    0030716498 5.322 1997     2700      1     3.410
## 8442    0030661687 5.440 1997     2700      1     3.528
## 8455    0030671503 4.307 1997     2700      1     2.599
## 8586    0030731486 5.178 1997     2700      1     3.371
## 8589    0030831628 4.379 1997     2700      1     2.671
## 8608    0030720886 5.370 1997     2700      1     3.662
## 8610    0030759072 5.217 1997     2700      1     3.305
## 8625    0030744945 5.294 1997     2700      1     3.382
## 8627    0030762081 4.480 1997     2700      1     2.568
## 8631    0030879325 4.935 1997     2700      1     3.023
## 8692    0030766770 4.941 1997     2700      1     3.233
## 8713    0030779037 5.224 1997     2700      1     3.212
## 8714    0030803131 4.352 1997     2700      1     2.644
## 8728    0030774499 4.679 1997     2700      1     2.872
## 8849    0030886175 4.489 1997     2700      1     2.577
## 8850    0030954873 5.207 1997     2700      1     3.499
## 8851    0030967953 6.065 1997     2700      1     4.153
## 8861    0006750594 4.483 1997     2700      1     2.571
## 8891    0030756517 4.679 1997     2700      1     2.767
```

```
## 8980    0030869269 5.624 1997     2700      1     3.712
## 8983    9844257578 4.711 1997     2700      1     3.003
## 9007    0030763532 5.703 1997     2700      1     3.691
## 9014    0030928107 4.550 1997     2700      1     2.842
## 9045    0030921893 4.869 1997     2700      1     2.957
## 9237    0030803718 4.703 1997     2700      1     2.896
## 9239    0030868293 5.188 1997     2700      1     3.276
## 9245    0030804947 4.885 1997     2700      1     3.177
## 9250    0030792120 4.836 1997     2700      1     2.924
## 9309    8544284052 4.624 1997     2700      1     2.712
## 9329    0030851442 4.840 1997     2700      1     2.928
## 9331    0030872063 4.662 1997     2700      1     2.750
## 9332    1842332651 4.618 1997     2700      1     2.706
## 9378    0030854951 4.445 1997     2700      1     2.533
## 9468    8544252402 4.498 1997     2700      1     2.790
## 9514    0030744599 4.744 1997     2700      1     2.832
## 9515    0030756101 4.907 1997     2700      1     3.100
## 9519    0030854780 4.755 1997     2700      1     2.843
## 9522    0030742587 4.609 1997     2700      1     2.901
## 9539    0030876805 5.124 1997     2700      1     3.212
## 9568    0030803395 4.639 1997     2700      1     2.931
## 9597    0030610461 5.331 1997     2700      1     3.319
## 9603    0038025895 4.877 1997     2700      1     2.965
## 9607    0030957310 5.247 1997     2700      1     3.335
## 9629    0030760729 4.653 1997     2700      1     2.741
## 9652    0030878260 4.429 1997     2700      1     2.517
## 9760    0031005933 5.176 1997     2700      1     3.264
## 9763    1842410169 4.665 1997     2700      1     2.653
## 9767    0030916401 5.117 1997     2700      1     3.205
## 9816    0030908055 5.171 1997     2700      1     3.259
## 9818    0030941817 4.703 1997     2700      1     2.691
## 9820    0030976173 4.495 1997     2700      1     2.787
## 9823    0031009275 4.344 1997     2700      1     2.537
## 9824    8244234470 4.344 1997     2700      1     2.636
## 9833    0343157354 5.144 1997     2700      1     3.232
## 9902    0030974188 4.972 1997     2700      1     3.060
## 10055   0030909527 4.552 1997     2700      1     2.745
## 10056   0030910022 5.719 1997     2700      1     4.011
## 10060   0030939253 4.627 1997     2700      1     2.715
## 10066   0031003334 5.309 1997     2700      1     3.397
## 10071   0343918505 4.869 1997     2700      1     3.161
## 10072   0346593922 4.853 1997     2700      1     2.941
## 10098   0031009871 4.652 1997     2700      1     2.845
## 10102   0030919511 5.119 1997     2700      1     3.207
## 10103   0030919667 4.467 1997     2700      1     2.555
## 10127   0030977450 4.957 1997     2700      1     3.045
## 10149   0030913316 4.821 1997     2700      1     2.909
## 10156   0030961921 5.444 1997     2700      1     3.736
## 10184   0030896520 5.060 1997     2700      1     3.048
## 10222   0030979720 4.965 1997     2700      1     3.053
```

```
## 10347   0030967165 5.834 1997     2700      1     3.822
## 10353   0030955080 4.722 1997     2700      1     3.014
## 10378   0012444519 6.220 1997     2700      1     4.308
## 10386   0030903440 4.583 1997     2700      1     2.875
## 10427   0030896996 4.738 1997     2700      1     3.030
## 10428   0030899940 5.040 1997     2700      1     3.028
## 10454   0030956673 6.413 1997     2700      1     4.501
## 10461   0030951198 5.067 1997     2700      1     3.155
## 10464   0031004940 4.520 1997     2700      1     2.812
## 10467   0030891988 4.695 1997     2700      1     2.683
## 10490   0030946366 4.576 1997     2700      1     2.564
## 10535   0031127072 5.869 1997     2700      1     4.062
## 10598   0030976067 5.207 1997     2700      1     3.195
## 10622   0030898397 4.415 1997     2700      1     2.503
## 10623   0030948376 4.687 1997     2700      1     2.775
## 10646   0030945186 4.445 1997     2700      1     2.533
## 10660   0031047918 4.572 1997     2700      1     2.560
## 10671   0031054674 5.776 1997     2700      1     3.864
## 10694   0031047361 4.710 1997     2700      1     2.798
## 10735   0030948337 4.434 1997     2700      1     2.627
## 10891   0013655172 4.847 1997     2700      1     2.935
## 10895   0031052369 4.993 1997     2700      1     3.081
## 10901   0031018269 4.695 1997     2700      1     2.783
## 10919   0031045652 5.274 1997     2700      1     3.362
## 10920   0031050740 4.653 1997     2700      1     2.741
## 10933   0031052862 5.331 1997     2700      1     3.319
## 10935   0031057614 4.810 1997     2700      1     2.798
## 10972   0031058448 4.616 1997     2700      1     2.809
## 10979   0031019745 4.552 1997     1000      2     2.640
## 10981   0031019923 4.600 1997     2700      1     2.688
## 10983   0031028712 4.450 1997     2700      1     2.538
## 10987   0031028804 5.501 1997     2700      1     3.589
## 10993   0030614495 4.450 1997     2700      1     2.538
## 11012   0030807358 4.886 1997     2700      1     2.874
## 11019   0030868077 4.859 1997     2700      1     2.947
## 11136   0031012726 5.498 1997     2700      1     3.586
## 11137   0031030450 4.891 1997     2700      1     2.979
## 11138   0031032055 5.515 1997     2700      1     3.603
## 11153   0031030620 5.167 1997     2700      1     3.255
## 11162   0031012532 4.615 1997     2700      1     2.703
## 11174   0031028365 4.604 1997     2700      1     2.592
## 11175   0031029379 4.969 1997     2700      1     2.957
## 11202   0031025101 4.552 1997     2700      1     2.844
## 11213   0031037065 4.605 1997     2700      1     2.593
## 11823   0030938817 4.429 1997     2700      1     2.721
## 11955   0031011193 4.851 1997     2700      1     2.939
## 11971   0031020373 4.539 1997     2700      1     2.527
## 11972   0031022490 4.618 1997     2700      1     2.606
## 13373   0031472453 4.648 1997     2700      1     2.736
## 13487   1842295778 5.282 1997     2700      1     3.370
```

```
## 13948   0032585514 4.968 1998     2700      1     3.012
## 13961   0032583492 5.604 1998     2700      1     3.852
## 13975   0032564658 4.473 1998     2700      1     2.517
## 13977   0032564703 4.515 1998     2700      1     2.763
## 13994   0032542385 4.640 1998     2700      1     2.684
## 14001   0032539164 4.551 1998     2700      1     2.700
## 14018   0032506614 5.304 1998     2700      1     3.348
## 14032   0032480930 4.637 1998     2700      1     2.681
## 14040   0032477294 5.316 1998     2700      1     3.564
## 14041   0032477305 5.311 1998     2700      1     3.559
## 14043   0032477331 5.005 1998     2700      1     3.154
## 14065   0031609707 4.571 1998     2700      1     2.615
## 14801   0032567115 5.271 1998     2700      1     3.315
## 14816   0032547938 6.002 1998     2700      1     4.250
## 14828   0032544946 4.306 1998     2700      1     2.554
## 14860   0032512097 4.594 1998     2700      1     2.638
## 14866   0032508952 6.390 1998     2700      1     4.434
## 14881   0032487923 5.141 1998     2700      1     3.185
## 14888   0032483685 5.248 1998     2700      1     3.292
## 14889   0032483690 4.752 1998     2700      1     2.796
## 15041   0032578789 4.449 1998     2700      1     2.697
## 15080   0032556180 5.427 1998     2700      1     3.675
## 15109   0032517251 4.717 1998     2700      1     2.965
## 15110   0032517258 4.889 1998     2700      1     3.137
## 15130   0032494396 4.981 1998     2700      1     2.925
## 15186   0032189897 5.522 1998     2700      1     3.566
## 15322   0032563824 4.832 1998     2700      1     2.876
## 15354   0032541616 5.053 1998     2700      1     3.097
## 15357   0032541672 4.883 1998     2700      1     2.927
## 15362   0032538047 4.487 1998     2700      1     2.531
## 15396   0032504985 4.662 1998     2700      1     2.706
## 15418   0032480346 4.888 1998     2700      1     2.932
## 15530   0032169485 5.282 1998     2700      1     3.326
## 15532   0032169557 4.561 1998     2700      1     2.605
## 15630   0032572929 4.615 1998     2700      1     2.659
## 15666   0032547326 6.369 1998     2700      1     4.413
## 15667   0032547328 4.489 1998     2700      1     2.533
## 15690   0032529103 5.002 1998     2700      1     3.046
## 15700   0008926519 5.466 1998     2700      1     3.510
## 15741   0032486734 4.466 1998     2700      1     2.714
## 15784   0032145836 4.610 1998     2700      1     2.554
## 15882   0032578110 4.511 1998     2700      1     2.660
## 15906  18744421762 4.597 1998     2700      1     2.641
## 15932   0032528169 4.517 1998     2700      1     2.666
## 15935   0032528180 5.239 1998     2700      1     3.283
## 15947   0032508297 5.453 1998     2700      1     3.497
## 15961   0032496880 4.515 1998     2700      1     2.559
## 15971   0032474694 4.931 1998     2700      1     2.975
## 16151   0032551162 4.753 1998     2700      1     2.797
## 16167   0032543663 5.069 1998     2700      1     3.113
```

```
## 16172   7144228604 4.850 1998     2700      1     2.894
## 16174   0009452157 4.306 1998     2700      1     2.554
## 16193   0032525901 4.517 1998     2700      1     2.765
## 16222   0032503637 4.575 1998     2700      1     2.823
## 16251   0032482329 5.835 1998     2700      1     4.083
## 16263   0032479034 4.911 1998     2700      1     2.855
## 16346   0032102292 5.131 1998     2700      1     3.175
## 16483   0032572043 5.088 1998     2700      1     3.336
## 16487   0032572086 6.274 1998     2700      1     4.318
## 16504   0032550626 5.681 1998     2700      1     3.725
## 16505   0032550630 5.027 1998     2700      1     3.275
## 16536   0032516267 4.935 1998     2700      1     3.183
## 16539   0032516296 5.565 1998     2700      1     3.609
## 16568   0032490147 4.494 1998     2700      1     2.538
## 16773   0032580484 4.324 1998     2700      1     2.572
## 16793   0032560037 4.393 1998     2700      1     2.542
## 16800   0032557174 4.358 1998     2700      1     2.507
## 16801   0032557175 4.491 1998     2700      1     2.535
## 16802   2642713359 4.651 1998     2700      1     2.695
## 16816   0032537068 4.638 1998     2700      1     2.682
## 16819   0032537191 4.969 1998     2700      1     3.118
## 16832   0032522736 4.715 1998     2700      1     2.659
## 16839   0032523018 4.548 1998     2700      1     2.592
## 16851   0032523214 4.715 1998     2700      1     2.659
## 16875   0032499126 4.411 1998     2700      1     2.659
## 16876   0032499141 4.532 1998     2700      1     2.576
## 16910   0008947293 4.653 1998     2700      1     2.697
## 16913   0032473922 4.668 1998     2700      1     2.712
## 17047   0032568257 4.810 1998     2700      1     2.854
## 17053   0032565112 5.107 1998     2700      1     3.151
## 17063   0032554552 4.707 1998     2700      1     2.955
## 17064   0032554569 4.610 1998     2700      1     2.554
## 17091   0032542709 5.010 1998     2700      1     3.054
## 17163   0032485350 5.150 1998     2700      1     3.194
## 17177   0032481673 4.592 1998     2700      1     2.536
## 17289   0032032575 4.574 1998     2700      1     2.618
## 17293   0032033112 4.403 1998     2700      1     2.651
## 17386   0032564902 5.256 1998     2700      1     3.200
## 17401   0032545939 4.480 1998     2700      1     2.524
## 17404   0032546036 4.281 1998     2700      1     2.529
## 17408   0032542516 4.484 1998     2700      1     2.528
## 17410   0032542541 4.339 1998     2700      1     2.587
## 17420   0032519431 4.715 1998     2700      1     2.659
## 17429   0032519925 5.694 1998     2700      1     3.638
## 17433   0032520153 4.772 1998     2700      1     2.816
## 17444   0032510076 5.618 1998     2700      1     3.562
## 17473   0032484953 4.572 1998     2700      1     2.516
## 17481   0032481152 4.867 1998     2700      1     2.911
## 17482   0032481323 4.638 1998     2700      1     2.682
## 17502   0031987322 4.598 1998     2700      1     2.642
```

```
## 17541   0032005791 4.628 1998     2700      1     2.672
## 17595   0032573850 4.468 1998     2700      1     2.512
## 17606   0032556992 4.604 1998     2700      1     2.648
## 17613   0032554096 4.586 1998     2700      1     2.630
## 17627   0032518022 6.088 1998     2700      1     4.132
## 17631   0032518290 5.115 1998     2700      1     3.159
## 17633   0032518420 4.457 1998     2700      1     2.501
## 17644   0032518878 4.711 1998     2700      1     2.755
## 17655   0032515386 4.679 1998     2700      1     2.623
## 17680   0032491837 4.969 1998     2700      1     3.118
## 17686   0032477553 4.484 1998     2700      1     2.528
## 17696   0009440989 4.352 1998     2700      1     2.600
## 18344   0031945285 5.058 1998     2700      1     3.102
## 18436   0031975335 5.254 1998     2700      1     3.403
## 18437   0031975401 4.758 1998     2700      1     2.802
## 20189   0033619959 5.403 1999     2700      1     3.607
## 20216   0033582097 4.593 1999     2700      1     2.692
## 20242   0033572972 5.596 1999     2700      1     3.695
## 20262   0033540004 5.090 1999     2700      1     3.189
## 20278   0033518193 4.549 1999     2700      1     2.548
## 20279   0033518231 5.873 1999     2700      1     3.972
## 20319   0032803699 4.811 1999     2700      1     2.910
## 20484   0033485508 4.616 1999     2700      1     2.715
## 20874   0032748385 5.646 1999     2700      1     3.645
## 20925   0033544340 6.198 1999     2700      1     4.502
## 20962   0033520771 4.299 1999     2700      1     2.503
## 20967   0032589826 4.410 1999     2700      1     2.509
## 21115   0032694263 5.196 1999     2700      1     3.295
## 21120   0032695482 5.349 1999     2700      1     3.448
## 21123   0032741518 4.596 1999     2700      1     2.595
## 21124   0032742418 5.935 1999     2700      1     3.934
## 21125   0032742419 4.335 1999     2700      1     2.639
## 21126   0032743289 5.343 1999     2700      1     3.442
## 21127   0032747280 5.109 1999     2700      1     3.208
## 21134   0033598598 5.651 1999     2700      1     3.855
## 21152   0032696770 5.142 1999     2700      1     3.241
## 21155   0032740874 4.729 1999     2700      1     3.033
## 21255   0033517494 4.412 1999     2700      1     2.616
## 21273   0032855570 4.367 1999     2700      1     2.671
## 21436   0033615307 4.467 1999     2700      1     2.566
## 21443   0033575997 4.748 1999     2700      1     2.847
## 21498   0033536463 4.644 1999     2700      1     2.743
## 21500   0033536483 4.408 1999     2700      1     2.712
## 21501   0033536501 4.406 1999     2700      1     2.710
## 21548   0032841392 4.784 1999     2700      1     2.883
## 21656   0033199833 4.945 1999     2700      1     3.044
## 21763   0033603792 4.795 1999     2700      1     2.999
## 21775   0342961310 4.661 1999     2700      1     2.965
## 21785   0033584440 5.727 1999     2700      1     3.726
## 21826   0033546647 5.256 1999     2700      1     3.255
```

```
## 21990   0033595079 5.033 1999     2700      1     3.032
## 21993   0033595120 4.514 1999     2700      1     2.613
## 22017   0033565321 4.419 1999     2700      1     2.518
## 22021   0033565955 5.009 1999     2700      1     3.008
## 22042   0033542870 5.579 1999     2700      1     3.578
## 22168   0033168952 4.690 1999     2700      1     2.894
## 22237   0033600275 4.833 1999     2700      1     3.137
## 22248   0033031252 4.771 1999     2700      1     2.870
## 22286   0033542393 4.685 1999     2700      1     2.989
## 22303   0033538266 4.432 1999     2700      1     2.531
## 22305   0033538314 4.793 1999     2700      1     2.892
## 22486   0033609374 4.633 1999     2700      1     2.732
## 22491   0033606238 4.793 1999     2700      1     2.892
## 22521   0033562593 4.474 1999     2700      1     2.573
## 22535   0033551370 4.720 1999     2700      1     2.819
## 22563   0033526309 5.723 1999     2700      1     3.822
## 22759   0033611975 4.434 1999     2700      1     2.638
## 22779   0033594375 4.872 1999     2700      1     2.971
## 22793   0033591010 4.634 1999     2700      1     2.733
## 22832   0033553188 4.783 1999     2700      1     3.087
## 22834   0033553194 4.617 1999     2700      1     2.716
## 22836   0033553207 4.636 1999     2700      1     2.735
## 22852   0033531690 4.688 1999     2700      1     2.787
## 22855   0033515827 5.255 1999     1000      2     3.354
## 22909   0033119011 4.749 1999     2700      1     2.953
## 23018   0033577290 4.397 1999     2700      1     2.701
## 23019   0033577336 4.245 1999     2700      1     2.549
## 23056   0033545541 5.796 1999     2700      1     3.895
## 23057   0033545542 4.631 1999     2700      1     2.730
## 23064   0033540993 4.594 1999     2700      1     2.693
## 23067   0033541047 4.233 1999     2700      1     2.537
## 23076   0033522146 5.479 1999     2700      1     3.478
## 23077   0033522206 5.337 1999     2700      1     3.436
## 23111   0033037922 4.231 1999     2700      1     2.535
## 23167   0033104847 4.549 1999     2700      1     2.548
## 23271   0033602049 5.262 1999     2700      1     3.361
## 23292   0033580206 4.603 1999     2700      1     2.807
## 23293   0033580211 4.651 1999     2700      1     2.750
## 23304   0033577041 4.749 1999     2700      1     2.748
## 23314   0033557202 4.498 1999     2700      1     2.597
## 23357   0033540644 4.402 1999     2700      1     2.501
## 23358   0033540646 5.996 1999     2700      1     4.095
## 23359   0033540680 4.688 1999     2700      1     2.687
## 23360   0033540714 4.973 1999     2700      1     3.177
## 23361   0033540719 4.883 1999     2700      1     2.882
## 23541   0033608182 4.940 1999     2700      1     2.939
## 23544   0033608191 4.549 1999     2700      1     2.648
## 23560   0033590516 4.464 1999     2700      1     2.563
## 23561   0033590525 4.644 1999     2700      1     2.743
## 23567   0033585499 4.878 1999     2700      1     2.977
```

```
## 23577   0032918414 4.830 1999     2700      1     3.134
## 23584   0032954778 5.005 1999     2700      1     3.104
## 23591   0033552854 4.631 1999     2700      1     2.935
## 23605   0033537343 4.731 1999     2700      1     2.830
## 23621   0033528072 4.427 1999     2700      1     2.526
## 23623   0033528101 4.622 1999     2700      1     2.721
## 23860   0032701521 4.971 1999     2700      1     3.070
## 23861   0032701642 4.548 1999     2700      1     2.647
## 23882   0032713075 5.527 1999     2700      1     3.626
## 23914   0032722149 4.823 1999     2700      1     3.027
## 23925   0032725185 4.746 1999     2700      1     2.845
## 23974   0032742992 5.265 1999     2700      1     3.364
## 23997   0032746174 4.892 1999     2700      1     2.991
## 24023   0032752579 4.462 1999     2700      1     2.561
## 24083   0032824349 4.787 1999     2700      1     2.886
## 24251   0032919402 4.464 1999     2700      1     2.563
## 24424   0033013684 4.775 1999     2700      1     2.774
## 26283   0034055392 0.000 2000     2700      1    -2.554
## 26314  24044533275 0.000 2000     1306      3    -2.759
## 26347  85018888882 0.000 2000     2700      3    -2.554
## 26354   0034735827 5.241 2000     2700      1     2.787
## 26368   0034626988 5.215 2000     2700      1     2.556
## 26410   0034305633 0.000 2000     2700      2    -2.659
## 26509  33746530366 0.000 2000     2700      1    -2.659
## 26516   0033853362 0.000 2000     2700      1    -2.659
## 26526   0034226622 0.000 2000     1300      2    -2.659
## 26692   0034048842 0.000 2000     2700      1    -2.659
## 26889  85008036014 0.000 2000     1704      3    -2.659
## 26890  85048650487 0.000 2000     2700      1    -2.659
## 26910  12944252960 0.000 2000     2700      1    -2.659
## 26948  85047454626 0.000 2000     2700      1    -2.659
## 26964   0034732201 5.487 2000     2700      1     2.828
## 26988   0034655836 0.000 2000     2700      1    -2.659
## 27018   0034102286 0.000 2000     2700      1    -2.659
## 27031   0034682247 5.466 2000     2700      1     2.707
## 27110   0035956483 5.056 2001     2700      1     3.100
## 27142   0035924765 5.181 2001     2700      1     3.225
## 27205   0035857147 4.489 2001     2700      1     2.633
## 27210   0035852034 5.010 2001     2700      1     3.054
## 27211   0035852037 5.008 2001     2700      1     3.152
## 28065   0035944839 4.480 2001     2700      1     2.828
## 28077   0035936007 4.392 2001     2700      1     2.536
## 28092   0035930101 4.593 2001     2700      1     2.737
## 28118   0035904776 4.828 2001     2700      1     2.972
## 28120   0035892018 5.798 2001     2700      1     3.942
## 28145   0035861037 4.689 2001     2700      1     2.733
## 28177   0035829840 4.821 2001     2700      1     3.169
## 28179   0035829842 6.472 2001     2700      1     4.616
## 28353   0035522330 5.300 2001     2700      1     3.444
## 28468   0035950685 4.432 2001     2700      1     2.576
```

```
## 28564   0035846315 4.580 2001     2700      1     2.724
## 28628   0035802216 4.544 2001     2700      1     2.588
## 28791   0035960116 4.693 2001     2700      1     2.837
## 28832   0035913589 5.497 2001     2700      1     3.641
## 28851   0035856017 5.495 2001     2700      1     3.639
## 28900   0035818048 5.736 2001     2700      1     3.880
## 29126   0035975408 4.632 2001     2700      1     2.776
## 29159   0035940039 4.392 2001     2700      1     2.536
## 29160   0035940040 4.567 2001     2700      1     2.711
## 29182   0034898468 4.742 2001     2700      1     2.786
## 29248   0035827865 4.721 2001     2700      1     2.865
## 29267   0034926528 4.251 2001     2700      1     2.599
## 29299   0035413919 4.889 2001     2700      1     3.237
## 29405   0035963919 4.481 2001     2700      1     2.625
## 29417   0035954660 4.486 2001     2700      1     2.630
## 29418   0035954670 4.624 2001     2700      1     2.768
## 29428   0035948630 4.447 2001     2700      1     2.591
## 29439   0034939686 4.476 2001     2700      1     2.620
## 29458   0035928419 4.926 2001     2700      1     3.274
## 29472   0035913228 4.532 2001     2700      1     2.676
## 29523   0035849988 4.543 2001     2700      1     2.687
## 29524   0035850010 5.106 2001     2700      1     3.150
## 29529   0034933989 4.909 2001     2700      1     3.158
## 29532   0034945516 5.330 2001     2700      1     3.474
## 29705   0035963494 4.396 2001     2700      1     2.540
## 29708   0035963529 5.289 2001     2700      1     3.433
## 29762   0035919183 4.463 2001     2700      1     2.507
## 29763   0035919184 4.459 2001     2700      1     2.503
## 29766   0035897882 4.216 2001     2700      1     2.564
## 29767   0035897888 4.789 2001     2700      1     2.933
## 29776   0035859201 4.805 2001     2700      1     3.054
## 29778   0035859233 4.976 2001     2700      1     3.324
## 29787   0035854054 5.981 2001     2700      1     4.329
## 29826   0035816007 5.037 2001     2700      1     3.181
## 29832   0035816032 5.067 2001     2700      1     3.211
## 30038   0035902194 4.674 2001     2700      1     2.818
## 30046   0035897686 4.612 2001     2700      1     2.756
## 30047   0035897688 5.037 2001     2700      1     3.181
## 30084   0035832261 6.117 2001     2700      1     4.261
## 30089   0035832341 4.478 2001     2700      1     2.826
## 30856   0035953646 4.483 2001     2700      1     2.831
## 30872   0035946664 5.610 2001     2700      1     3.654
## 30947   0035857967 4.501 2001     2700      1     2.645
## 30954   0035849144 4.588 2001     2700      1     2.936
## 30967   0035843629 4.350 2001     2700      1     2.698
## 31007   0035804852 4.734 2001     2700      1     2.878
## 31008   0035804857 4.620 2001     2700      1     2.869
## 31151   0035962077 4.916 2001     2700      1     3.264
## 31177   0035925774 4.765 2001     2700      1     3.014
## 31204   0035869435 4.410 2001     2700      1     2.758
```

```
## 31220   0035857786 4.465 2001     2700      1     2.813
## 31237   0035835981 4.789 2001     2700      1     2.933
## 31257   0035826089 5.495 2001     2700      1     3.639
## 31264   0035819910 4.759 2001     2700      1     2.803
## 31434   0035961566 5.156 2001     2700      1     3.300
## 31461   0035931959 4.840 2001     2700      1     2.884
## 31463   0035931973 4.682 2001     2700      1     2.826
## 31477   0035925169 4.717 2001     2700      1     2.761
## 31485   0035901051 5.395 2001     2700      1     3.539
## 31491   0035864943 4.678 2001     2700      1     2.822
## 31492   0035864977 4.470 2001     2700      1     2.614
## 31541   0035825726 5.310 2001     2700      1     3.454
## 31569   0035819481 4.496 2001     2700      1     2.540
## 31715   0035945667 4.738 2001     2700      1     2.882
## 31734   0035915688 4.647 2001     2700      1     2.691
## 31742   0035905889 4.581 2001     2700      1     2.725
## 31744   0035905908 4.896 2001     2700      1     3.040
## 31773   0035843243 5.197 2001     2700      1     3.446
## 31775   0035843254 4.324 2001     2700      1     2.672
## 31782   0035835027 4.362 2001     2700      1     2.611
## 31783   0035835029 4.403 2001     2700      1     2.547
## 31795   0035814656 4.673 2001     2700      1     2.817
## 31799   0035804277 5.895 2001     2700      1     4.039
## 31801   0035804283 4.647 2001     2700      1     2.791
## 31933   0034795654 4.634 2001     2700      1     2.778
## 31937   0034799675 4.729 2001     2700      1     2.873
## 32022   0034857255 4.388 2001     2700      1     2.532
## 32059   0034903710 4.656 2001     2700      1     2.800
## 32248   0035013499 4.468 2001     2700      1     2.512
## 32609   0035174173 4.507 2001     2700      1     2.651
## 33892   0035960866 5.082 2001     2700      1     3.126
## 34423   0037132575 4.351 2002     2700      1     2.973
## 34424   0037132651 3.900 2002     2700      1     2.627
## 34438   0037079352 4.660 2002     2700      1     3.183
## 34445   0037069763 4.294 2002     2700      1     2.916
## 34446   0037069783 4.449 2002     2700      1     3.276
## 34454   0037065306 5.195 2002     2700      1     3.817
## 34457   0037065340 4.209 2002     2700      1     2.831
## 34467   0037028001 4.145 2002     2700      1     2.767
## 34474   0037021538 3.880 2002     2700      1     2.502
## 34477   0037021550 3.922 2002     2700      1     2.544
## 34553   0036908576 3.908 2002     2700      1     2.530
## 34596   0036950665 4.266 2002     2700      1     2.888
## 34958   0037202854 4.238 2002     2700      1     2.860
## 34973   0037191726 4.195 2002     2700      1     2.817
## 34977   0037184416 4.294 2002     2700      1     2.916
## 34979   0037184418 4.896 2002     2700      1     3.518
## 35001   0037153024 4.533 2002     2700      1     3.155
## 35004   0037153037 4.517 2002     2700      1     3.139
## 35005   0037153042 5.145 2002     2700      1     3.872
```

```
## 35013   0037145856 5.083 2002     2700      1     3.705
## 35014   0037145863 5.578 2002     2700      1     4.200
## 35032   0037078969 5.588 2002     2700      1     4.415
## 35039   0037073256 4.972 2002     2700      1     3.699
## 35043   0037073289 4.418 2002     2700      1     2.941
## 35079   0037032401 4.648 2002     2700      1     3.270
## 35130   0036840210 3.984 2002     2700      1     2.606
## 35358   0037206364 5.141 2002     2700      1     3.763
## 35359   0037206368 4.285 2002     2700      1     2.907
## 35370   0037180233 4.412 2002     2700      1     3.034
## 35377   0037167982 4.361 2002     2700      1     2.983
## 35385   0037164045 4.478 2002     2700      1     3.001
## 35386   0037164054 5.560 2002     2700      1     4.083
## 35397   0037126201 4.294 2002     2700      1     2.916
## 35398   0037126332 4.518 2002     2700      1     3.041
## 35400   0037126341 5.773 2002     2700      1     4.395
## 35408   0037120864 4.463 2002     2700      1     3.085
## 35411   0037120907 4.813 2002     2700      1     3.640
## 35446   0037048640 4.161 2002     2700      1     2.783
## 35448   0037048669 6.012 2002     2700      1     4.535
## 35451   0037048695 5.664 2002     2700      1     4.187
## 35452   0037048697 4.170 2002     2700      1     2.897
## 35457   0037027050 4.327 2002     2700      1     3.154
## 35462   0037015543 4.380 2002     2700      1     3.207
## 35481   0036771775 4.268 2002     2700      1     2.791
## 35699   0037179578 4.011 2002     2700      1     2.633
## 35709   0037174346 4.951 2002     2700      1     3.573
## 35730   0037130645 3.986 2002     2700      1     2.608
## 35758   0037063394 3.715 2002     2700      1     2.542
## 35759   0037063403 4.739 2002     2700      1     3.466
## 35760   0037063415 4.234 2002     2700      1     2.856
## 35778  18644385210 4.777 2002     2700      1     3.300
## 35867   0036733312 4.085 2002     2700      1     2.707
## 35911   0036745457 3.816 2002     2700      1     2.643
## 36123   0037190086 4.360 2002     2700      1     2.982
## 36181   0037158614 4.762 2002     2700      1     3.384
## 36244   0037125569 4.079 2002     2700      1     2.602
## 36254   0037103420 4.673 2002     2700      1     3.295
## 36269   0037077489 4.481 2002     2700      1     3.004
## 36270   0037077490 4.958 2002     2700      1     3.785
## 36292   0037043652 4.862 2002     2700      1     3.385
## 36295   0037043682 4.263 2002     2700      1     2.990
## 36306   0037036787 4.703 2002     2700      1     3.325
## 36318   0037014941 4.189 2002     2700      1     2.712
## 36365   0036681988 5.232 2002     2700      1     3.854
## 36367   0036682163 4.454 2002     2700      1     3.076
## 36369   0036682267 4.857 2002     2700      1     3.479
## 36524   0037178579 4.374 2002     2700      1     2.897
## 36554   0037166985 4.791 2002     2700      1     3.413
## 36555   0037166986 4.893 2002     2700      1     3.515
```

```
## 36583   0037130113 4.121 2002     2700      1     2.848
## 36586   0037130183 4.068 2002     2700      1     2.690
## 36593   0037125379 6.631 2002     2700      1     5.458
## 36594   0037125417 3.979 2002     2700      1     2.601
## 36597   0037125454 4.380 2002     2700      1     3.207
## 36619   0037055011 4.224 2002     2700      1     2.846
## 36620   0037055016 4.251 2002     2700      1     2.873
## 36658   0037019322 4.383 2002     2700      1     3.005
## 36666   0037014584 5.135 2002     2700      1     3.862
## 36667   0037014630 4.589 2002     2700      1     3.211
## 36668   0037014644 4.131 2002     2700      1     2.753
## 36692   0035983642 4.125 2002     2700      1     2.747
## 36821   0037182768 4.258 2002     2700      1     2.985
## 36860   0037134824 4.216 2002     2700      1     2.739
## 36864   0037134872 4.163 2002     2700      1     2.686
## 36865   0037134927 3.726 2002     2700      1     2.553
## 36867   0037134933 3.807 2002     2700      1     2.534
## 36901   0037071793 4.430 2002     2700      1     3.157
## 36904   0037071808 5.799 2002     2700      1     4.421
## 36910   0037067128 4.048 2002     2700      1     2.670
## 36912   0037067150 4.184 2002     2700      1     2.806
## 36940   0037030694 3.959 2002     2700      1     2.581
## 36956   0037024313 4.125 2002     2700      1     2.952
## 36996   0036595171 4.043 2002     2700      1     2.566
## 37144   0037198420 4.789 2002     2700      1     3.411
## 37146   0037198432 5.119 2002     2700      1     3.946
## 37175   0036251741 4.327 2002     2700      1     3.154
## 37181   0037162115 4.877 2002     2700      1     3.499
## 37182   0037162116 5.260 2002     2700      1     3.882
## 37188   0037157095 3.984 2002     2700      1     2.606
## 37191   0037157169 3.999 2002     2700      1     2.621
## 37208   0037129633 4.406 2002     2700      1     3.028
## 37241   0037092912 4.464 2002     2700      1     3.086
## 37243   0037092914 4.102 2002     2700      1     2.724
## 37246   0037093012 4.590 2002     2700      1     3.417
## 37286   0037042231 4.892 2002     2700      1     3.514
## 37288   0037042250 3.905 2002     2700      1     2.632
## 37305   0037018763 3.994 2002     2700      1     2.616
## 37309   0037007679 4.195 2002     2700      1     2.817
## 37311   0037007683 4.588 2002     2700      1     3.111
## 37362   0036570099 4.672 2002     2700      1     3.195
## 37452   0037181697 3.871 2002     2700      1     2.698
## 37453   0037181727 4.139 2002     2700      1     2.662
## 37462   0037171703 4.324 2002     2700      1     2.946
## 37475   0037165667 4.466 2002     2700      1     2.989
## 37476   0037165703 4.734 2002     2700      1     3.356
## 37477   0037165713 4.125 2002     2700      1     2.747
## 37478   0037165728 4.888 2002     2700      1     3.510
## 37479   0037165740 5.429 2002     2700      1     4.156
## 37527   0037123168 4.123 2002     2700      1     2.646
```

```
## 37528   0037123206 4.022 2002     2700      1     2.849
## 37570   0037061912 5.703 2002     2700      1     4.325
## 37581   0037051930 4.197 2002     2700      1     2.819
## 37582   0037051975 4.690 2002     2700      1     3.517
## 37626   0037018496 4.615 2002     2700      1     3.342
## 37638   0037012465 4.786 2002     2700      1     3.408
## 37640   0037012476 4.433 2002     2700      1     3.055
## 37744   0037197039 3.880 2002     2700      1     2.502
## 37745   0037197044 4.518 2002     2700      1     3.140
## 37752   0037187907 5.013 2002     2700      1     3.635
## 37753   0037187926 4.094 2002     2700      1     2.716
## 37758   0037181496 4.070 2002     2700      1     2.692
## 37759   0037181511 4.841 2002     2700      1     3.463
## 37760   0037181515 4.728 2002     2700      1     3.350
## 37778   0036119735 4.484 2002     2700      1     3.106
## 37782   0036122659 4.001 2002     2700      1     2.524
## 37790   0037160936 4.064 2002     2700      1     2.686
## 37819   0037149714 4.285 2002     2700      1     2.907
## 37826   0037139438 4.401 2002     2700      1     3.023
## 37868   0037076025 5.024 2002     2700      1     3.751
## 37875   0037070525 4.224 2002     2700      1     2.846
## 37876   0037070528 4.231 2002     2700      1     3.058
## 37900   0037035121 4.647 2002     2700      1     3.374
## 37913   0037029069 3.902 2002     2700      1     2.524
## 38083   0037181148 4.810 2002     2700      1     3.432
## 38109   0037148928 4.712 2002     2700      1     3.235
## 38143   0037116832 5.558 2002     2700      1     4.180
## 38145   0037116836 4.256 2002     2700      1     2.779
## 38161   0037075257 5.488 2002     2700      1     4.011
## 38163   0037075268 3.927 2002     2700      1     2.654
## 38164   0037075269 3.951 2002     2700      1     2.573
## 38165   0037075271 4.109 2002     2700      1     2.632
## 38172   0037070259 4.711 2002     2700      1     3.333
## 38192   0036155211 4.109 2002     2700      1     2.936
## 38200   0037045561 4.075 2002     2700      1     2.697
## 38208   0037034257 6.602 2002     2700      1     5.224
## 38217   0037028761 4.268 2002     2700      1     2.995
## 38218   0037028766 3.946 2002     2700      1     2.773
## 38220   0037028769 4.250 2002     2700      1     2.872
## 38247   0037006397 4.366 2002     2700      1     2.988
## 38328   0037204196 4.718 2002     2700      1     3.340
## 38357   0037160567 3.981 2002     2700      1     2.603
## 38367   0037132805 3.880 2002     2700      1     2.707
## 38375   0037122950 4.667 2002     2700      1     3.394
## 38377   0037122987 3.981 2002     2700      1     2.603
## 38381   0037116641 6.031 2002     2700      1     4.653
## 38382   0037116642 5.093 2002     2700      1     3.715
## 38383   0037116647 4.485 2002     2700      1     3.212
## 38384   0037116658 3.807 2002     2700      1     2.634
## 38390   0037065519 4.048 2002     2700      1     2.875
```

```
## 38398   0037050352 5.876 2002     2700      1     4.399
## 38407   0037045439 4.562 2002     2700      1     3.184
## 38421   0037012134 4.436 2002     2700      1     3.058
## 38430   0037006110 4.025 2002     2700      1     2.548
## 38432   0037006117 3.976 2002     2700      1     2.598
## 38529   0036083829 4.036 2002     2700      1     2.658
## 38535   0036086409 4.229 2002     2700      1     2.752
## 38536   0036086413 3.999 2002     2700      1     2.726
## 38553   0036106054 4.887 2002     2700      1     3.714
## 38572   0036123845 3.948 2002     2700      1     2.570
## 38606   0036143018 4.750 2002     2700      1     3.273
## 38662   0036168150 3.908 2002     2700      1     2.635
## 38755   0036210052 4.381 2002     2700      1     3.108
## 38789   0036239507 4.691 2002     2700      1     3.214
## 39652   0036731990 3.940 2002     2700      1     2.767
## 39758   0036800759 3.956 2002     2700      1     2.578
## 39896   0036894541 4.121 2002     2700      1     2.743
## 40605   0346030568 5.028 2003     2700      1     3.817
## 40608   0347882750 5.233 2003     2700      1     3.917
## 40610   0348011697 4.079 2003     2700      1     2.763
## 40638  10744222633 4.103 2003     2700      1     2.787
## 40646   0344235454 4.984 2003     2700      1     3.873
## 40657   0344738736 4.084 2003     2700      1     2.973
## 40659   0345168925 4.208 2003     2700      1     2.892
## 40660   0345492466 4.735 2003     2700      1     3.419
## 40682   0345690179 5.350 2003     2700      1     4.034
## 40689   0344630326 3.954 2003     2700      1     2.638
## 40731   0346340042 4.157 2003     2700      1     2.841
## 40749   0346969977 4.442 2003     2700      1     3.231
## 40771   0348230958 6.275 2003     2700      1     4.959
## 41119   0345283207 3.840 2003     2700      1     2.524
## 41122   0345714860 5.499 2003     2700      1     4.084
## 41127   0344081179 4.283 2003     2700      1     2.868
## 41128   0344861847 4.621 2003     2700      1     3.305
## 41129   0345293127 3.833 2003     2700      1     2.517
## 41131   0345293130 4.162 2003     2700      1     3.051
## 41139   0345414665 4.343 2003     2700      1     2.928
## 41148   0344926414 5.006 2003     2700      1     3.895
## 41151   0642278662 3.806 2003     2700      1     2.595
## 41156   0344943245 4.973 2003     2700      1     3.657
## 41157   0345374595 4.903 2003     2700      1     3.587
## 41159   0642272544 4.675 2003     2700      1     3.260
## 41161   0242551979 3.904 2003     2700      1     2.588
## 41176   0242658928 4.458 2003     2700      1     3.142
## 41183   0242410368 4.973 2003     2700      1     3.657
## 41185   0242493774 4.400 2003     2700      1     2.985
## 41210   0242493661 4.273 2003     2700      1     2.957
## 41237   0344444858 4.520 2003     2700      1     3.105
## 41251   0345824713 4.352 2003     2700      1     3.036
## 41258   0346969978 5.002 2003     2700      1     3.587
```

```
## 41458   0142025460 3.894 2003     2700      1     2.578
## 41460   0142089171 4.888 2003     2700      1     3.473
## 41463   0142123411 4.222 2003     2700      1     2.906
## 41468   0142087597 4.833 2003     2700      1     3.517
## 41484   0142009533 5.306 2003     2700      1     3.891
## 41493   0142024742 4.567 2003     2700      1     3.152
## 41496   0142088521 4.407 2003     2700      1     3.091
## 41531   0141863194 4.217 2003     2700      1     2.901
## 41575   0141653838 4.222 2003     2700      1     2.906
## 41587   0141705375 4.884 2003     2700      1     3.469
## 41593   0141816759 4.768 2003     2700      1     3.353
## 41594   0141816761 4.287 2003     2700      1     3.076
## 41647   0242285697 4.478 2003     2700      1     3.162
## 41655   0242380323 3.866 2003     2700      1     2.755
## 41771   0141425718 4.211 2003     2700      1     2.796
## 41772   0141484564 4.167 2003     2700      1     2.851
## 41777   0141462439 4.563 2003     2700      1     3.148
## 41778   0141573545 5.008 2003     2700      1     3.593
## 41779   0141573546 3.904 2003     2700      1     2.793
## 41781   0141685592 4.301 2003     2700      1     2.886
## 41782   0141796735 4.455 2003     2700      1     3.244
## 41783   0141796739 4.693 2003     2700      1     3.377
## 41807   0141611906 4.138 2003     2700      1     2.822
## 41813   0141834950 4.316 2003     2700      1     3.000
## 41829   0042387879 4.992 2003     2700      1     3.577
## 41838   0042413423 4.335 2003     2700      1     3.124
## 41850   0042360213 5.273 2003     2700      1     3.957
## 41855   0041381147 3.891 2003     2700      1     2.780
## 41856   0041381153 4.095 2003     2700      1     2.779
## 41858   0042884162 3.758 2003     2700      1     2.647
## 41886   0141612004 3.855 2003     2700      1     2.539
## 41945   0142217904 3.947 2003     2700      1     2.532
## 42124   0041327804 5.015 2003     2700      1     3.699
## 42140   0042422040 4.594 2003     2700      1     3.483
## 42142   0042922806 4.186 2003     2700      1     2.975
## 42150   0041464854 4.417 2003     2700      1     3.101
## 42171   0042195833 4.746 2003     2700      1     3.430
## 42172   0042195862 4.003 2003     2700      1     2.792
## 42175   0042697063 4.958 2003     2700      1     3.642
## 42188   0041912565 4.113 2003     2700      1     2.797
## 42190   0042413567 4.065 2003     2700      1     2.650
## 42191   0042914547 3.963 2003     2700      1     2.852
## 42203   0042661252 4.135 2003     2700      1     2.720
## 42210   0041735992 4.538 2003     2700      1     3.222
## 42216   0042093742 5.388 2003     2700      1     4.177
## 42217   0042594633 4.172 2003     2700      1     2.856
## 42222   0041708062 4.398 2003     2700      1     3.287
## 42223   0041708064 4.291 2003     2700      1     3.080
## 42228   0042125511 4.751 2003     2700      1     3.336
## 42236   0041385951 4.038 2003     2700      1     2.927
```

```
## 42246   0041886776 3.994 2003     2700      1     2.678
## 42267   0042888943 4.611 2003     2700      1     3.295
## 42397   0041842634 4.262 2003     2700      1     2.946
## 42398   0042343801 5.646 2003     2700      1     4.330
## 42404   0041698076 4.135 2003     2700      1     2.720
## 42412   0038497523 4.054 2003     2700      1     2.843
## 42420   0038455694 4.334 2003     2700      1     3.018
## 42438   0038825532 3.870 2003     2700      1     2.554
## 42444   0038601952 4.724 2003     2700      1     3.613
## 42449   0038304776 5.520 2003     2700      1     4.204
## 42452   0038679758 4.587 2003     2700      1     3.172
## 42462   0038013919 4.018 2003     2700      1     2.702
## 42464   0038352047 3.855 2003     2700      1     2.744
## 42466   0038690424 5.414 2003     2700      1     4.098
## 42467   0038690437 4.549 2003     2700      1     3.338
## 42489   0038504054 4.322 2003     2700      1     2.907
## 42515   0041302382 4.281 2003     2700      1     2.965
## 42539   0042304086 4.467 2003     2700      1     3.052
## 42551   0042743965 3.866 2003     2700      1     2.550
## 42564   0043245147 4.497 2003     2700      1     3.181
## 42684   0037830097 3.972 2003     2700      1     2.656
## 42690   0037973279 5.594 2003     2700      1     4.278
## 42691   0038482206 5.915 2003     2700      1     4.704
## 42699   0038167811 4.243 2003     2700      1     3.132
## 42713   0038130715 5.029 2003     2700      1     3.713
## 42719   0037493499 4.842 2003     2700      1     3.526
## 42722   0038507413 4.230 2003     2700      1     3.119
## 42729   0038434032 3.859 2003     2700      1     2.543
## 42736   0037840165 4.892 2003     2700      1     3.576
## 42745   0038242951 4.006 2003     2700      1     2.690
## 42746   0038242968 5.038 2003     2700      1     3.722
## 42754   0037805278 4.410 2003     2700      1     2.995
## 42756   0038142845 4.115 2003     2700      1     2.904
## 42757   0038142850 3.852 2003     2700      1     2.741
## 42772   0037986208 4.721 2003     2700      1     3.405
## 42773   0038323914 4.575 2003     2700      1     3.259
## 42780   0038314212 3.960 2003     2700      1     2.644
## 42893   0042532322 4.480 2003     2700      1     3.164
## 43069   0038526363 4.506 2003     2700      1     3.091
## 43074   0037638884 5.206 2003     2700      1     3.890
## 43077   0038724280 4.848 2003     2700      1     3.737
## 43085   0037737900 5.165 2003     2700      1     3.849
## 43086   0037737901 3.822 2003     2700      1     2.506
## 43088   0037947320 3.982 2003     2700      1     2.666
## 43089   0038075468 5.272 2003     2700      1     3.956
## 43091   0038751994 4.929 2003     2700      1     3.818
## 43115   0038663167 4.361 2003     2700      1     3.045
## 43116   0038663174 4.006 2003     2700      1     2.895
## 43129   0038316599 4.797 2003     2700      1     3.382
## 43130   0038655478 3.979 2003     2700      1     2.868
```

```
## 43132   0037906573 4.466 2003     2700      1     3.051
## 43138   0012868624 4.671 2003     2700      1     3.460
## 43146   0038824056 4.162 2003     2700      1     3.051
## 43153   0038298787 4.370 2003     2700      1     3.054
## 43194   0037988905 4.148 2003     2700      1     2.832
## 43240   0038651907 4.577 2003     2700      1     3.261
## 43241   0038651918 3.988 2003     2700      1     2.672
## 43286   0242600542 4.485 2003     2700      1     3.070
## 43287   0242684416 5.438 2003     2700      1     4.227
## 43416   0037464536 4.084 2003     2700      1     2.873
## 43420   0345636017 4.427 2003     2700      1     3.111
## 43428   0038528237 4.980 2003     2700      1     3.664
## 43435   0037451905 5.245 2003     2700      1     3.929
## 43438   0037451929 6.034 2003     2700      1     4.718
## 43443   0037448932 3.966 2003     2700      1     2.650
## 43479   0037502809 5.082 2003     2700      1     3.971
## 43481   0037840394 4.140 2003     2700      1     3.029
## 43483   0037986313 5.056 2003     2700      1     3.740
## 43485   0038516861 4.525 2003     2700      1     3.314
## 43498   0037417219 4.629 2003     2700      1     3.418
## 43510   0037414165 3.901 2003     2700      1     2.585
## 43511   0037414194 4.269 2003     2700      1     2.953
## 43512   0037414217 4.123 2003     2700      1     2.807
## 43552   0037396734 4.424 2003     2700      1     3.108
## 43561   0037399439 5.125 2003     2700      1     3.809
## 43733   0037468682 3.829 2003     2700      1     2.718
## 43734   0037468691 4.429 2003     2700      1     3.113
## 43739   0037467321 3.688 2003     2700      1     2.577
## 43766   0037454273 4.160 2003     2700      1     2.844
## 43767   0037454282 4.138 2003     2700      1     2.723
## 43768   0037454283 4.138 2003     2700      1     2.822
## 43785   0037434851 3.802 2003     2700      1     2.691
## 43787   0037434895 4.553 2003     2700      1     3.237
## 43807   0037420270 4.012 2003     2700      1     2.696
## 43808   0037420274 5.142 2003     2700      1     3.826
## 43884   0037364369 4.444 2003     2700      1     3.128
## 44099   0037468409 3.866 2003     2700      1     2.550
## 44100   0037468422 4.369 2003     2700      1     3.053
## 44104   0037466917 4.108 2003     2700      1     2.693
## 44120   0037456351 4.578 2003     2700      1     3.262
## 44121   0037456358 4.402 2003     2700      1     3.086
## 44125   0037453950 4.520 2003     2700      1     3.204
## 44127   0037453976 4.581 2003     2700      1     3.166
## 44133   0037440670 3.950 2003     2700      1     2.634
## 44150   0037434515 4.739 2003     2700      1     3.423
## 44159   0037433178 4.012 2003     2700      1     2.597
## 44160   0037433181 3.866 2003     2700      1     2.550
## 44166   0037425790 3.859 2003     2700      1     2.648
## 44178   0037419908 4.233 2003     2700      1     2.917
## 44182   0037289753 4.297 2003     2700      1     2.981
```

```
## 44192   0037308404 5.103 2003     2700      1     3.787
## 44201   0037311377 4.989 2003     2700      1     3.574
## 44212   0037313797 3.966 2003     2700      1     2.650
## 44261   0037330518 3.985 2003     2700      1     2.669
## 44428   0037438792 4.376 2003     2700      1     3.060
## 44429   0037438809 4.264 2003     2700      1     3.053
## 44432   0037438936 3.985 2003     2700      1     2.669
## 44433   0037438998 4.038 2003     2700      1     2.722
## 44454   0037425535 5.533 2003     2700      1     4.118
## 44455   0037425558 4.402 2003     2700      1     3.086
## 44456   0037425564 5.704 2003     2700      1     4.388
## 44457   0037425578 5.336 2003     2700      1     4.020
## 44461   0037417523 5.839 2003     2700      1     4.628
## 44468   0037413484 4.422 2003     2700      1     3.211
## 44469   0037413492 5.094 2003     2700      1     3.883
## 44472   0037413628 3.962 2003     1000      2     2.547
## 44490   0037216353 5.144 2003     2700      1     3.828
## 44501   0037218814 4.737 2003     2700      1     3.421
## 44525   0037237980 3.908 2003     2700      1     2.592
## 44529   0037241112 4.447 2003     2700      1     3.131
## 44541   0037248459 3.918 2003     2700      1     2.707
## 44543   0037249283 3.684 2003     2700      1     2.573
## 44546   0037251810 4.453 2003     2700      1     3.137
## 44690   0037330279 4.140 2003     2700      1     2.824
## 44719   0037382166 3.982 2003     2700      1     2.666
## 44720   0037382807 4.266 2003     2700      1     2.950
## 44745   0037498595 4.430 2003     2700      1     3.219
## 44793   0037685168 4.190 2003     2700      1     2.874
## 44806   0037716856 4.108 2003     2700      1     2.792
## 44889   0038165476 3.954 2003     2700      1     2.539
## 44918   0038303393 3.711 2003     2700      1     2.600
## 45018   0038731023 4.110 2003     2700      1     2.695
## 45080   0041878535 4.148 2003     2700      1     2.832
## 45330   0242515752 3.911 2003     2700      1     2.595
## 45365   0345374590 4.847 2003     2700      1     3.531
## 45366   0345374591 4.435 2003     2700      1     3.324
## 45373   0345690174 3.988 2003     2700      1     2.672
## 45403   0346843100 3.991 2003     2700      1     2.675
## 45422   0347986777 4.410 2003     2700      1     3.094
## 46148   0042737443 4.087 2003     2700      1     2.672
## 46153   0141499228 4.645 2003     2700      1     3.434
## 46333  11144239923 5.214 2004     2700      1     3.986
## 46334  19944395825 4.649 2004     2700      1     3.421
## 46341  10844262691 4.397 2004     2700      1     3.374
## 46342  10844281210 4.459 2004     2700      1     3.231
## 46367  10344231441 4.676 2004     2700      1     3.348
## 46368  10344239881 4.585 2004     2700      1     3.357
## 46369  10344242939 4.474 2004     2700      1     3.246
## 46375  10344221567 4.188 2004     2700      1     2.960
## 46376  10344229444 4.159 2004     2700      1     3.036
```

```
## 46381  11344275794 3.847 2004     2700      1     2.619
## 46424  19744365355 3.936 2004     2700      1     2.913
## 46467  10044271620 4.226 2004     2700      1     2.998
## 46589  15244363856 4.164 2004     2700      1     2.836
## 46593  15744363493 3.986 2004     2700      1     2.863
## 46597  15744399190 4.100 2004     2700      1     2.872
## 46820   9644252909 4.467 2004     2700      1     3.239
## 47011   9244240268 3.964 2004     2700      1     2.736
## 47018   9244240769 4.129 2004     2700      1     2.801
## 47019   9244247612 4.578 2004     2700      1     3.555
## 47020   9244264413 3.803 2004     2700      1     2.575
## 47044   8544252449 4.680 2004     2700      1     3.557
## 47045   8544258102 4.526 2004     2700      1     3.403
## 47053   8444225132 4.176 2004     2700      1     2.948
## 47062  19644400578 4.870 2004     2700      1     3.542
## 47069   7744231805 4.691 2004     2700      1     3.463
## 47070   7744237066 4.382 2004     2700      1     3.154
## 47072   7744239901 4.468 2004     2700      1     3.345
## 47085   7644221218 3.947 2004     2700      1     2.824
## 47140  13844273390 4.129 2004     2700      1     3.006
## 47147  13944251605 4.572 2004     2700      1     3.344
## 47168  15244339164 4.231 2004     2700      1     3.003
## 47393   6944244875 4.867 2004     2700      1     3.639
## 47395   6944248992 4.460 2004     2700      1     3.232
## 47399   6944229468 4.497 2004     2700      1     3.269
## 47407   6944232728 4.311 2004     2700      1     3.083
## 47443   5044248578 3.831 2004     2700      1     2.603
## 47458  19544385747 5.348 2004     2700      1     4.120
## 47601   7444240833 3.982 2004     2700      1     2.859
## 47764   4544305451 4.264 2004     2700      1     3.241
## 47790   4544279029 4.188 2004     2700      1     2.960
## 47806   4544332903 5.079 2004     2700      1     3.851
## 47815   4544222914 3.776 2004     2700      1     2.548
## 47825   4444358501 4.553 2004     2700      1     3.325
## 47843   4544329012 3.814 2004     2700      1     2.586
## 47851   4344690525 4.410 2004     2700      1     3.082
## 47995   4944239035 4.273 2004     2700      1     3.045
## 48006   5444255241 3.855 2004     2700      1     2.627
## 48052   8544244084 4.077 2004     2700      1     2.849
## 48054   9644268242 3.755 2004     2700      1     2.527
## 48197   4143052707 4.043 2004     2700      1     2.815
## 48207   4143067005 4.774 2004     2700      1     3.546
## 48208   4143069253 3.924 2004     2700      1     2.696
## 48227   4143071570 4.102 2004     2700      1     2.774
## 48228   4143094988 5.453 2004     2700      1     4.225
## 48236   4043076922 3.999 2004     2700      1     2.771
## 48238   4043153049 4.489 2004     2700      1     3.261
## 48239   4043156247 4.890 2004     2700      1     3.662
## 48367   4043055316 4.632 2004     2700      1     3.404
## 48373   4043152981 3.933 2004     2700      1     2.705
```

```
## 48385   4344583554 4.318 2004     2700      1     3.090
## 48512   3342892905 4.641 2004     2700      1     3.413
## 48522   3242759883 3.736 2004     2700      1     2.508
## 48553   3142735110 4.217 2004     2700      1     3.194
## 48554   3142745348 4.790 2004     2700      1     3.462
## 48585   3042728480 4.043 2004     2700      1     3.020
## 48665   3042642128 5.510 2004     2700      1     4.282
## 48672   3042731163 4.134 2004     2700      1     2.906
## 48674   3042781155 3.884 2004     2700      1     2.656
## 48679   3042821849 3.670 2004     2700      1     2.647
## 48702   3242749074 4.079 2004     2700      1     2.851
## 48705   3242754344 3.746 2004     2700      1     2.623
## 48718   4043070821 3.779 2004     2700      1     2.551
## 48720   4043082182 3.928 2004     2700      1     2.700
## 48721   4043092238 4.451 2004     2700      1     3.223
## 48951   2942537772 4.149 2004     2700      1     3.026
## 48952   2942554887 5.618 2004     2700      1     4.495
## 48957   2942627194 4.077 2004     2700      1     2.849
## 49006   2542548063 5.360 2004     2700      1     4.337
## 49069   2942709937 3.782 2004     2700      1     2.554
## 49074   2942720933 4.617 2004     2700      1     3.289
## 49281   2442665224 4.280 2004     2700      1     3.052
## 49283   2442715038 5.135 2004     2700      1     3.907
## 49290   2442696436 4.547 2004     2700      1     3.319
## 49306   2342471392 6.269 2004     2700      1     5.041
## 49307   2442479695 5.758 2004     2700      1     4.530
## 49309   2442572117 5.004 2004     2700      1     3.776
## 49313   2442590642 4.673 2004     2700      1     3.445
## 49338   2342464257 4.234 2004     2700      1     3.006
## 49345   2342486731 4.168 2004     2700      1     2.940
## 49369   2342517421 5.147 2004     2700      1     4.124
## 49373   3843094224 4.764 2004     2700      1     3.536
## 49380   2342501856 4.398 2004     2700      1     3.170
## 49381   2342564429 4.008 2004     2700      1     2.780
## 49424   2342488880 3.755 2004     2700      1     2.527
## 49427   2342516208 3.700 2004     2700      1     2.677
## 49428   2342524107 3.852 2004     2700      1     2.624
## 49434   2342646956 4.407 2004     2700      1     3.179
## 49472   2942536636 4.358 2004     2700      1     3.130
## 49475   2942561024 4.094 2004     2700      1     2.866
## 49649   2142649221 4.212 2004     2700      1     3.189
## 49651   2142758687 4.063 2004     2700      1     2.835
## 49664   1942436013 4.103 2004     2700      1     2.875
## 49674  11144354823 4.207 2004     2700      1     2.979
## 49696  12144291219 3.697 2004     2700      1     2.574
## 49710   3042762336 4.314 2004     2700      1     3.086
## 49718  11144357519 3.986 2004     2700      1     2.658
## 49722   1842864234 5.018 2004     2700      1     3.790
## 49739   1842815777 4.502 2004     2700      1     3.174
## 49743  11144356419 3.866 2004     2700      1     2.638
```

```
## 49830   2442509789 3.581 2004     2700      1     2.558
## 49967  12144288049 4.590 2004     2700      1     3.362
## 49970   1642400686 4.214 2004     2700      1     2.986
## 49977   1642296706 3.739 2004     2700      1     2.511
## 50015  11144355828 4.350 2004     2700      1     3.122
## 50019  10744229257 3.730 2004     2700      1     2.502
## 50026   1642369930 4.229 2004     2700      1     3.001
## 50055   1442355581 4.481 2004     2700      1     3.253
## 50062  10744220250 4.551 2004     2700      1     3.223
## 50063  10744225301 5.224 2004     2700      1     3.996
## 50065   1442353066 3.979 2004     2700      1     2.751
## 50071  11144357189 3.839 2004     2700      1     2.511
## 50087   1442357943 3.847 2004     2700      1     2.619
## 50332  10744233940 4.677 2004     2700      1     3.449
## 50335   1342331884 3.850 2004     2700      1     2.622
## 50388   1442314663 3.997 2004     2700      1     2.669
## 50413  10744223871 4.045 2004     2700      1     2.817
## 50414   1342288777 4.142 2004     2700      1     3.119
## 50468   1642327683 4.090 2004     2700      1     2.967
## 50613   1642499234 4.500 2004     2700      1     3.272
## 50622   1642540483 4.321 2004     2700      1     2.993
## 50623   9144252520 4.184 2004     2700      1     2.956
## 50643   0345872128 4.849 2004     2700      1     3.621
## 50657   0347948542 3.871 2004     2700      1     2.643
## 50661   0345824715 4.381 2004     2700      1     3.153
## 50674   0346599193 4.602 2004     2700      1     3.374
## 50675   0347416893 3.842 2004     2700      1     2.614
## 50678   0348047527 4.767 2004     2700      1     3.439
## 50695   0345830478 3.919 2004     2700      1     2.591
## 50712   0346787782 3.666 2004     2700      1     2.543
## 50713   0346787909 3.820 2004     2700      1     2.592
## 50725   0347418277 4.315 2004     2700      1     3.292
## 50732   0742306162 3.544 2004     2700      1     2.521
## 50771  10944261840 3.779 2004     2700      1     2.656
## 50818   1342332130 3.828 2004     2700      1     2.600
## 50848  13744263906 4.057 2004     2700      1     3.034
## 50998   1542571327 3.817 2004     2700      1     2.589
## 51428   2142717404 4.031 2004     2700      1     2.803
## 51821   3242733807 3.931 2004     2700      1     2.703
## 51895   3442899889 4.734 2004     2700      1     3.506
## 51927   4043085036 3.742 2004     2700      1     2.514
## 52262  85047690933 4.353 2004     2700      1     3.125
## 52323   9644283066 4.178 2004     2700      1     2.950
## 52387   3142514201 3.624 2004     1300      2     2.601
## 52521  29544447206 3.996 2005     2700      1     2.697
## 52524  29544437848 4.305 2005     2700      1     3.006
## 52539  29144451858 4.279 2005     2700      1     3.184
## 52543  29144490030 3.866 2005     2700      1     2.567
## 52545  29144533834 4.497 2005     2700      1     3.297
## 52558  28944434329 4.369 2005     2700      1     3.169
```

```
## 52561  28944445445 4.451 2005     2700      1     3.356
## 52565  28944433025 3.552 2005     2700      1     2.557
## 52569  28944447646 4.919 2005     2700      1     3.620
## 52607  28944437578 4.673 2005     2700      1     3.678
## 52614  28844472594 3.996 2005     2700      1     2.697
## 52618  28844501802 4.241 2005     2700      1     3.041
## 52767  30144444279 4.733 2005     2700      1     3.533
## 52796  31044442642 3.817 2005     2700      1     2.617
## 52797  31044444463 4.156 2005     2700      1     2.956
## 52803  31044451759 4.265 2005     2700      1     3.065
## 52804  31044452189 4.001 2005     2700      1     2.702
## 53762  28144433147 4.679 2005     2700      1     3.380
## 53764  28144453057 4.091 2005     2700      1     2.891
## 53771  28144443357 3.807 2005     2700      1     2.712
## 53772  28144451163 4.332 2005     2700      1     3.132
## 53773  28144453509 3.901 2005     2700      1     2.806
## 53794  27744431927 4.342 2005     2700      1     3.142
## 53795  27744494434 4.618 2005     2700      1     3.623
## 53807  27744496592 4.514 2005     2700      1     3.314
## 53820  27744606737 4.954 2005     2700      1     3.754
## 53828  27744477622 3.728 2005     2700      1     2.528
## 53845  27644443333 4.575 2005     2700      1     3.580
## 53863  27644452293 3.749 2005     2700      1     2.549
## 53865  27644461904 3.702 2005     2700      1     2.707
## 53875  27644513772 4.335 2005     2700      1     3.240
## 53883  27644568878 4.218 2005     2700      1     3.223
## 53918  28444478439 4.387 2005     2700      1     3.088
## 54155  27244440305 4.455 2005     2700      1     3.360
## 54170  26844536978 5.830 2005     2700      1     4.630
## 54171  26844552488 5.412 2005     2700      1     4.212
## 54178  26844494491 4.633 2005     2700      1     3.538
## 54179  26844538114 4.966 2005     2700      1     3.766
## 54180  26844544418 3.601 2005     2700      1     2.506
## 54208  26444452073 5.212 2005     2700      1     3.913
## 54210  26444477603 4.006 2005     2700      1     3.011
## 54211  26444506232 4.042 2005     2700      1     2.842
## 54217  26444574824 4.419 2005     2700      1     3.120
## 54243  25844488342 3.971 2005     2700      1     2.771
## 54246  25844509466 4.220 2005     2700      1     2.921
## 54285  25844438380 4.442 2005     2700      1     3.242
## 54287  25844440999 3.660 2005     2700      1     2.665
## 54305  26444515022 3.926 2005     2700      1     2.627
## 54312  26444551183 3.878 2005     2700      1     2.678
## 54314  26444577543 3.881 2005     2700      1     2.681
## 54319  26444598506 4.478 2005     2700      1     3.278
## 54557  25444461935 3.898 2005     2700      1     2.698
## 54562  25144470719 3.920 2005     2700      1     2.720
## 54575  25144456112 5.883 2005     2700      1     4.683
## 54580  25144502974 3.881 2005     2700      1     2.681
## 54581  25144505285 3.742 2005     2700      1     2.542
```

```
## 54582  25144518364 4.074 2005     2700      1     2.874
## 54614  24944531368 3.878 2005     2700      1     2.678
## 54632  24644439190 4.674 2005     2700      1     3.474
## 54633  24644443217 4.833 2005     2700      1     3.633
## 54677  24344483122 3.953 2005     2700      1     2.753
## 54701  24044484768 3.823 2005     2700      1     2.728
## 54742  24744470476 4.035 2005     2700      1     2.835
## 55090  23944436856 3.575 2005     2700      1     2.580
## 55101  23944436934 4.091 2005     2700      1     2.891
## 55113  23844539714 4.976 2005     2700      1     3.776
## 55121  23844454293 3.800 2005     2700      1     2.600
## 55123  23844494864 4.211 2005     2700      1     2.912
## 55133  23844533906 4.026 2005     2700      1     2.826
## 55149  23744500093 4.222 2005     2700      1     2.923
## 55162  23444448145 3.866 2005     2700      1     2.771
## 55163  23444452077 4.300 2005     2700      1     3.100
## 55164  23744459272 3.820 2005     2700      1     2.620
## 55189  23044480500 3.937 2005     2700      1     2.737
## 55193  23044442114 4.460 2005     2700      1     3.260
## 55235  23644435612 3.976 2005     2700      1     2.677
## 55238  23644439061 4.516 2005     2700      1     3.316
## 55240  23644446866 3.989 2005     2700      1     2.789
## 55244  23644456321 3.901 2005     2700      1     2.701
## 55245  23644460325 3.942 2005     2700      1     2.947
## 55428  22844445230 4.774 2005     2700      1     3.475
## 55430  22844447051 4.585 2005     2700      1     3.385
## 55438  22844434877 4.780 2005     2700      1     3.580
## 55442  22844448780 4.222 2005     2700      1     3.022
## 55448  22144468433 3.996 2005     2700      1     2.796
## 55466  22344443114 4.392 2005     2700      1     3.192
## 55468  22344450570 4.102 2005     2700      1     2.902
## 55471  22344457679 4.119 2005     2700      1     2.919
## 55485  22244446183 4.420 2005     2700      1     3.425
## 55486  22244485721 4.183 2005     2700      1     3.088
## 55521  22144499588 3.713 2005     2700      1     2.618
## 55536  21444447949 4.377 2005     2700      1     3.177
## 55569  21444441271 4.299 2005     2700      1     3.000
## 55570  21444447966 4.369 2005     2700      1     3.169
## 55571  21444450527 4.501 2005     2700      1     3.406
## 55591  22144440464 4.362 2005     2700      1     3.162
## 55605  22144492953 4.207 2005     2700      1     2.908
## 55606  22144495740 3.790 2005     2700      1     2.590
## 55869  20644465409 3.857 2005     2700      1     2.657
## 55886  20444444376 4.187 2005     2700      1     2.987
## 55888  20444501831 4.934 2005     2700      1     3.734
## 55913  20444398121 3.931 2005     2700      1     2.632
## 55914  20544462061 4.452 2005     2700      1     3.357
## 55957  21144448879 4.965 2005     2700      1     3.765
## 55970  18844374582 3.884 2005     2700      1     2.889
## 55986  19644379430 3.989 2005     2700      1     2.789
```

```
## 55989  19644391940 4.484 2005     2700      1     3.185
## 56001  20444454585 3.817 2005     2700      1     2.617
## 56002  20444457518 4.596 2005     2700      1     3.601
## 56009  20444499355 4.459 2005     2700      1     3.259
## 56310  19444363154 3.931 2005     2700      1     2.731
## 56315  21144457072 4.436 2005     2700      1     3.236
## 56318  18944367562 4.915 2005     2700      1     3.616
## 56337  18644370841 4.807 2005     2700      1     3.812
## 56339  18744382834 3.953 2005     2700      1     2.958
## 56341  20844438843 4.222 2005     2700      1     3.022
## 56349  18244371651 5.531 2005     2700      1     4.331
## 56350  20844432890 4.096 2005     2700      1     3.001
## 56396  18744401357 3.950 2005     2700      1     2.651
## 56397  18744406319 3.706 2005     2700      1     2.506
## 56407  17944362664 4.751 2005     2700      1     3.452
## 56409  17944380530 3.760 2005     2700      1     2.560
## 56410  20944448998 3.953 2005     2700      1     2.753
## 56491  20944436157 4.708 2005     2700      1     3.508
## 56656  17644392071 4.823 2005     2700      1     3.524
## 56664  17544364774 4.269 2005     2700      1     3.069
## 56684  17844402965 4.271 2005     2700      1     3.071
## 56693  17244379193 4.136 2005     2700      1     2.936
## 56702  17144385790 5.319 2005     2700      1     4.020
## 56705  17144430418 4.592 2005     2700      1     3.392
## 56707  21544449115 4.018 2005     2700      2     2.719
## 56712  17144398838 3.739 2005     2700      1     2.539
## 56719  20244366093 3.760 2005     2700      1     2.560
## 56730  20144386747 4.154 2005     2700      1     2.954
## 56731  20144388867 4.033 2005     2700      1     2.833
## 56738  16444381731 3.829 2005     2700      1     2.734
## 56740  20144387681 4.355 2005     2700      1     3.155
## 56748  15744372810 5.001 2005     2700      1     3.801
## 56757  20144386837 4.994 2005     2700      1     3.999
## 56762  15944387060 3.940 2005     2700      1     2.641
## 56763  15944406488 4.216 2005     2700      1     3.016
## 56824  20144387499 3.895 2005     2700      1     2.596
## 56965  15744378579 5.147 2005     2700      1     4.152
## 56967  20144366696 4.939 2005     2700      1     3.739
## 56992  15044344420 3.717 2005     2700      1     2.517
## 57041  14744275842 4.576 2005     2700      1     3.376
## 57048  14544304095 5.144 2005     2700      1     3.944
## 57060  20044369898 3.543 2005     2700      1     2.548
## 57105  14644435711 3.499 2005     2700      1     2.504
## 57189  20044368802 4.056 2005     2700      1     2.856
## 57354  13844315270 3.929 2005     2700      1     2.630
## 57359  13944250706 3.610 2005     2700      1     2.615
## 57391  13744253924 3.839 2005     2700      1     2.744
## 57392  13744261333 3.974 2005     2700      1     2.774
## 57400  13844265752 4.225 2005     2700      1     3.025
## 57402  13844274980 3.923 2005     2700      1     2.723
```

```
## 57417  13444256401 4.619 2005     2700      1     3.419
## 57420  13444293219 5.349 2005     2700      1     4.149
## 57426  13444309949 4.491 2005     2700      1     3.291
## 57427  13444310598 3.728 2005     2700      1     2.528
## 57437  13444263619 4.684 2005     2700      1     3.484
## 57495  14944348752 3.863 2005     2700      1     2.663
## 57674  19944432606 4.121 2005     2700      1     3.126
## 57678  12544253745 4.628 2005     2700      1     3.428
## 57683  12544253835 4.548 2005     2700      1     3.348
## 57684  12544257550 4.190 2005     2700      1     2.990
## 57685  19944431657 4.004 2005     2700      1     2.804
## 57698  12144284532 3.878 2005     2700      1     2.883
## 57714  19944426812 4.047 2005     2700      1     2.748
## 57732  11844268070 4.117 2005     2700      1     2.917
## 57734  13744249546 3.742 2005     2700      1     2.647
## 57739  11844269164 4.196 2005     2700      1     3.101
## 57784  19944429519 4.900 2005     2700      1     3.700
## 57787  11244285095 3.904 2005     2700      1     2.605
## 57788  11244331742 4.906 2005     2700      1     3.706
## 57831  12744263350 4.085 2005     2700      1     3.090
## 57836  12744278427 3.829 2005     2700      1     2.530
## 57868  13444260835 3.732 2005     2700      1     2.637
## 57870  13444268127 4.350 2005     2700      1     3.355
## 57881  13744254861 3.749 2005     2700      1     2.549
## 57883  13744256812 4.438 2005     2700      1     3.139
## 57919  14144253829 3.857 2005     2700      1     2.657
## 57965  14644425916 3.631 2005     2700      1     2.536
## 58080  16244415880 3.909 2005     2700      1     2.914
## 58139  17644377222 4.115 2005     2700      1     3.120
## 58156  17844391245 3.929 2005     2700      1     2.729
## 58196  18244390229 3.826 2005     2700      1     2.831
## 58197  18244393432 3.533 2005     2700      1     2.538
## 58304  20144388724 3.857 2005     2700      1     2.657
## 58349  20944446791 4.391 2005     2700      1     3.191
## 58709  26944497954 3.884 2005     2700      1     2.684
## 59586  33845866114 3.911 2006     2700      1     2.847
## 59588  33845880922 5.714 2006     2700      1     4.749
## 59600  33845866055 4.159 2006     2700      1     2.990
## 59601  33845890326 4.968 2006     2700      1     3.799
## 59635  33845693164 3.800 2006     2700      1     2.631
## 59636  33845707784 4.606 2006     2700      1     3.437
## 59643  33845717310 4.910 2006     2700      1     3.846
## 59678  33845483482 3.948 2006     2700      1     2.983
## 59679  33845490014 6.122 2006     2700      1     4.953
## 59680  33845497716 3.958 2006     2700      1     2.689
## 59681  33845500545 4.941 2006     2700      1     3.772
## 59686  33845478495 4.306 2006     2700      1     3.037
## 59687  33845485717 3.595 2006     2700      1     2.531
## 59688  33845491101 3.880 2006     2700      1     2.611
## 59689  33947325340 4.051 2006     2700      1     2.987
```

```
## 59705  33845449051 4.316 2006     2700      1     3.147
## 59714  33845450660 4.586 2006     2700      1     3.317
## 59746  33748560432 4.653 2006     2700      1     3.484
## 59787  33845297656 3.725 2006     2700      1     2.556
## 59788  33845299821 3.784 2006     2700      1     2.615
## 59792  33845317604 4.246 2006     2700      1     3.077
## 59793  33845322277 3.595 2006     2700      1     2.630
## 59798  33845354407 3.759 2006     2700      1     2.590
## 59849  33845926387 3.808 2006     2700      1     2.639
## 59883  33846317116 3.734 2006     2700      1     2.769
## 60010  33847190697 3.751 2006     2700      1     2.582
## 61367  33751529506 4.659 2006     2700      1     3.390
## 61370  33751545838 5.100 2006     2700      1     3.831
## 61389  33751185526 3.780 2006     2700      1     2.716
## 61412  33751206860 4.493 2006     2700      1     3.324
## 61413  33751217149 4.658 2006     2700      1     3.489
## 61441  33750983605 5.495 2006     2700      1     4.326
## 61448  33750978404 4.084 2006     2700      1     2.915
## 61491  33750694575 4.082 2006     2700      1     2.913
## 61499  33750726707 4.632 2006     2700      1     3.568
## 61500  33750731595 4.693 2006     2700      1     3.524
## 61521  33750630655 3.751 2006     2700      1     2.786
## 61532  33750500326 4.054 2006     2700      1     2.885
## 61535  33750507268 4.113 2006     2700      1     2.844
## 61564  33750507246 3.788 2006     2700      1     2.619
## 61566  33750518186 4.246 2006     2700      1     3.281
## 61570  33750532312 5.148 2006     2700      1     4.183
## 61574  33750584214 5.548 2006     2700      1     4.279
## 61582  33750597734 4.320 2006     2700      1     3.151
## 62170  33750346251 4.432 2006     2700      1     3.368
## 62188  33750300528 4.116 2006     2700      1     2.947
## 62212  33750035507 4.320 2006     2700      1     3.355
## 62216  33750081777 3.983 2006     2700      1     2.714
## 62223  33750106228 4.641 2006     2700      1     3.472
## 62224  33750121615 4.499 2006     2700      1     3.330
## 62227  33749636443 3.738 2006     2700      1     2.773
## 62260  33749597936 4.523 2006     2700      1     3.254
## 62261  33749599695 5.460 2006     2700      1     4.495
## 62263  33749618085 4.879 2006     2700      1     3.710
## 62273  33749613119 5.139 2006     2700      1     4.174
## 62274  33749659678 3.989 2006     2700      1     2.820
## 62294  33749436870 4.405 2006     2700      1     3.236
## 62297  33749445317 5.971 2006     2700      1     4.802
## 62319  33749020857 4.306 2006     2700      1     3.137
## 62325  33749430001 3.938 2006     2700      1     2.973
## 62331  33749441325 4.657 2006     2700      1     3.488
## 62335  33749446861 3.772 2006     2700      1     2.603
## 62472  33846664981 3.928 2006     2700      1     2.759
## 62539  39049174296 3.751 2006     2700      1     2.582
## 62724  43749107283 4.870 2006     2700      1     3.701
```

```
## 62921  33749040947 3.626 2006     2700      1     2.562
## 62924  33749079234 3.958 2006     2700      1     2.789
## 63010  33748354709 3.584 2006     2700      1     2.619
## 63021  33748316794 3.928 2006     2700      1     2.963
## 63047  33748437811 4.028 2006     2700      1     2.859
## 63048  33748438134 4.637 2006     2700      1     3.468
## 63124  33748292955 4.111 2006     2700      1     2.942
## 63132  33748312093 4.065 2006     2700      1     2.896
## 63137  33748331335 3.890 2006     2700      1     2.621
## 63152  33748747557 3.876 2006     2700      1     2.707
## 63833  33748161812 4.051 2006     2700      1     2.882
## 63860  33747834620 3.716 2006     2700      1     2.547
## 63864  33747870163 5.283 2006     2700      1     4.114
## 63866  33747584933 4.308 2006     2700      1     3.139
## 63887  33747135682 4.969 2006     2700      1     3.800
## 63904  33747343208 5.434 2006     2700      1     4.265
## 63910  33747102040 3.816 2006     2700      1     2.647
## 63948  33746875641 4.520 2006     2700      1     3.351
## 63985  33746768834 3.832 2006     2700      1     2.563
## 63987  33746587393 4.057 2006     2700      1     2.993
## 64006  33746437130 4.877 2006     2700      1     3.608
## 64017  33746667851 4.079 2006     2700      1     2.910
## 64020  33746672599 3.964 2006     2700      1     2.695
## 64021  33746675685 4.031 2006     2700      1     2.862
## 64033  33746724157 3.887 2006     2700      1     2.618
## 64083  33748046126 3.747 2006     2700      1     2.578
## 64148  33750630109 3.948 2006     2700      1     2.679
## 64376  33746387027 3.828 2006     2700      1     2.659
## 64377  33746430561 3.742 2006     2700      1     2.573
## 64378  33746430578 4.001 2006     2700      1     2.832
## 64379  33746452905 4.192 2006     2700      1     3.128
## 64395  33748513704 3.579 2006     2700      1     2.614
## 64421  33746075560 4.642 2006     2700      1     3.473
## 64422  33746088001 5.656 2006     2700      1     4.691
## 64482  33745890309 4.593 2006     2700      1     3.324
## 64484  33745913326 3.707 2006     2700      1     2.643
## 64512  33745698681 4.248 2006     2700      1     3.079
## 64514  33745698685 4.688 2006     2700      1     3.624
## 64520  33745611449 5.100 2006     2700      1     4.135
## 64521  33745612500 3.983 2006     2700      1     2.714
## 64522  33745614361 5.082 2006     2700      1     4.117
## 64523  33745632422 4.146 2006     2700      1     2.877
## 64530  33745812011 4.129 2006     2700      1     2.960
## 64531  33745817034 3.992 2006     2700      1     3.027
## 64535  33745840137 3.992 2006     2700      1     2.823
## 64537  33745851479 4.385 2006     2700      1     3.216
## 64548  33744914821 3.540 2006     2700      1     2.575
## 64979  33745661633 3.776 2006     2700      1     2.712
## 65008  33745227399 4.917 2006     2700      1     3.748
## 65009  33745227751 3.734 2006     2700      1     2.565
```

```
## 65010  33745255382 4.457 2006     2700      1     3.393
## 65030  33646360865 4.352 2006     2700      1     3.183
## 65051  33745081608 5.495 2006     2700      1     4.530
## 65054  33745102555 5.215 2006     2700      1     4.046
## 65066  33745065872 3.734 2006     2700      1     2.769
## 65093  33744718154 3.600 2006     2700      1     2.635
## 65106  33744966030 4.777 2006     2700      1     3.608
## 65135  33646825385 4.208 2006     2700      1     3.039
## 65148  33744472167 3.702 2006     2700      1     2.533
## 65149  33744472168 4.090 2006     2700      1     2.921
## 65151  33744498063 4.280 2006     2700      1     3.111
## 65190  33745210374 3.862 2006     2700      1     2.693
## 65191  33745212925 3.876 2006     2700      1     2.707
## 65192  33745217611 3.764 2006     2700      1     2.595
## 65230  33745685701 3.747 2006     2700      1     2.578
## 65338  33750089034 3.980 2006     2700      1     3.015
## 65365  33846406448 3.854 2006     2700      1     2.685
## 65669  33646859471 4.708 2006     2700      1     3.743
## 65672  33646875293 4.246 2006     2700      1     2.977
## 65680  33646828953 3.862 2006     2700      1     2.897
## 65716  33646676272 4.484 2006     2700      1     3.315
## 65717  33646678189 4.410 2006     2700      1     3.241
## 65764  33646482407 3.887 2006     2700      1     2.618
## 65765  33646488054 4.825 2006     2700      1     3.556
## 65772  33646450281 3.851 2006     2700      1     2.582
## 65773  33646452922 4.459 2006     2700      1     3.290
## 65790  33646475441 3.584 2006     2700      1     2.619
## 65816  30044438368 5.433 2006     2700      1     4.369
## 65837  33646177653 3.505 2006     2700      1     2.540
## 65838  33646178951 5.085 2006     2700      1     3.916
## 65847  33646392755 4.004 2006     2700      1     2.835
## 65850  33646401549 3.977 2006     2700      1     2.808
## 65975  33745489614 3.646 2006     2700      1     2.582
## 66274  33646164892 4.031 2006     2700      1     2.862
## 66300  33645865234 4.803 2006     2700      1     3.634
## 66319  33646052556 5.164 2006     2700      1     3.995
## 66338  33645770270 4.219 2006     2700      1     3.155
## 66344  33645729203 4.579 2006     2700      1     3.410
## 66346  33645741222 4.267 2006     2700      1     2.998
## 66372  33645523067 6.444 2006     2700      1     5.380
## 66392  29944434110 3.964 2006     2700      1     2.795
## 66393  29944434742 4.119 2006     2700      1     2.850
## 66394  29944442726 4.794 2006     2700      1     3.625
## 66408  33645222620 3.812 2006     2700      1     2.543
## 66409  33645236930 3.684 2006     2700      1     2.515
## 66423  33645530745 4.561 2006     2700      1     3.392
## 66490  33646420626 3.788 2006     2700      1     2.619
## 66492  33646432226 4.312 2006     2700      1     3.248
## 66777  33645390604 3.621 2006     2700      1     2.656
## 66778  33645399959 4.643 2006     2700      1     3.474
```

```
## 66806  33645102811 3.897 2006     2700      1     2.833
## 66807  33645103550 5.432 2006     2700      1     4.467
## 66813  33645068471 4.040 2006     2700      1     2.771
## 66834  33745029895 3.832 2006     2700      1     2.768
## 66845  33644979597 4.290 2006     2700      1     3.325
## 66847  33644984085 4.144 2006     2700      1     2.975
## 66856  33644967633 3.995 2006     2700      1     3.030
## 66872  33644957630 4.031 2006     2700      1     2.862
## 66880  33644833147 4.678 2006     2700      1     3.509
## 66883  33644807437 5.185 2006     2700      1     3.916
## 66925  33644584065 4.544 2006     2700      1     3.275
## 66926  33644605043 5.148 2006     2700      1     3.979
## 66930  33644638676 3.816 2006     2700      1     2.647
## 66932  33644642974 3.921 2006     2700      1     2.752
## 66942  33644654777 4.106 2006     2700      1     2.937
## 66944  33644660402 3.689 2006     2700      1     2.625
## 67001  33645355399 4.176 2006     2700      1     3.007
## 67354  33344467770 4.351 2006     2700      1     3.182
## 67363  33644536465 4.565 2006     2700      1     3.396
## 67394  32644457434 4.090 2006     2700      1     2.921
## 67395  32644467389 5.178 2006     2700      1     3.909
## 67432  32144436098 4.713 2006     2700      1     3.444
## 67435  32144455695 4.106 2006     2700      1     2.937
## 67438  32544459770 3.748 2006     1000      2     2.579
## 67445  32144432645 4.141 2006     2700      1     3.077
## 67446  32144441493 4.605 2006     2700      1     3.640
## 67447  32144443648 4.776 2006     2700      1     3.712
## 67455  32544444748 3.983 2006     2700      1     2.714
## 67461  32044452414 4.031 2006     2700      1     2.862
## 67474  31444436871 3.689 2006     2700      1     2.520
## 67506  32444436464 3.716 2006     2700      1     2.547
## 67508  32444438623 3.828 2006     2700      1     2.659
## 67510  32444443951 4.627 2006     2700      1     3.358
## 67517  32444450909 3.534 2006     2700      1     2.569
## 67538  33644842676 4.351 2006     2700      1     3.182
## 67781  31544462628 3.679 2006     1000      2     2.510
## 67786  31344467254 4.933 2006     2700      1     3.968
## 67787  31344470705 5.219 2006     2700      1     3.950
## 67790  31344479326 4.396 2006     2700      1     3.227
## 67795  31344454603 3.772 2006     2700      1     2.603
## 67804  30844432417 3.675 2006     2700      1     2.506
## 67813  30944452384 4.432 2006     2700      1     3.263
## 67815  30944457531 4.478 2006     2700      1     3.414
## 67876  32244433703 3.636 2006     2700      1     2.671
## 67920  30944433554 4.252 2006     2700      1     3.083
## 67933  31044431832 4.425 2006     2700      1     3.256
## 67938  31044436630 4.688 2006     2700      1     3.519
## 67941  31044442783 3.734 2006     2700      1     2.565
## 67942  31044446194 4.154 2006     2700      1     3.189
## 67944  31044453008 4.376 2006     2700      1     3.207
```

```
## 67945  31044453148 4.025 2006     2700      1     2.856
## 67950  31044456529 5.039 2006     2700      1     3.870
## 67972  32044434020 3.847 2006     2700      1     2.678
## 67974  32044435429 4.349 2006     2700      1     3.080
## 67991  32144461605 3.995 2006     2700      1     2.726
## 68363  33748146129 3.511 2006     2700      1     2.546
## 68847  37549036689 4.103 2007     2700      1     3.151
## 68848  37549039095 3.740 2007     2700      1     2.888
## 68868  37349092105 3.950 2007     2700      1     2.893
## 68877  37349011273 3.858 2007     2700      1     2.701
## 68878  37349028816 4.130 2007     2700      1     3.073
## 68923  37249041477 3.789 2007     2700      1     2.732
## 68924  37249061691 4.883 2007     2700      1     3.826
## 68931  37149028276 4.444 2007     2700      1     3.592
## 68952  38449085341 3.654 2007     2700      1     2.597
## 68959  36549078679 4.337 2007     2700      1     3.280
## 68962  37249065591 3.858 2007     2700      1     2.906
## 69125  36048934283 3.807 2007     2700      1     2.955
## 69128  36049048260 3.795 2007     2700      1     2.638
## 69155  36849011221 4.476 2007     2700      1     3.419
## 69168  36849035575 3.973 2007     2700      1     2.916
## 69181  36849068108 4.107 2007     2700      1     2.950
## 69184  36849079348 3.564 2007     2700      1     2.612
## 69237  38049023539 4.325 2007     2700      1     3.373
## 69238  38049038389 3.992 2007     2700      1     2.935
## 69239  38049053061 4.415 2007     2700      1     3.358
## 69241  38049087576 4.711 2007     2700      1     3.654
## 69968  56149101731 3.783 2007     2700      1     2.626
## 70668  36549006511 3.668 2007     2700      1     2.611
## 70677  36549033449 4.134 2007     2700      1     2.977
## 70678  36549066931 4.045 2007     2700      1     3.193
## 70725  36348961929 4.189 2007     2700      1     3.132
## 70727  36348982774 4.230 2007     2700      1     3.173
## 70728  36349035495 4.062 2007     2700      1     2.905
## 70766  36148981435 4.360 2007     2700      1     3.303
## 70798  36148951701 3.556 2007     2700      1     2.604
## 70799  36148966515 4.227 2007     2700      1     3.070
## 70800  36149000149 5.380 2007     2700      1     4.323
## 70832  35848963822 4.711 2007     2700      1     3.654
## 70833  35848964343 4.234 2007     2700      1     3.282
## 70835  35848969753 4.853 2007     2700      1     3.796
## 70842  35848935201 5.343 2007     2700      1     4.186
## 70844  35848964571 4.402 2007     2700      1     3.450
## 70845  35848968871 6.256 2007     2700      1     5.199
## 70888  35248896228 3.992 2007     2700      1     2.835
## 70910  35748974550 4.877 2007     2700      1     3.820
## 70923  36048931015 4.213 2007     2700      1     3.156
## 70927  36048938340 3.610 2007     2700      1     2.553
## 70929  36048953291 3.747 2007     2700      1     2.590
## 70932  36048958880 3.498 2007     2700      1     2.646
```

```
## 70934  36048963838 4.213 2007     2700      1     3.156
## 70935  36048968753 4.402 2007     2700      1     3.345
## 70950  36049037185 3.835 2007     2700      1     2.778
## 70951  36049042525 3.835 2007     2700      1     2.778
## 70959  36248973101 3.783 2007     2700      1     2.626
## 71265  41149161956 4.099 2007     2700      1     3.147
## 71308  48849113878 4.849 2007     2700      1     3.792
## 71611  35348968910 3.761 2007     1000      2     2.704
## 71623  35348981185 3.708 2007     2700      1     2.856
## 71624  35348997893 4.319 2007     2700      1     3.467
## 71653  35148831175 4.279 2007     2700      1     3.427
## 71667  34948830633 3.376 2007     2700      1     2.524
## 71669  35248829599 4.645 2007     2700      1     3.693
## 71673  35248895346 4.762 2007     2700      1     3.705
## 71681  35348866794 3.930 2007     2700      1     2.773
## 71714  34948849235 3.498 2007     2700      1     2.546
## 71717  34948891641 4.829 2007     2700      1     3.772
## 71720  34948845507 4.340 2007     2700      1     3.283
## 71722  34948882262 4.028 2007     2700      1     2.871
## 71753  34548827323 4.070 2007     2700      1     3.013
## 71755  34548835744 4.466 2007     2700      1     3.409
## 71788  34948860191 3.564 2007     2700      1     2.507
## 71791  34948872683 3.695 2007     2700      1     2.638
## 71809  35348828533 4.122 2007     2700      1     2.965
## 71831  35348985396 3.688 2007     2700      1     2.531
## 71842  35448937632 3.617 2007     2700      1     2.765
## 71847  35648946842 3.807 2007     2700      1     2.955
## 71864  35649003880 4.876 2007     2700      1     3.719
## 72101  40649100791 3.868 2007     2700      1     2.811
## 72327  34648832073 4.784 2007     2700      1     3.727
## 72334  34748875855 4.118 2007     2700      1     3.061
## 72398  34548833077 4.091 2007     2700      1     3.034
## 72400  34548849816 4.178 2007     2700      1     3.121
## 72454  34548776934 3.753 2007     2700      1     2.696
## 72465  34548570632 4.484 2007     2700      1     3.427
## 72494  34547426346 3.681 2007     2700      1     2.624
## 72517  34548419640 3.595 2007     2700      1     2.743
## 72546  34548418912 4.594 2007     2700      1     3.537
## 72548  34548433987 4.078 2007     2700      1     3.021
## 72555  34548426480 4.019 2007     2700      1     2.862
## 72556  34548430952 4.304 2007     2700      1     3.147
## 72568  34848816502 4.152 2007     2700      1     3.200
## 72569  34848818486 3.879 2007     2700      1     2.822
## 72572  34848831462 4.045 2007     2700      1     2.988
## 72575  34848838902 4.019 2007     2700      1     2.962
## 72578  34848858523 4.244 2007     2700      1     3.187
## 72581  34848881905 4.352 2007     2700      1     3.295
## 72582  34848881907 3.681 2007     2700      1     2.524
## 72583  34848888949 3.954 2007     2700      1     2.897
## 72587  34848901974 4.316 2007     2700      1     3.259
```

```
## 72588  34848902937 4.337 2007     2700      1     3.280
## 72591  34848920368 3.801 2007     2700      1     2.644
## 72606  34547502731 3.531 2007     2700      1     2.579
## 72608  34547541570 3.406 2007     2700      1     2.554
## 72609  34547545269 4.719 2007     2700      1     3.767
## 72613  34548040666 3.695 2007     2700      1     2.638
## 72619  34548172901 3.625 2007     2700      1     2.568
## 72623  34548262715 4.493 2007     2700      1     3.336
## 72673  34648836403 4.569 2007     2700      1     3.617
## 72681  34748862577 3.830 2007     2700      1     2.673
## 72807  36048990829 3.632 2007     2700      1     2.575
## 72822  36248973631 3.841 2007     2700      1     2.784
## 73355  34548221486 3.438 2007     1704      3     2.586
## 73492  34548066571 3.759 2007     2700      1     2.702
## 73493  34548075820 5.741 2007     2700      1     4.684
## 73494  34548083742 4.058 2007     2700      1     3.001
## 73500  34548092397 4.740 2007     2700      1     3.683
## 73551  34547950500 4.540 2007     2700      1     3.588
## 73558  34547876797 4.649 2007     2700      1     3.592
## 73559  34547906405 4.663 2007     2700      1     3.606
## 73563  34548361951 3.587 2007     2700      1     2.530
## 73594  34547757915 5.372 2007     2700      1     4.215
## 73599  34547730740 4.006 2007     2700      1     2.949
## 73600  34547755634 3.747 2007     2700      1     2.895
## 73649  34547647628 4.354 2007     2700      1     3.297
## 73681  34547614056 4.149 2007     2700      1     3.092
## 73683  34547630852 3.983 2007     2700      1     3.031
## 73684  34547634884 3.728 2007     2700      1     2.671
## 73685  34547643177 3.734 2007     2700      1     2.882
## 73686  34547652884 4.074 2007     2700      1     2.917
## 73697  34547681974 3.954 2007     2700      1     2.897
## 73699  34547686398 4.247 2007     2700      1     3.090
## 73747  34548397297 3.789 2007     2700      1     2.732
## 73760  34548456959 3.595 2007     2700      1     2.538
## 74193  34547148985 4.266 2007     2700      1     3.209
## 74194  34547170053 6.085 2007     2700      1     5.028
## 74199  34547230736 4.019 2007     2700      1     2.962
## 74203  34547159840 4.045 2007     2700      1     2.988
## 74204  34547167603 3.695 2007     2700      1     2.638
## 74226  34447521458 4.639 2007     2700      1     3.482
## 74241  34447544253 3.830 2007     2700      1     2.773
## 74246  34447300446 4.669 2007     2700      1     3.612
## 74247  34447512756 4.710 2007     2700      1     3.858
## 74249  34447520136 5.342 2007     2700      1     4.285
## 74250  34447558140 5.428 2007     2700      1     4.371
## 74312  34447340945 3.992 2007     2700      1     2.935
## 74314  34548165110 3.506 2007     2700      1     2.554
## 74322  34447257725 3.920 2007     2700      1     2.763
## 74323  34447291144 4.282 2007     2700      1     3.225
## 74369  34447116875 3.695 2007     2700      1     2.638
```

```
## 74375  34447121852 3.852 2007     2700      1     2.795
## 74385  34447137331 4.461 2007     2700      1     3.304
## 74398  34250840592 3.920 2007     2700      1     2.863
## 74704  35248881664 4.126 2007     2700      1     3.069
## 75040  34548567516 3.824 2007     2700      1     2.972
## 75044  34347210370 4.041 2007     2700      1     2.984
## 75045  34347237664 4.019 2007     2700      1     2.962
## 75047  34347255039 4.665 2007     2700      1     3.508
## 75057  34347272329 3.753 2007     2700      1     2.901
## 75085  34250696907 4.873 2007     2700      1     3.921
## 75086  34250722018 4.152 2007     2700      1     3.200
## 75089  34250779802 4.066 2007     2700      1     3.009
## 75092  34250835050 3.709 2007     1000      2     2.652
## 75142  34250212715 6.307 2007     2700      1     5.455
## 75144  34250309164 4.292 2007     2700      1     3.235
## 75168  34249899377 4.920 2007     2700      1     3.763
## 75183  34250024383 3.783 2007     2700      1     2.831
## 75202  34249982970 4.196 2007     2700      1     3.139
## 75203  34250027957 3.983 2007     2700      1     2.826
## 75236  34248149838 3.765 2007     2700      1     2.708
## 75239  34248169298 3.795 2007     2700      1     2.738
## 75242  34248201151 4.476 2007     2700      1     3.419
## 75243  34248206205 3.688 2007     2700      1     2.631
## 75256  34249037494 3.356 2007     2700      1     2.504
## 75260  34249662628 3.925 2007     2700      1     2.868
## 75262  34249863368 3.852 2007     2700      1     3.000
## 75274  34249913494 4.647 2007     2700      1     3.490
## 75275  34249915334 4.224 2007     2700      1     3.167
## 75382  34447648357 3.807 2007     2700      1     2.855
## 75793  34547309949 3.854 2007     2700      2     2.797
## 75956  34249302025 3.580 2007     2700      1     2.523
## 75959  34249803312 4.122 2007     2700      1     3.270
## 75961  34249830160 4.754 2007     2700      1     3.902
## 75973  34249027978 3.959 2007     2700      1     2.902
## 75990  34249657825 4.383 2007     2700      1     3.326
## 75993  34249673868 4.807 2007     2700      1     3.750
## 76036  34249047454 4.895 2007     2700      1     3.838
## 76042  34248551882 4.301 2007     2700      1     3.144
## 76043  34248559834 4.710 2007     2700      1     3.653
## 76044  34248564820 3.531 2007     2700      1     2.679
## 76093  34248399089 4.292 2007     2700      1     3.235
## 76121  34247615981 3.835 2007     2700      1     2.778
## 76133  34247862190 5.281 2007     2700      1     4.224
## 76164  34247268567 3.610 2007     2700      1     2.553
## 76199  34248184445 3.702 2007     2700      1     2.645
## 76200  34248185926 4.203 2007     2700      1     3.046
## 76214  34248214561 3.765 2007     2700      1     2.913
## 76218  34248227278 3.734 2007     2700      1     2.577
## 76835  34247863975 5.154 2007     2700      1     3.997
## 76849  34247476744 5.028 2007     2700      1     3.971
```

```
## 76864  34247470278 4.213 2007     2700      1     3.056
## 76865  34247498668 4.800 2007     2700      1     3.643
## 76883  34247167109 3.675 2007     2700      1     2.518
## 76903  34247241630 5.117 2007     2700      1     4.060
## 76945  34247118878 3.813 2007     2700      1     2.961
## 76946  34247144499 4.185 2007     2700      1     3.128
## 76947  34247149303 4.041 2007     2700      1     2.984
## 76950  34247523621 4.713 2007     2700      1     3.656
## 76959  34047096257 4.791 2007     2700      1     3.939
## 76970  34047243811 4.396 2007     2700      1     3.544
## 76972  34047268898 3.900 2007     2700      1     2.843
## 76981  34047237367 5.526 2007     2700      1     4.674
## 76988  34147096980 4.461 2007     2700      1     3.404
## 76993  34147124264 4.538 2007     2700      1     3.381
## 76995  34147139988 3.610 2007     2700      1     2.553
## 76998  34147153376 3.564 2007     2700      1     2.507
## 76999  34147153377 4.107 2007     2700      1     3.050
## 77015  33847667949 4.049 2007     2700      1     3.197
## 77028  33947671432 3.747 2007     2700      1     2.690
## 77077  34247476430 3.747 2007     2700      1     2.590
## 77081  34247498181 3.668 2007     2700      1     2.511
## 77084  34247545653 3.639 2007     2700      1     2.687
## 77570  34247282706 4.091 2007     2700      1     3.139
## 77574  34147113263 3.593 2007     1000      2     2.536
## 77582  34047185285 3.728 2007     2700      1     2.671
## 77604  33947210018 3.884 2007     2700      1     2.827
## 77607  33947595236 5.510 2007     2700      1     4.558
## 77643  33947519044 3.632 2007     2700      1     2.680
## 77661  33947196096 4.621 2007     2700      1     3.564
## 77664  33947279831 4.110 2007     2700      1     3.158
## 77685  33947271991 3.734 2007     2700      1     2.677
## 77712  33847705701 5.514 2007     2700      1     4.357
## 77720  33847723397 4.950 2007     2700      1     4.098
## 77730  33847393654 3.396 2007     2700      1     2.544
## 77752  33846926962 3.715 2007     2700      1     2.763
## 77761  33847253945 4.880 2007     2700      1     3.723
## 77762  33847267289 3.396 2007     2700      1     2.544
## 77772  33847368173 5.051 2007     2700      1     3.994
## 77773  33847369469 4.655 2007     2700      1     3.598
## 77776  33847381116 5.213 2007     2700      1     4.156
## 77777  33847391669 4.160 2007     2700      1     3.103
## 77779  33847404482 4.431 2007     2700      1     3.374
## 77780  33847406983 3.747 2007     2700      1     2.590
## 77781  33847407101 3.987 2007     2700      1     2.930
## 77786  33847413701 3.920 2007     2700      1     2.863
## 77819  33947356563 3.540 2007     2700      1     2.688
## 77839  33947710793 4.749 2007     2700      1     3.692
## 78473  34547138943 3.721 2007     2700      2     2.664
## 78607  33847353773 4.810 2007     2700      1     3.858
## 78631  33847692398 3.435 2007     2700      1     2.583
```

```
## 78651  33847165011 4.126 2007     2700      1     3.069
## 78654  33847103656 4.307 2007     2700      1     3.455
## 78655  33847174197 4.203 2007     2700      1     3.351
## 78704  33846964515 3.819 2007     2700      1     2.762
## 78707  33846996345 5.465 2007     2700      1     4.613
## 78728  33846857305 3.602 2007     2700      1     2.545
## 78737  33846866271 4.955 2007     2700      1     3.898
## 78738  33846875851 5.005 2007     2700      1     3.948
## 78746  33846863146 4.895 2007     2700      1     3.838
## 78749  33846878293 3.715 2007     2700      1     2.558
## 78780  33846571916 4.217 2007     2700      1     3.160
## 78792  33846075235 4.032 2007     2700      1     2.975
## 78810  33846661903 3.959 2007     2700      1     2.902
## 78812  33846673816 4.935 2007     2700      1     3.878
## 78813  33846688887 5.379 2007     2700      1     4.222
## 78816  33846783284 3.715 2007     2700      1     2.658
## 78821  33846787125 3.835 2007     2700      1     2.678
## 78822  33846794896 5.202 2007     2700      1     4.145
## 78825  33846810181 3.564 2007     2700      1     2.507
## 78832  33846817495 4.130 2007     2700      1     3.278
## 78836  33846829811 3.983 2007     2700      1     3.131
## 78839  33846839628 3.915 2007     2700      1     2.758
## 78843  33846850600 3.777 2007     2700      1     2.620
## 78847  33846913993 3.531 2007     2700      1     2.679
## 78899  33847610733 4.122 2007     2700      1     3.065
## 78900  33847617509 3.987 2007     2700      1     2.830
## 79061  34248581736 3.523 2007     2700      1     2.571
## 79359  33846524262 3.396 2007     2700      1     2.544
## 79377  33846294746 5.096 2007     2700      1     4.039
## 79386  33846413027 3.572 2007     2700      1     2.515
## 79416  33846126939 4.946 2007     2700      1     3.789
## 79420  33846175081 3.777 2007     2700      1     2.720
## 79427  33846125284 5.173 2007     2700      1     4.116
## 79428  33846129459 4.066 2007     2700      1     3.114
## 79489  33845970192 5.126 2007     2700      1     4.069
## 79491  33845972969 3.789 2007     2700      1     2.632
## 79492  33845979811 3.889 2007     2700      1     2.732
## 79495  33845994359 4.653 2007     2700      1     3.496
## 79496  33846006533 3.661 2007     2700      1     2.504
## 79498  33846012847 3.841 2007     2700      1     2.684
## 79499  33846017279 4.078 2007     2700      1     2.921
## 79510  33845984538 4.757 2007     2700      1     3.700
## 79511  33846021262 4.182 2007     2700      1     3.125
## 79512  33846026341 4.599 2007     2700      1     3.542
## 79513  33846027094 4.070 2007     2700      1     2.913
## 79540  33845870422 3.734 2007     2700      1     2.882
## 79560  33846037400 3.935 2007     2700      1     2.983
## 79602  33846668526 3.874 2007     2700      1     2.817
## 79607  33846671273 3.654 2007     2700      1     2.597
## 79609  33846672853 4.091 2007     2700      1     3.034
```

```
## 79671  33847184284 3.632 2007     2700      1     2.680
## 79978  34250801440 3.925 2007     2700      1     3.073
## 80011  34347336319 3.646 2007     2700      1     2.589
## 80013  34347343739 3.668 2007     2700      1     2.611
## 80106  35649016098 4.629 2007     2700      1     3.572
## 80111  35848953625 3.728 2007     2700      1     2.671
## 80112  35848954757 4.777 2007     2700      1     3.825
## 80217  42449096529 3.935 2007     2700      1     3.083
## 80636  58149083708 4.778 2008     2700      1     3.724
## 80643  58149218477 3.923 2008     2700      1     3.073
## 80683  58049213696 4.583 2008     2700      1     3.529
## 80684  58049216683 3.839 2008     2700      1     2.785
## 80728  57349192265 3.722 2008     2700      1     2.668
## 80744  57449112519 4.616 2008     2700      1     3.562
## 80797  57349093766 3.663 2008     2700      1     2.609
## 80814  57349133922 3.823 2008     2700      1     2.769
## 80816  57349184260 3.522 2008     2700      1     2.672
## 80943  56149095348 3.843 2008     2700      1     2.893
## 80980  57449084208 4.389 2008     2700      1     3.335
## 80983  57449095883 3.663 2008     2700      1     2.609
## 81062  58149387739 4.794 2008     2700      1     3.740
## 81349  64349102964 3.423 2008     2700      1     2.573
## 82714  57049128384 3.717 2008     2700      1     2.563
## 82785  56649112752 4.452 2008     2700      1     3.398
## 82845  56249128276 4.627 2008     2700      1     3.573
## 82846  56249141853 3.916 2008     2700      1     2.862
## 82880  55849115999 3.659 2008     1000      2     2.605
## 82899  55549095137 3.894 2008     2700      1     2.840
## 82911  55849084700 3.909 2008     2700      1     2.855
## 82920  55849139852 3.698 2008     2700      1     2.748
## 82950  54049103289 3.570 2008     2700      1     2.516
## 83109  58149467012 3.754 2008     2700      1     2.700
## 83609  55249119765 4.514 2008     2700      1     3.460
## 83619  54349091094 3.478 2008     2700      1     2.528
## 83648  55549139481 4.069 2008     2700      1     3.015
## 83661  54949144335 4.216 2008     2700      1     3.162
## 83714  54049121079 3.886 2008     2700      1     2.832
## 83728  54049104424 3.768 2008     2700      1     2.818
## 83729  54049106963 4.294 2008     2700      1     3.240
## 83794  53749095835 4.145 2008     2700      1     3.091
## 83849  54849146500 4.137 2008     2700      1     3.083
## 83859  54949141082 3.589 2008     1704      3     2.535
## 83861  50249096174 4.078 2008     2700      1     3.128
## 83872  51749084267 3.451 2008     2700      1     2.601
## 83891  52949103244 3.781 2008     2700      1     2.831
## 83892  52949106040 4.278 2008     2700      1     3.224
## 83897  52949144507 4.137 2008     2700      1     3.083
## 83898  52949144863 3.722 2008     2700      1     2.772
## 83901  53149119591 3.401 2008     2700      1     2.551
## 83969  55849112296 3.599 2008     2700      1     2.649
```

```
## 83975  55849150592 4.342 2008     2700      1     3.288
## 84213  65549087532 3.839 2008     2700      1     2.889
## 84545  52649151888 3.916 2008     2700      1     2.966
## 84550  52049088955 4.048 2008     2700      1     2.894
## 84553  52049108262 4.242 2008     2700      1     3.188
## 84555  52449113766 3.626 2008     2700      1     2.776
## 84585  52649120886 4.627 2008     2700      1     3.473
## 84598  52249089796 3.843 2008     2700      1     2.789
## 84607  51949095058 4.868 2008     2700      1     3.918
## 84682  51449097277 4.098 2008     2700      1     2.944
## 84728  50949089029 4.489 2008     2700      1     3.435
## 84734  51649088233 4.453 2008     2700      1     3.399
## 84761  51349169034 3.673 2008     2700      1     2.619
## 84831  52749095854 3.593 2008     2700      1     2.539
## 85169  59049103004 3.503 2008     2700      1     2.653
## 85277  65249106599 3.570 2008     2700      1     2.620
## 85688  50149098605 3.746 2008     1000      2     2.692
## 85697  50449091665 3.772 2008     2700      1     2.618
## 85698  50449093260 4.360 2008     2700      1     3.306
## 85709  50449095362 4.693 2008     2700      1     3.639
## 85728  49749143019 3.576 2008     2700      1     2.726
## 85743  50149121231 4.030 2008     1000      2     2.976
## 85752  49949087905 4.118 2008     2700      1     3.168
## 85753  49949091204 3.678 2008     2700      1     2.624
## 85765  49449098216 3.958 2008     2700      1     3.008
## 85828  49449100639 4.042 2008     2700      1     2.988
## 85831  49249087104 4.356 2008     2700      1     3.406
## 85886  49249104701 5.150 2008     2700      1     4.096
## 85962  48149090717 3.855 2008     2700      1     2.801
## 85972  48149110000 3.647 2008     2700      1     2.797
## 86469  47849130092 4.213 2008     2700      1     3.159
## 86470  47849130608 3.731 2008     2700      1     2.677
## 86483  47849100908 3.663 2008     2700      1     2.509
## 86505  47949092540 4.171 2008     2700      1     3.221
## 86507  47949114668 5.232 2008     2700      1     4.078
## 86564  47549096598 3.768 2008     2700      1     2.818
## 86565  47549105407 4.511 2008     2700      1     3.357
## 86620  46949093558 3.886 2008     2700      1     2.732
## 86634  47049126504 3.827 2008     2700      1     2.773
## 86680  46449113799 4.877 2008     2700      1     3.723
## 86684  46449128469 3.546 2008     1000      2     2.596
## 86729  45749146166 3.855 2008     2700      1     3.005
## 86742  46449098821 3.879 2008     2700      1     2.929
## 86743  46749084688 3.843 2008     2700      1     2.789
## 86746  46749099010 4.045 2008     2700      1     3.195
## 86749  46749117203 3.712 2008     2700      1     2.658
## 86755  46749156706 4.460 2008     2700      1     3.306
## 86780  47949131252 3.387 2008     2700      1     2.537
## 86795  48749129542 4.693 2008     2700      1     3.743
## 87045  55149104425 3.768 2008     2700      1     2.818
```

```
## 87540  45949092534 4.637 2008     2700      1     3.583
## 87543  45949107235 3.831 2008     2700      1     2.777
## 87544  46749131798 5.204 2008     2700      1     4.150
## 87592  45549084293 4.115 2008     2700      1     3.061
## 87602  45349084317 4.501 2008     2700      1     3.447
## 87603  45349103702 4.595 2008     2700      1     3.441
## 87663  45149127545 4.115 2008     2700      1     3.061
## 87691  44849103815 4.757 2008     2700      1     3.907
## 87696  44849115945 4.501 2008     2700      1     3.447
## 87735  45749100010 3.951 2008     2700      1     2.897
## 87739  45749114895 4.710 2008     2700      1     3.556
## 87740  45749124735 3.409 2008     2700      1     2.559
## 87745  45749149476 3.642 2008     2700      1     2.588
## 87792  44949119632 3.451 2008     2700      1     2.601
## 87828  46349104324 3.831 2008     2700      1     2.777
## 88491  44349145194 4.455 2008     2700      1     3.401
## 88496  44449106055 3.641 2008     1000      2     2.587
## 88499  44349153009 3.978 2008     2700      1     2.924
## 88500  44349155992 4.235 2008     2700      1     3.285
## 88501  44349165737 5.327 2008     2700      1     4.377
## 88546  44249122799 5.169 2008     2700      1     4.115
## 88634  43549085056 3.948 2008     2700      1     3.098
## 88637  43549110959 4.509 2008     2700      1     3.355
## 88680  43049170245 4.574 2008     2700      1     3.624
## 88682  43049174930 3.968 2008     2700      1     2.814
## 88683  43049176549 4.334 2008     2700      1     3.384
## 88750  42949149810 3.901 2008     2700      1     3.051
## 88757  43049090817 4.174 2008     2700      1     3.120
## 88759  43049094836 3.740 2008     2700      1     2.790
## 88760  43049095682 4.414 2008     2700      1     3.360
## 88784  43049136576 3.570 2008     2700      1     2.516
## 88857  44449096577 3.731 2008     2700      1     2.677
## 88859  44449122269 3.944 2008     2700      1     3.094
## 89498  42449106956 4.533 2008     2700      1     3.683
## 89500  42449141936 3.802 2008     2700      1     2.648
## 89501  44149123099 4.666 2008     2700      1     3.816
## 89566  42249112261 3.621 2008     1000      2     2.567
## 89567  42249098105 4.181 2008     2700      1     3.127
## 89601  42249083269 4.686 2008     2700      1     3.836
## 89602  42249092533 4.619 2008     2700      1     3.769
## 89605  48149094782 3.698 2008     2700      1     2.544
## 89617  42049123233 4.048 2008     2700      1     3.198
## 89645  40849108663 3.574 2008     1000      2     2.520
## 89663  41649112174 3.663 2008     2700      1     2.509
## 89739  41849105806 3.930 2008     2700      1     2.776
## 89751  41849148289 3.992 2008     2700      1     2.938
## 89783  43249086840 3.683 2008     2700      1     2.529
## 89784  43249087616 3.859 2008     2700      1     2.805
## 89878  46349092759 3.712 2008     2700      1     2.558
## 90277  43049134750 3.647 2008     2700      1     2.593
```

```
## 90284  41449104685 3.576 2008     2700      1     2.522
## 90287  41449112372 5.371 2008     2700      1     4.317
## 90289  41449113183 4.795 2008     2700      1     3.741
## 90338  40949100835 3.944 2008     2700      1     2.994
## 90347  40949102607 5.086 2008     2700      1     3.932
## 90361  40949127393 4.633 2008     2700      1     3.783
## 90407  40849124009 4.384 2008     2700      1     3.330
## 90409  40849142102 4.916 2008     2700      1     3.762
## 90412  40749132230 3.356 2008     1000      2     2.506
## 90415  40749141596 4.476 2008     2700      1     3.526
## 90432  41649090186 4.800 2008     2700      1     3.746
## 90458  40449141326 4.620 2008     2700      1     3.566
## 90464  40449104676 4.174 2008     2700      1     3.324
## 90509  39749103691 3.615 2008     2700      1     2.765
## 90521  40449086087 4.445 2008     2700      1     3.291
## 90523  40549084318 3.879 2008     2700      1     2.825
## 90545  40749105777 3.541 2008     2700      1     2.691
## 91382  39549099909 3.693 2008     2700      1     2.639
## 91383  39549100109 3.988 2008     2700      1     2.934
## 91384  39549102834 4.683 2008     2700      1     3.629
## 91386  39549111874 3.958 2008     2700      1     2.904
## 91388  39549096658 4.328 2008     2700      1     3.174
## 91415  39749102149 4.416 2008     2700      1     3.566
## 91426  39349097864 3.575 2008     1000      2     2.521
## 91488  38949105879 4.256 2008     2700      1     3.306
## 91489  38949106566 3.754 2008     2700      1     2.700
## 91490  38949188680 3.871 2008     2700      1     2.817
## 91499  38949152338 4.285 2008     2700      1     3.231
## 91520  38649125875 4.145 2008     2700      1     3.091
## 91534  38649100024 3.604 2008     2700      1     2.754
## 91552  38849130851 3.827 2008     2700      1     2.773
## 91565  38849174979 4.368 2008     2700      1     3.418
## 91567  38849182472 3.972 2008     2700      1     2.918
## 91602  39849097577 4.313 2008     2700      1     3.159
## 91603  39849099704 3.516 2008     2700      1     2.666
## 91605  39849102541 3.615 2008     2700      1     2.665
## 91725  43349085680 3.955 2008     2700      1     2.901
## 91726  43349087466 3.772 2008     2700      1     2.718
## 91730  43349096215 3.707 2008     2700      1     2.653
## 91738  43349107306 3.975 2008     2700      1     2.921
## 91845  38949168818 3.621 2008     1300      2     2.771
## 91993  38749111066 5.445 2008     2700      1     4.391
## 91995  38749127156 4.616 2008     2700      1     3.562
## 92001  38749133726 3.560 2008     1000      2     2.506
## 92033  38549088096 4.842 2008     2700      1     3.788
## 92037  38549134052 4.324 2008     2700      1     3.270
## 92040  38549177642 4.453 2008     2700      1     3.399
## 92046  38349164176 4.531 2008     2700      1     3.477
## 92068  38349049478 5.357 2008     2700      1     4.303
## 92111  38049146378 4.148 2008     2700      1     3.094
```

```
## 92113  38049169559 4.935 2008     2700      1     3.881
## 92134  38049062200 4.629 2008     2700      1     3.575
## 92136  38049077991 3.843 2008     2700      1     2.789
## 92140  38049082149 4.298 2008     2700      1     3.244
## 92148  38149064606 3.631 2008     2700      1     2.681
## 92161  38149141767 3.958 2008     2700      1     2.804
## 92193  37549009899 4.066 2008     2700      1     3.116
## 92197  37549020305 3.582 2008     2700      1     2.528
## 92199  37549027612 4.244 2008     2700      1     3.190
## 92203  37549064327 3.776 2008     2700      1     2.722
## 92219  38149074110 3.934 2008     2700      1     2.880
## 92231  38349085066 3.610 2008     2700      1     2.556
## 92494  41149157013 3.547 2008     2700      1     2.597
## 92576  42149177179 4.216 2008     2700      1     3.162
## 92681  44949253343 4.140 2008     2700      1     2.986
## 92735  46749123025 3.731 2008     2700      1     2.677
## 92922  57149102701 3.912 2008     2700      1     3.062
## 93538  74049108635 3.752 2009     2700      1     2.653
## 93541  73949086506 3.919 2009     2700      1     2.721
## 93592  73549083853 3.805 2009     2700      1     2.706
## 93602  77954601328 4.143 2009     2700      1     3.044
## 93632  70350733579 3.757 2009     2700      1     2.559
## 93633  70449636557 4.842 2009     2700      1     3.644
## 93634  71849097187 3.961 2009     2700      1     2.763
## 93635  71849098875 3.947 2009     2700      1     2.848
## 93652  70350733425 3.985 2009     2700      1     2.991
## 93654  70350738289 4.729 2009     2700      1     3.630
## 93659  72549091935 4.039 2009     2700      1     3.145
## 93707  71849118976 3.968 2009     2700      1     2.869
## 93720  71549150905 3.500 2009     2700      1     2.606
## 93760  73349094732 4.184 2009     2700      1     3.190
## 93764  73349134686 4.036 2009     2700      1     2.937
## 93775  84870055817 4.573 2009     2700      1     3.679
## 93843  65549167028 3.801 2009     2700      1     2.702
## 93889  67650760521 3.826 2009     2700      1     2.727
## 93962  70749114018 3.645 2009     2700      1     2.546
## 93964  70749145982 3.814 2009     2700      1     2.715
## 93975  72049129932 3.889 2009     2700      1     2.790
## 93995  72849125358 3.858 2009     2700      1     2.864
## 94224  77049105843 3.904 2009     2700      1     2.910
## 94281  77649213935 3.838 2009     2700      1     2.844
## 95604  72049129430 4.683 2009     2700      1     3.584
## 95638  70449652591 3.805 2009     2700      1     2.706
## 95723  70349873173 5.082 2009     2700      1     4.088
## 95726  70449636163 5.302 2009     2700      1     4.308
## 95743  77950669604 5.246 2009     2700      1     4.147
## 95828  70449355084 3.784 2009     2700      1     2.685
## 95831  70449377960 3.743 2009     2700      1     2.644
## 95834  70449394580 4.039 2009     2700      1     2.940
## 95842  70449448312 4.397 2009     2700      1     3.298
```

```
## 96014  73949153085 4.118 2009     2700      1     3.019
## 96406  73349109845 3.667 2009     2700      2     2.568
## 96572  84881264341 4.415 2009     2700      1     3.316
## 96673  70350064091 3.705 2009     2700      1     2.711
## 96696  70350458703 4.870 2009     2700      1     3.876
## 96716  70350212398 4.751 2009     2700      1     3.652
## 96717  70350220313 3.629 2009     2700      1     2.530
## 96718  70350227015 3.975 2009     2700      1     2.777
## 96719  70350236228 4.659 2009     2700      1     3.765
## 96753  70349506141 4.798 2009     2700      1     3.904
## 96773  70349943834 3.625 2009     1000      2     2.526
## 96854  65749111503 3.762 2009     2700      1     2.663
## 96867  69949132983 3.549 2009     2700      1     2.655
## 96870  69949144253 3.696 2009     2700      1     2.597
## 96895  70349481938 4.575 2009     2700      1     3.476
## 96897  70349610473 4.853 2009     2700      1     3.754
## 96899  70349625757 5.493 2009     2700      1     4.394
## 96900  70349629019 5.086 2009     2700      1     3.987
## 96919  70349684791 4.187 2009     2700      1     2.989
## 96920  70349687102 4.017 2009     2700      1     2.918
## 96953  70350455774 3.691 2009     2700      1     2.697
## 97145  77952988108 4.600 2009     2700      1     3.501
## 97523  70349437154 3.951 2009     2700      1     2.852
## 97545  70349308645 3.601 2009     2700      1     2.502
## 97546  70349308646 4.248 2009     2700      1     3.149
## 97548  70349319816 3.705 2009     2700      1     2.711
## 97550  70349335796 4.765 2009     2700      1     3.666
## 97633  70349238733 4.938 2009     2700      1     4.044
## 97634  70349244812 4.857 2009     2700      1     3.963
## 97636  70349254065 3.850 2009     2700      1     2.751
## 97666  66849121699 3.710 2009     2700      1     2.611
## 97668  66849128097 3.947 2009     2700      1     2.848
## 97740  70149093912 5.343 2009     2700      1     4.145
## 97741  70149098511 3.748 2009     2700      1     2.649
## 97748  70249111091 3.839 2009     1000      2     2.641
## 97755  70149112350 4.282 2009     2700      1     3.084
## 97852  68949150765 3.725 2009     2700      1     2.731
## 97861  69849097352 3.958 2009     2700      1     2.859
## 97897  70349205537 3.506 2009     2700      1     2.512
## 97906  70349213381 3.691 2009     2700      1     2.797
## 97912  70349220917 3.757 2009     2700      1     2.658
## 97930  70349636711 4.014 2009     2700      1     2.915
## 98067  73449108112 3.846 2009     2700      1     2.852
## 98073  73449123799 4.061 2009     2700      1     2.962
## 98210  75349114036 3.766 2009     2700      1     2.667
## 98756  69249219296 5.111 2009     2700      1     4.012
## 98759  69249229866 3.850 2009     2700      1     2.751
## 98778  69449098588 4.214 2009     2700      1     3.115
## 98813  68949090752 4.850 2009     2700      1     3.652
## 98814  68949092632 4.197 2009     2700      1     3.303
```

```
## 98815  68949102467 3.830 2009     2700      1     2.731
## 98816  68949133346 4.250 2009     2700      1     3.052
## 98817  68949171916 3.775 2009     2700      1     2.881
## 98921  68849084949 4.231 2009     2700      1     3.033
## 98930  68849119553 4.573 2009     2700      1     3.474
## 98931  68849123400 4.335 2009     2700      1     3.236
## 98992  68249154875 5.117 2009     2700      1     4.018
## 99008  68349117223 3.919 2009     2700      1     3.025
## 99032  68849104003 3.613 2009     2700      1     2.514
## 99033  68849112640 3.607 2009     2700      1     2.508
## 99081  67650635227 3.991 2009     2700      1     2.793
## 99104  67651160419 4.452 2009     2700      1     3.558
## 99146  69349094143 4.819 2009     2700      1     3.720
## 99564  68049123594 4.967 2009     2700      1     3.868
## 99616  68149144251 3.826 2009     2700      1     2.727
## 99644  67651218924 4.011 2009     2700      1     2.912
## 99664  67650391945 3.818 2009     2700      1     2.620
## 99679  67650318719 3.926 2009     2700      1     2.728
## 99686  67650899270 4.135 2009     2700      1     3.036
## 99688  67650909125 3.968 2009     2700      1     2.869
## 99748  68149112408 3.440 2009     2700      1     2.546
## 99827  67649664439 3.975 2009     2700      1     3.081
## 99858  67649814915 3.743 2009     2700      1     2.545
## 99874  67649967683 4.202 2009     2700      1     3.308
## 99913  68049122014 3.725 2009     2700      1     2.831
## 99961  68849122050 3.988 2009     2700      1     2.889
## 99962  68849124811 4.132 2009     2700      1     3.238
## 100692 67649390890 3.995 2009     2700      1     2.896
## 100705 67649218781 4.174 2009     2700      1     2.976
## 100727 67149116154 4.250 2009     2700      1     3.052
## 100761 67449110743 5.734 2009     2700      1     4.536
## 100762 67449152290 4.549 2009     2700      1     3.655
## 100764 67449164373 3.944 2009     2700      1     2.950
## 100831 67149142042 5.168 2009     2700      1     4.069
## 100844 67049116139 3.882 2009     2700      1     2.783
## 100938 66649113371 4.055 2009     2700      1     2.857
## 100947 66449100700 4.172 2009     2700      1     2.974
## 101089 67651005390 3.748 2009     2700      1     2.754
## 101093 67651006589 3.433 2009     2700      1     2.539
## 101786 67650169237 4.101 2009     2700      1     3.002
## 101788 65849187703 5.312 2009     2700      1     4.213
## 101794 66249084112 4.991 2009     2700      1     3.892
## 101808 66149152644 4.179 2009     2700      1     2.981
## 101854 66149118225 3.814 2009     2700      1     2.715
## 101861 66249137759 3.789 2009     1000      2     2.690
## 101904 66149123748 3.968 2009     1000      2     2.869
## 101942 65549163328 4.148 2009     2700      1     3.154
## 101994 65449136284 3.759 2009     1303      7     2.660
## 102096 66449111715 3.766 2009     2700      1     2.667
## 102501 65549092261 3.656 2009     1000      2     2.557
```

```
## 102867 65449152185 5.067 2009     2700      1     3.968
## 102988 64749093541 4.983 2009     2700      1     3.884
## 103006 64749086505 4.036 2009     2700      1     2.937
## 103008 64749114159 4.724 2009     2700      1     3.526
## 103061 62549134121 3.909 2009     1000      2     2.810
## 103846 63249124498 3.838 2009     2700      1     2.739
## 103851 63349094964 3.846 2009     2700      1     2.952
## 103903 62749097289 4.481 2009     2700      1     3.587
## 103904 62749131642 4.113 2009     2700      1     2.915
## 103921 62649136134 4.286 2009     2700      1     3.292
## 103922 62649155822 4.030 2009     2700      1     3.036
## 103984 61849137328 3.830 2009     2700      1     2.731
## 103998 62349094336 3.870 2009     2700      1     2.771
## 104003 62249133709 4.143 2009     1000      2     3.044
## 104012 62149120632 4.393 2009     2700      1     3.399
## 104068 61849118968 4.245 2009     2700      1     3.251
## 104126 60049098815 4.857 2009     2700      1     3.758
## 104132 61549094077 5.418 2009     2700      1     4.524
## 104148 62449262646 3.640 2009     2700      1     2.646
## 105045 60649118584 4.654 2009     2700      1     3.660
## 105055 60649090430 4.316 2009     2700      1     3.217
## 105061 61449183682 5.109 2009     2700      1     4.010
## 105074 61449162171 3.752 2009     2700      1     2.653
## 105118 60849115270 5.804 2009     2700      1     4.705
## 105126 60749089639 3.858 2009     2700      1     2.759
## 105140 59849099670 3.904 2009     2700      1     3.010
## 105210 60249096283 3.797 2009     2700      1     2.803
## 105211 60249099307 4.478 2009     2700      1     3.280
## 105240 59249084303 4.172 2009     2700      1     3.073
## 105242 59449101337 4.234 2009     2700      1     3.036
## 105251 59749096464 4.506 2009     2700      1     3.407
## 105271 58949085856 3.525 2009     2700      1     2.631
## 105282 61749094289 4.261 2009     2700      1     3.063
## 105283 61749094852 3.775 2009     2700      1     2.577
## 105852 59449094890 4.255 2009     2700      1     3.261
## 105884 58749094444 5.536 2009     2700      1     4.338
## 105903 58749109459 3.607 2009     2700      1     2.613
## 105947 58349091054 3.908 2009     2700      1     2.710
## 105956 58249110796 3.725 2009     1000      2     2.527
## 105970 58249088833 3.991 2009     2700      1     2.793
## 105990 58149385407 4.564 2009     2700      1     3.465
## 105992 58149389215 5.900 2009     2700      1     4.801
## 106003 58149385649 4.940 2009     2700      1     3.742
## 106031 61749090233 4.005 2009     2700      1     2.906
## 106037 61749102618 4.145 2009     2700      1     3.046
## 106063 57749092004 3.743 2009     2700      1     2.644
## 106066 57749093663 3.623 2009     2700      1     2.524
## 106107 58249087710 3.985 2009     2700      1     2.886
## 106187 61449122079 3.671 2009     2700      1     2.777
## 106195 61449149483 4.101 2009     2700      1     3.002
```

```
## 107838 78650858488 4.531 2010     2700      1     3.388
## 107864 78650664067 3.685 2010     2700      1     2.542
## 107918 78649906060 4.743 2010     2700      1     3.600
## 108011 78649880804 4.124 2010     2700      1     2.981
## 108234 78649822277 3.520 2010     2700      1     2.581
## 108250 78649888666 3.441 2010     2700      1     2.502
## 108252 78649895980 3.904 2010     2700      1     2.761
## 109519 78649429727 4.440 2010     2700      1     3.297
## 109522 78649439268 4.337 2010     2700      1     3.194
## 109593 78549288927 3.760 2010     2700      1     2.821
## 109846 78149330579 3.601 2010     2700      1     2.563
## 109891 78049510428 4.327 2010     2700      1     3.084
## 109892 78049515807 4.522 2010     2700      1     3.379
## 109919 78049407029 4.053 2010     2700      1     2.910
## 110787 77957768543 4.019 2010     2700      1     2.876
## 110790 78049428879 4.101 2010     2700      1     2.958
## 110803 78049337151 3.858 2010     2700      1     2.715
## 110805 78049351929 4.511 2010     2700      1     3.368
## 110864 77958149840 4.229 2010     2700      1     2.986
## 111024 77957934893 3.955 2010     2700      1     2.712
## 111045 77957928301 3.775 2010     2700      1     2.632
## 111115 77957725001 3.841 2010     2700      1     2.698
## 111130 77957671700 3.646 2010     2700      1     2.503
## 111485 80051785529 3.594 2010     2700      1     2.556
## 112042 77957333252 3.808 2010     2700      1     2.665
## 112043 77957341503 4.430 2010     2700      1     3.187
## 112124 77957139539 3.680 2010     1000      2     2.537
## 112173 77956929653 3.485 2010     2700      1     2.546
## 112174 77956942602 3.749 2010     2700      1     2.506
## 112259 77956634601 3.549 2010     2700      1     2.610
## 113861 77955877759 5.447 2010     2700      1     4.204
## 113873 77955760684 3.730 2010     2700      1     2.587
## 114067 77955357246 3.858 2010     2700      1     2.715
## 114203 77955296243 3.591 2010     2700      1     2.553
## 114219 77955286305 3.791 2010     2700      1     2.648
## 114324 77955573674 4.157 2010     2700      1     3.218
## 114349 77956057548 3.989 2010     2700      1     2.846
## 115221 77955066199 5.339 2010     2700      1     4.196
## 115230 77954988618 4.006 2010     2700      1     2.863
## 115361 77954635020 4.231 2010     2700      1     3.088
## 115455 77954421915 4.566 2010     2700      1     3.423
## 115465 77954374543 4.143 2010     2700      1     2.900
## 115552 77954183132 3.685 2010     2700      1     2.542
## 115604 77954974268 3.717 2010     2700      1     2.574
## 116866 77953956205 4.405 2010     2700      1     3.367
## 117020 77953577951 4.518 2010     2700      1     3.375
## 117076 77952962769 3.929 2010     2700      1     2.686
## 117084 77952977625 3.827 2010     2700      1     2.684
## 117093 77952998648 3.957 2010     2700      1     3.018
## 117101 77953013128 3.801 2010     2700      1     2.658
```

```
## 117127 77953173228 4.019 2010     2700      1     2.981
## 117148 77952995935 3.649 2010     2700      1     2.506
## 117177 77952766987 3.851 2010     2700      1     2.708
## 117245 77953360245 3.933 2010     2700      1     2.994
## 118168 77952757705 4.297 2010     2700      1     3.358
## 118169 77952760196 3.757 2010     2700      1     2.514
## 118172 77952784840 4.247 2010     2700      1     3.308
## 118242 77952470787 3.736 2010     2700      1     2.797
## 118247 77952356171 3.714 2010     2700      1     2.571
## 118249 77952389110 3.851 2010     2700      1     2.708
## 118307 77952301687 4.418 2010     2700      1     3.275
## 118351 77951868945 4.370 2010     2700      1     3.227
## 118352 77951874018 4.849 2010     2700      1     3.910
## 118367 77951894040 4.247 2010     2700      1     3.308
## 118392 77951589703 4.220 2010     2700      1     3.077
## 118410 77951861324 3.897 2010     2700      1     2.754
## 119263 77951639385 4.329 2010     2700      1     3.086
## 119325 77951179995 4.040 2010     2700      1     3.101
## 119326 77951195763 4.197 2010     2700      1     3.159
## 119394 77950894766 3.527 2010     2700      1     2.588
## 119487 77950847452 3.722 2010     2700      1     2.579
## 119497 77950477601 4.327 2010     2700      1     3.184
## 119499 77950486825 3.875 2010     2700      1     2.732
## 119582 77950475726 4.241 2010     2700      1     3.098
## 119583 77950478147 4.036 2010     2700      1     3.097
## 119585 77950494853 4.319 2010     2700      1     3.076
## 120218 77950246404 4.319 2010     2700      1     3.176
## 120282 77949511388 3.778 2010     2700      1     2.635
## 120453 77649329576 4.566 2010     2700      1     3.323
## 120619 77950355514 3.621 2010     2700      1     2.682
## 121458 77249162677 4.349 2010     2700      1     3.206
## 121459 77249165024 3.722 2010     2700      1     2.579
## 121517 77149173502 3.864 2010     2700      1     2.826
## 121526 77149120471 3.773 2010     2700      1     2.530
## 121527 77149136099 3.803 2010     2700      1     2.560
## 121596 76649086071 4.375 2010     2700      1     3.436
## 122262 74549178560 4.208 2010     1315      2     3.065
## 122299 74849098405 4.260 2010     2700      1     3.321
## 122328 74549140993 5.722 2010     2700      1     4.479
## 122329 74549159752 5.207 2010     2700      1     4.169
## 122330 74549184608 3.581 2010     2700      1     2.642
## 122358 74949117631 4.260 2010     2700      1     3.222
## 122403 74149095074 3.682 2010     2700      1     2.743
## 122451 74049098673 4.147 2010     2700      1     3.109
## 122453 74049114688 4.434 2010     2700      1     3.291
## 122489 74949089659 3.781 2010     2700      1     2.638
## 122867 77950651061 4.285 2010     2700      1     3.346
## 122994 77952974496 3.813 2010     2700      1     2.670
## 123000 77952987107 3.798 2010     2700      1     2.655
## 123134 77955050878 3.858 2010     2700      1     2.615
```

```
## 125455 84855370035 4.953 2011     2700      1     3.917
## 125459 84855425106 5.298 2011     2700      1     4.361
## 125468 84555187324 3.856 2011     2700      1     2.615
## 125541 84255194001 4.947 2011     2700      1     3.806
## 125573 83755178774 3.993 2011     2700      1     2.957
## 125605 83355163991 3.976 2011     2700      1     2.940
## 125606 83455210809 3.567 2011     2700      1     2.630
## 125608 83455235074 3.980 2011     2700      1     3.043
## 125659 83255182742 3.895 2011     2700      1     2.958
## 125678 83155182848 4.324 2011     2700      1     3.183
## 125681 83055177179 3.952 2011     1000      2     2.811
## 125689 83055168450 3.885 2011     2700      1     2.948
## 125690 83055173186 4.057 2011     2700      1     2.916
## 125692 83155168389 3.830 2011     2700      1     2.893
## 125693 83155177076 3.989 2011     2700      1     2.953
## 125777 79960457303 3.482 2011     2700      1     2.545
## 125944 82555168440 4.771 2011     2700      1     3.630
## 125993 84055200171 3.731 2011     2700      1     2.590
## 125994 84055217017 3.890 2011     2700      1     2.749
## 125995 84055217292 3.856 2011     2700      1     2.715
## 127205 82255175382 5.249 2011     2700      1     4.108
## 127216 81855218115 4.488 2011     2700      1     3.347
## 127225 81855161048 3.950 2011     2700      1     3.013
## 127277 81555213207 3.799 2011     2700      1     2.763
## 127287 81455154874 4.549 2011     2700      1     3.408
## 127289 81455159320 4.127 2011     2700      1     2.986
## 127302 81255208675 4.283 2011     2700      1     3.142
## 127362 80755130317 4.428 2011     2700      1     3.491
## 127364 80755188166 4.426 2011     2700      1     3.285
## 127437 80455162464 4.949 2011     2700      1     3.808
## 127457 80355135138 4.850 2011     2700      1     3.913
## 127458 80355148407 4.102 2011     2700      1     2.861
## 127461 80455179713 3.713 2011     2700      1     2.572
## 127572 80555136815 4.124 2011     2700      1     3.088
## 128312 80054769456 4.528 2011     2700      1     3.387
## 128352 80054714783 3.880 2011     2700      1     2.739
## 128374 80054003232 4.343 2011     2700      1     3.202
## 128462 80053469106 4.076 2011     2700      1     3.139
## 128473 80053510713 3.861 2011     2700      1     2.620
## 128565 80053402552 3.851 2011     2700      1     2.610
## 128577 80053426842 4.065 2011     2700      1     2.824
## 129148 80053359692 4.030 2011     2700      1     2.889
## 129173 84555215993 3.637 2011     2700      1     2.601
## 129220 80053069652 4.106 2011     2700      1     2.965
## 129236 80053065502 3.945 2011     2700      1     2.804
## 129326 80052834654 5.140 2011     2700      1     3.999
## 129352 80052792378 4.382 2011     2700      1     3.241
## 129501 80052462089 3.945 2011     2700      1     2.804
## 129504 80052515621 3.713 2011     2700      1     2.572
## 129538 80052482300 3.725 2011     2700      1     2.584
```

```
## 129610 80052371597 3.972 2011     2700      1     2.831
## 129611 80052376347 3.895 2011     2700      1     2.754
## 129619 80052436335 4.001 2011     2700      1     2.860
## 130467 80052358982 3.650 2011     2700      1     2.713
## 130507 80051720194 5.682 2011     2700      1     4.541
## 130522 80052047215 4.316 2011     2700      1     3.175
## 130523 80052063328 3.904 2011     2700      1     2.763
## 130603 84860389821 4.428 2011     2700      1     3.287
## 130622 80051723862 4.452 2011     2700      1     3.311
## 130627 80051831092 4.268 2011     2700      1     3.127
## 130629 84867338003 4.503 2011     2700      1     3.262
## 130677 80051640525 4.042 2011     2700      1     2.901
## 130688 80051571901 4.590 2011     2700      1     3.554
## 130690 80051775476 5.397 2011     2700      1     4.256
## 130735 80051468836 4.057 2011     2700      1     2.916
## 130778 79961102782 3.743 2011     2700      1     2.707
## 130779 79961108629 6.254 2011     2700      1     5.218
## 130780 79961116424 4.087 2011     2700      1     2.946
## 130798 79961040968 3.804 2011     2700      1     2.663
## 130883 79961006281 3.941 2011     2700      1     2.700
## 131346 79960855656 4.776 2011     2700      1     3.635
## 131363 79960704300 3.825 2011     2700      1     2.584
## 131364 79960709532 3.890 2011     2700      1     2.749
## 131435 79960575079 4.596 2011     2700      1     3.455
## 131438 79960611010 3.670 2011     2700      1     2.529
## 131494 79960230129 3.950 2011     2700      1     2.809
## 131599 79960056173 4.241 2011     2700      1     3.100
## 131600 79960066694 3.552 2011     2700      1     2.516
## 131724 79959997891 3.895 2011     2700      1     2.654
## 131727 79960015327 4.590 2011     2700      1     3.554
## 131728 79960015997 5.714 2011     2700      1     4.777
## 131731 79960027556 4.472 2011     2700      1     3.436
## 132437 79959794787 5.141 2011     2700      1     4.000
## 132474 79959772357 4.719 2011     2700      1     3.578
## 132530 79959478455 5.314 2011     2700      1     4.173
## 132531 79959478969 4.354 2011     2700      1     3.213
## 132542 79959392796 3.748 2011     2700      1     2.507
## 132545 79959426992 4.106 2011     2700      1     2.965
## 132546 79959487586 4.565 2011     2700      1     3.424
## 132606 79959363092 3.574 2011     1000      2     2.637
## 132615 79959297709 3.871 2011     2700      1     2.835
## 132662 79958724181 4.832 2011     2700      1     3.591
## 132668 79959316645 3.976 2011     2700      1     2.835
## 132736 79958195203 4.483 2011     2700      1     3.342
## 132738 79958226299 4.639 2011     2700      1     3.498
## 132745 79958066836 4.674 2011     2700      1     3.533
## 132790 79957795474 4.944 2011     2700      1     3.803
## 132794 79957877087 4.014 2011     2700      1     3.077
## 132801 79952593642 3.682 2011     2700      1     2.541
## 132832 79957520670 3.959 2011     2700      1     2.718
```

```
## 132838 79957805100 4.354 2011     2700      1     3.113
## 132839 79957808383 4.120 2011     2700      1     2.979
## 132850 79957889288 3.777 2011     2700      1     2.636
## 132905 79958782230 3.713 2011     2700      1     2.572
## 133601 79957444504 4.382 2011     2700      1     3.346
## 133602 79957445348 4.662 2011     2700      1     3.725
## 133644 79956204057 3.846 2011     2700      1     2.705
## 133684 79956311926 3.968 2011     1000      2     2.727
## 133688 79956189911 4.080 2011     2700      1     2.939
## 133690 79956218863 5.583 2011     2700      1     4.442
## 133696 79956258982 3.993 2011     2700      1     2.752
## 133759 79955909191 4.421 2011     2700      1     3.280
## 133811 79955823001 3.760 2011     2700      1     2.619
## 133813 79955877895 3.880 2011     2700      1     2.739
## 133897 79955624770 4.409 2011     2700      1     3.268
## 133939 79955517235 4.745 2011     2700      1     3.504
## 133996 79955115145 3.695 2011     2700      1     2.554
## 134076 79958006776 3.885 2011     2700      1     2.948
## 134662 79955376345 3.993 2011     2700      1     2.752
## 134719 79955009821 4.244 2011     2700      1     3.003
## 134722 79955016374 3.663 2011     2700      1     2.726
## 134723 79955035027 3.993 2011     2700      1     2.852
## 134724 79955038968 3.967 2011     2700      1     2.826
## 134759 79954561234 4.639 2011     2700      1     3.498
## 134775 79954487919 5.185 2011     2700      1     4.248
## 134778 79954553689 4.428 2011     2700      1     3.287
## 134783 79953848750 3.799 2011     2700      1     2.763
## 134788 79954542852 3.731 2011     2700      1     2.590
## 134883 79953748230 4.910 2011     2700      1     3.669
## 134959 79953297488 4.005 2011     2700      1     2.864
## 134968 79953307878 4.700 2011     2700      1     3.559
## 134978 79953329970 4.175 2011     2700      1     3.034
## 134979 79953331206 4.492 2011     2700      1     3.351
## 135559 79953225169 4.658 2011     2700      1     3.721
## 135593 79952941296 3.793 2011     2700      1     2.856
## 135629 79953232077 4.567 2011     2700      1     3.426
## 135637 79953118839 5.685 2011     2700      1     4.444
## 135705 79952795093 4.650 2011     2700      1     3.713
## 135762 79952376771 4.188 2011     2700      1     3.047
## 135766 79952424506 4.158 2011     2700      1     3.017
## 135793 79952476762 4.836 2011     2700      1     3.800
## 135911 79952363727 5.013 2011     2700      1     3.872
## 135923 79952267051 4.395 2011     2700      1     3.154
## 135924 79952271716 3.900 2011     2700      1     2.759
## 136005 79952219919 4.001 2011     2700      1     2.860
## 136036 79952570027 3.688 2011     2700      1     2.547
## 136042 79952749048 3.603 2011     2700      1     2.567
## 136941 79951961755 3.835 2011     2700      1     2.694
## 136954 79951978171 3.825 2011     2700      1     2.584
## 136956 79952204045 3.688 2011     2700      1     2.547
```

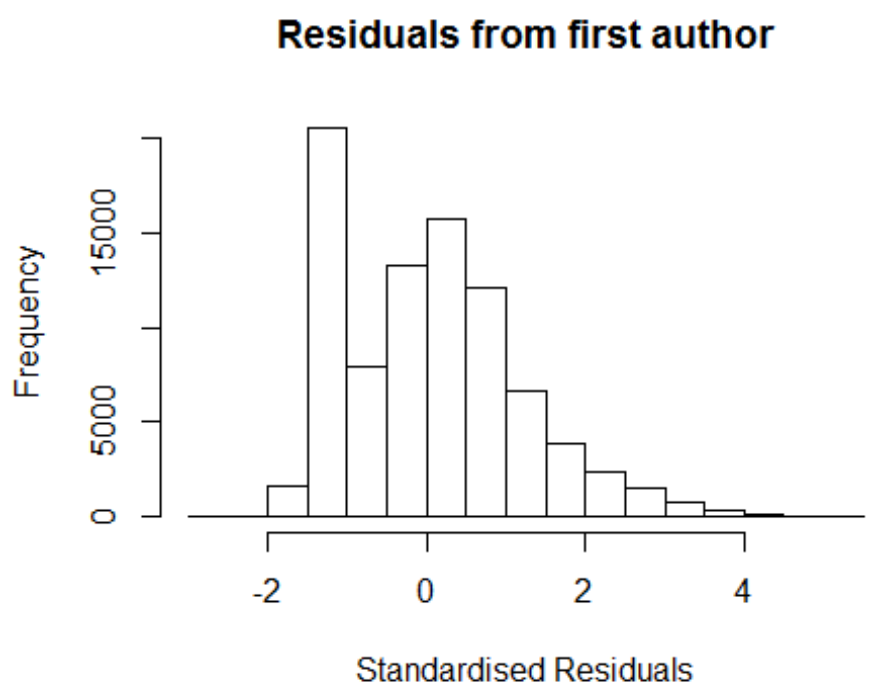
```
## 136959 79951778987 3.914 2011     2700      1     2.773
## 136973 79951850287 4.863 2011     2700      1     3.722
## 136988 79951821800 4.881 2011     2700      1     3.845
## 136997 79951699237 4.608 2011     2700      1     3.367
## 137041 79951514365 3.458 2011     1000      2     2.521
## 137056 79851493086 4.470 2011     2700      1     3.229
## 137064 79751522951 5.633 2011     2700      1     4.696
## 137065 79851475199 3.707 2011     2700      1     2.770
## 137169 79551566712 4.247 2011     2700      1     3.106
## 137251 79251491762 3.799 2011     2700      1     2.658
## 137265 79551523377 4.106 2011     2700      1     3.169
## 137278 79551666741 3.927 2011     2700      1     2.686
## 137775 79251501962 3.820 2011     2700      1     2.679
## 137776 79251525181 3.985 2011     2700      1     2.844
## 137777 79251527865 4.001 2011     2700      1     2.760
## 137786 78751642476 3.804 2011     2700      1     2.663
## 137800 78751610848 4.907 2011     2700      1     3.666
## 137805 79952148734 4.771 2011     2700      1     3.630
## 137813 78751519245 3.731 2011     2700      1     2.695
## 137814 78751532610 3.719 2011     2700      1     2.683
## 137817 78751667739 3.989 2011     2700      1     2.848
## 137857 78651351928 3.959 2011     2700      1     3.022
## 137860 78651378996 3.950 2011     2700      1     2.809
## 137861 78651379500 5.349 2011     2700      1     4.208
## 137865 78651393550 4.594 2011     2700      1     3.453
## 137867 78651505523 3.777 2011     2700      1     2.636
## 137896 78650893810 3.993 2011     2700      1     2.752
## 137905 78650940402 3.574 2011     2700      1     2.538
## 137909 78650949272 4.632 2011     2700      1     3.391
## 137988 78650987230 3.560 2011     2700      1     2.623
## 137995 78651245300 4.046 2011     2700      1     3.010
## 138009 79151481493 3.670 2011     2700      1     2.529
## 138381 79955523146 3.731 2011     2700      1     2.590
## 138610 80051964418 3.980 2011     2700      1     2.839
## 139175 79952183161 3.858 2011     2700      2     2.717
## 139231 79958822104 4.047 2011     2700      2     3.110
## 139857 84871293779 4.110 2012     2700      1     2.991
## 139901 84871314601 3.955 2012     2700      1     2.836
## 140049 84870877380 3.955 2012     2700      1     2.836
## 140050 84870887067 4.610 2012     2700      1     3.491
## 140140 84870549609 4.439 2012     2700      1     3.320
## 141706 84870012939 4.380 2012     2700      1     3.261
## 141721 84869792699 3.633 2012     2700      1     2.619
## 141725 84870312106 4.184 2012     2700      1     2.965
## 141912 84868691029 4.249 2012     2700      1     3.130
## 142043 84868305308 4.241 2012     2700      1     3.022
## 142202 84868609657 3.465 2012     2700      1     2.551
## 142208 84868613782 3.623 2012     2700      1     2.504
## 142976 84867828250 4.526 2012     2700      1     3.512
## 143180 84867221377 4.243 2012     2700      1     3.024
```

```
## 143205 84867437406 3.667 2012     2700      1     2.548
## 143276 84867273101 3.658 2012     2700      1     2.644
## 144201 84866555961 4.254 2012     2700      1     3.135
## 144204 84866628742 3.794 2012     2700      1     2.780
## 144331 84866098277 4.037 2012     2700      1     2.918
## 144357 84866271313 3.691 2012     2700      1     2.677
## 144361 84866395310 4.216 2012     2700      1     3.097
## 144362 84866395319 3.995 2012     2700      1     2.876
## 144505 84865395104 3.895 2012     2700      1     2.776
## 144517 84865987121 3.945 2012     2700      1     2.826
## 145514 84865454335 3.901 2012     2700      1     2.987
## 145517 84865487218 4.045 2012     2700      1     2.826
## 145540 84865701980 4.297 2012     2700      1     3.178
## 145542 84865820981 3.939 2012     2700      1     2.720
## 145641 84865320778 4.511 2012     2700      1     3.392
## 145642 84865323065 4.426 2012     2700      1     3.512
## 145875 84864842736 4.100 2012     2700      1     2.881
## 146034 84864286532 3.526 2012     2700      1     2.512
## 146664 84864242034 4.379 2012     2700      1     3.260
## 146677 84864138716 3.932 2012     2700      1     2.813
## 146785 84864008031 5.104 2012     2700      1     3.985
## 146790 84930197839 3.672 2012     2700      1     2.553
## 146935 84863923855 3.855 2012     2700      1     2.736
## 147014 84863524725 3.731 2012     2700      1     2.817
## 147015 84863524855 5.149 2012     2700      1     3.930
## 147018 84863533887 4.258 2012     2700      1     3.039
## 147075 84863550146 3.806 2012     2700      1     2.587
## 147080 84863554398 3.686 2012     2700      1     2.772
## 147811 84862859820 6.239 2012     2700      1     5.020
## 147814 84862903106 5.924 2012     2700      1     4.705
## 147913 84862511733 3.543 2012     2700      1     2.629
## 147933 84862859054 3.790 2012     2700      1     2.571
## 148038 84862226987 4.247 2012     2700      1     3.333
## 148062 84862489245 3.731 2012     2700      1     2.717
## 148141 84862118837 3.961 2012     2700      1     2.842
## 148146 84862260833 4.332 2012     2700      1     3.113
## 148915 84861689703 3.695 2012     2700      1     2.576
## 149011 84861394764 4.694 2012     2700      1     3.575
## 149012 84861409254 3.735 2012     2700      1     2.616
## 149013 84861409551 4.560 2012     2700      1     3.546
## 149126 84861018546 4.042 2012     2700      1     3.128
## 149127 84861022041 4.793 2012     2700      1     3.879
## 149136 84862777742 3.759 2012     1000      2     2.540
## 149215 84860914047 3.560 2012     2700      1     2.546
## 149238 84860741191 4.704 2012     2700      1     3.585
## 149259 84860776122 4.355 2012     2700      1     3.236
## 149264 84860910413 3.658 2012     2700      1     2.539
## 149362 84860718270 4.087 2012     2700      1     2.968
## 149363 84860720391 4.234 2012     2700      1     3.115
## 150121 84860344255 3.718 2012     2700      1     2.599
```

```
## 150154 84860116634 3.813 2012     2700      1     2.694
## 150155 84860136615 5.159 2012     2700      1     4.040
## 150175 84860200127 4.100 2012     2700      1     2.981
## 150177 84860358233 3.618 2012     2700      1     2.604
## 150256 84859819731 4.273 2012     2700      1     3.154
## 150277 84859834632 4.025 2012     2700      1     2.906
## 150279 84859846433 4.100 2012     2700      1     2.981
## 150384 84859555074 4.313 2012     2700      1     3.194
## 150445 84860383862 4.156 2012     2700      1     3.037
## 150503 84858665432 4.608 2012     2700      1     3.489
## 150508 84859544287 4.701 2012     2700      1     3.582
## 150550 84859092587 3.942 2012     2700      1     2.823
## 150563 84859718265 4.683 2012     2700      1     3.564
## 151282 84859406106 3.740 2012     2700      1     2.621
## 151344 84858638369 4.465 2012     2700      1     3.346
## 151347 84863337617 4.454 2012     2700      1     3.335
## 151349 84858796262 3.993 2012     1000      2     2.874
## 151408 84858331964 4.546 2012     1000      2     3.427
## 151434 84858265970 4.576 2012     2700      1     3.457
## 151459 84858329412 3.813 2012     2700      1     2.694
## 151554 84863229939 4.007 2012     2700      1     3.093
## 151641 84858019974 4.876 2012     2700      1     3.757
## 151895 84863267488 3.633 2012     2700      1     2.514
## 152472 84857848576 3.905 2012     2700      1     2.686
## 152564 84857427752 4.330 2012     2700      1     3.211
## 152567 84863115197 3.667 2012     2700      1     2.548
## 152654 84857132739 4.189 2012     2700      1     3.070
## 152684 84857065859 4.573 2012     2700      1     3.354
## 152856 84856433589 5.714 2012     2700      1     4.700
## 152866 84856487711 5.900 2012     2700      1     4.886
## 152870 84856552278 3.704 2012     2700      1     2.585
## 152908 84857643783 4.842 2012     2700      1     3.723
## 153052 84863012865 4.087 2012     2700      1     3.073
## 153390 84856245226 4.422 2012     2700      1     3.303
## 153392 84856249964 3.802 2012     2700      1     2.683
## 153393 84856253589 4.184 2012     2700      1     3.065
## 153404 84856133161 3.633 2012     2700      1     2.619
## 153473 84856159009 4.542 2012     2700      1     3.323
## 153475 84862908900 3.581 2012     2700      1     2.567
## 153490 84855993116 3.748 2012     2700      1     2.734
## 153532 84855843640 3.790 2012     2700      1     2.876
## 153533 84855854046 4.736 2012     2700      1     3.617
## 153548 84856176782 4.001 2012     2700      1     3.087
## 153549 84862909100 4.062 2012     2700      1     2.943
## 153663 84855426656 3.662 2012     2700      1     2.543
## 153672 84855459760 4.667 2012     2700      1     3.548
## 153732 84055199809 4.449 2012     2700      1     3.330
## 153835 84856628070 3.840 2012     2700      1     2.721
## 153838 84856657579 4.089 2012     2700      1     2.870
## 154252 84863718620 3.870 2012     2700      1     2.956
```

```
## 154329 84864861863 4.382 2012     2700      1     3.263
## 154561 84870478439 4.172 2012     2700      1     3.053
## 154562 84870494510 4.054 2012     2700      1     2.935
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##       Min        1Q    Median        3Q       Max
## -2.758597 -0.949531  0.000149  0.710260  5.457976
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.91976    0.03110   61.73   <2e-16 ***
## FirstAuthorFemale1  0.09971    0.00843   11.83   <2e-16 ***
## LastAuthorFemale1  -0.20466    0.00884  -23.14   <2e-16 ***
## Year1997           -0.00739    0.04247   -0.17     0.86
## Year1998            0.03665    0.04237    0.86     0.39
## Year1999           -0.01880    0.04212   -0.45     0.66
## Year2000            0.73913    0.09366    7.89    3e-15 ***
## Year2001           -0.06343    0.04274   -1.48     0.14
## Year2002           -0.54207    0.03666  -14.79   <2e-16 ***
## Year2003           -0.60405    0.03768  -16.03   <2e-16 ***
## Year2004           -0.69163    0.03570  -19.37   <2e-16 ***
## Year2005           -0.72006    0.03544  -20.32   <2e-16 ***
## Year2006           -0.75036    0.03402  -22.05   <2e-16 ***
## Year2007           -0.86287    0.03335  -25.87   <2e-16 ***
## Year2008           -0.86528    0.03285  -26.34   <2e-16 ***
## Year2009           -0.82098    0.03271  -25.10   <2e-16 ***
## Year2010           -0.77634    0.03218  -24.13   <2e-16 ***
## Year2011           -0.77860    0.03297  -23.62   <2e-16 ***
## Year2012           -0.80094    0.03238  -24.73   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 1.07
## Multiple R-squared:  0.0753, Adjusted R-squared:  0.0751
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
##  9 observations c(14272,15323,22486,33665,36647,38880,39540,73247,82361)
##   are outliers with |weight| = 0 ( < 1.2e-06);
##  6961 weights are ~= 1. The remaining 79688 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0002  0.8960  0.9380  0.9060  0.9830  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
```

```
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.15e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.005  1           1.003
## Year              1.005 16           1.000
```




```
## [1] "List of  2719 outliers with residuals above  2.5"
##           ScopusId  NLCS Year OneField Fields residuals
## 28      0029855881 4.801 1996     2700      1     2.881
## 30      0029906954 4.525 1996     2700      1     2.605
## 31      0141469033 5.783 1996     2700      1     3.863
## 47      0030453430 5.115 1996     2700      1     3.195
## 79     10544241545 4.738 1996     2700      1     2.719
## 87     10544235697 5.572 1996     2700      1     3.652
## 95      0029911547 5.746 1996     2700      1     4.031
```

```
## 113     0030445619 5.158 1996     2700      1     3.139
## 153     0029683640 4.469 1996     2700      1     2.549
## 179     0029844930 4.941 1996     2700      1     3.021
## 197     0029979690 4.533 1996     2700      1     2.514
## 1288    0029827650 4.623 1996     2700      1     2.703
## 1305    0030599533 4.661 1996     2700      1     2.642
## 1309    0029828983 4.411 1996     2700      1     2.596
## 1355    0030590745 4.426 1996     2700      1     2.506
## 1359    0029808003 5.038 1996     2700      1     3.019
## 1369    0029845013 4.799 1996     2700      1     2.879
## 1394    0029859887 4.512 1996     2700      1     2.797
## 1397    0029919459 5.204 1996     2700      1     3.185
## 1399    3042828081 4.528 1996     2700      1     2.509
## 1414   10544220023 4.705 1996     2700      1     2.785
## 1424    0029907556 5.146 1996     2700      1     3.331
## 1425    0029908290 4.742 1996     2700      1     2.822
## 1649    0029849462 4.874 1996     2700      1     2.954
## 1672    0030603334 4.697 1996     2700      1     2.678
## 1677    0029862114 4.329 1996     2700      1     2.614
## 1708    0029823520 4.701 1996     2700      1     2.781
## 1717   10144258656 4.587 1996     2700      1     2.667
## 1725    0029861097 4.833 1996     2700      1     3.118
## 1728    0029861579 4.669 1996     2700      1     2.650
## 1755    0030581586 5.042 1996     2700      1     3.327
## 1756    0030581587 4.964 1996     2700      1     2.945
## 1757    0030581590 4.755 1996     2700      1     2.736
## 1758    0242409954 4.538 1996     2700      1     2.618
## 1763    0005150171 4.797 1996     2700      1     2.877
## 1764    0007410805 5.440 1996     2700      1     3.520
## 1766   10144244674 5.218 1996     2700      1     3.298
## 1772    0029758899 4.682 1996     2700      1     2.867
## 1773    0029759549 4.361 1996     2700      1     2.646
## 1787    0010496839 5.103 1996     2700      1     3.183
## 1788    0029795224 5.088 1996     2700      1     3.069
## 1798    0029773862 4.819 1996     2700      1     2.899
## 1958   10144241022 5.066 1996     2700      1     3.047
## 1977    0029798819 4.729 1996     1300      2     2.809
## 1979    0029842830 5.287 1996     1300      2     3.367
## 2023    0030604909 4.869 1996     2700      1     2.850
## 2032    0029790755 4.602 1996     2700      1     2.583
## 2033    0029840653 5.036 1996     2700      1     3.116
## 2070    9544253883 4.485 1996     2700      1     2.770
## 2071    0029740950 5.202 1996     2700      1     3.282
## 2072    0029741921 5.309 1996     2700      1     3.389
## 2074    9544228424 5.880 1996     2700      1     3.960
## 2107    0029835263 5.098 1996     2700      1     3.079
## 2108    0029835392 5.259 1996     2700      1     3.339
## 2109    0029838499 4.691 1996     2700      1     2.672
## 2110    0242435873 5.375 1996     2700      1     3.660
## 2112    9544222721 4.784 1996     2700      1     3.069
```

```
## 2113    0029739543 4.452 1996     2700      1     2.532
## 2115    0029771618 4.842 1996     2700      1     3.027
## 2130    0029818344 4.305 1996     2700      1     2.590
## 2131    0029822091 4.840 1996     2700      1     2.920
## 2159    0029737324 4.674 1996     2700      1     2.754
## 2162    0029739970 5.663 1996     2700      1     3.743
## 2167    0029758146 4.466 1996     2700      1     2.546
## 2354    0030250166 4.259 1996     2700      1     2.544
## 2485    0029741822 5.190 1996     2700      1     3.171
## 2486    0029746491 4.535 1996     2700      1     2.615
## 2489    9444268678 5.105 1996     2700      1     3.185
## 2490    0029740582 4.658 1996     2700      1     2.738
## 2500    0030017720 4.990 1996     2700      1     3.070
## 2503    0029783235 4.667 1996     2700      1     2.747
## 2510    0029739053 4.836 1996     2700      1     2.916
## 2513    0029759025 5.617 1996     2700      1     3.598
## 2526    0029830177 4.802 1996     2700      1     2.987
## 2537    0029737310 4.288 1996     2700      1     2.573
## 2543   15844420661 5.335 1996     2700      1     3.415
## 2574    0029830551 4.585 1996     2700      1     2.665
## 2582    0029862164 4.517 1996     2700      1     2.702
## 2588    0029916051 4.499 1996     2700      1     2.684
## 2741    0030594830 5.438 1996     2700      1     3.518
## 2743    0030000230 4.826 1996     2700      1     2.906
## 2744    8944228913 5.408 1996     2700      1     3.389
## 2748    0030056083 5.607 1996     2700      1     3.687
## 2750    0030057106 5.671 1996     2700      1     3.751
## 2797    8944233864 5.465 1996     2700      1     3.545
## 2806    0029902170 4.900 1996     2700      1     3.085
## 2830    0030015522 4.848 1996     2700      1     2.928
## 3004    0029944290 4.689 1996     1300      2     2.769
## 3042    0029786164 4.533 1996     2700      1     2.613
## 3058    0029900294 5.422 1996     2700      1     3.502
## 3059    0029942003 5.780 1996     2700      1     3.860
## 3085    0030594604 4.947 1996     2700      1     3.027
## 3109    0029948212 5.514 1996     2700      1     3.594
## 3111    0029948967 4.647 1996     2700      1     2.727
## 3154    9344267138 4.432 1996     2700      1     2.717
## 3166   15844364006 4.499 1996     2700      1     2.579
## 3185    0029888972 4.889 1996     2700      1     3.174
## 3214    0030000088 4.916 1996     2700      1     2.996
## 3462    0029892519 4.994 1996     2700      1     3.074
## 3468    0029939573 4.570 1996     2700      1     2.650
## 3471    0029878679 4.693 1996     2700      1     2.773
## 3474    0029897909 5.220 1996     2700      1     3.201
## 3479    0029886630 5.409 1996     2700      1     3.489
## 3487    0029889718 4.606 1996     2700      1     2.686
## 3498    0029870954 4.842 1996     2700      1     2.922
## 3503    0029932125 4.845 1996     2700      1     2.826
## 3521    0029925567 5.575 1996     2700      1     3.655
```

```
## 3537    0029871858 4.847 1996     2700      1     2.828
## 3539    0030008301 4.452 1996     2700      1     2.532
## 3552    9244262406 6.270 1996     2700      1     4.350
## 3560    0029863684 5.262 1996     2700      1     3.342
## 3572    0029913381 4.769 1996     2700      1     2.750
## 3573    0029913643 4.471 1996     2700      1     2.551
## 3578    0029929617 5.256 1996     2700      1     3.541
## 3579    0029929618 4.964 1996     2700      1     3.249
## 3795    0029875770 5.139 1996     2700      1     3.219
## 3797    0029982081 4.757 1996     2700      1     2.837
## 3806    0029988316 4.987 1996     2700      1     3.067
## 3830    0029920903 4.358 1996     2700      1     2.643
## 3866    0029881352 4.449 1996     2700      1     2.529
## 3877    0029869817 4.507 1996     2700      1     2.792
## 3890    0029993565 5.060 1996     2700      1     3.140
## 3892   13344285352 4.993 1996     2700      1     2.974
## 4094    0029915834 4.443 1996     2700      1     2.523
## 4105    0029970342 5.515 1996     2700      1     3.595
## 4111    0029917496 5.333 1996     2700      1     3.314
## 4150    0029993531 4.882 1996     2700      1     2.962
## 4157    0030009316 4.530 1996     2700      1     2.610
## 4168    0029873844 4.915 1996     2700      1     2.896
## 4169    0029987128 4.757 1996     2700      1     2.738
## 4171    0029870311 4.393 1996     2700      1     2.678
## 4178    9044236527 5.250 1996     2700      1     3.535
## 4195    0030070405 4.558 1996     2700      1     2.743
## 4217    0029866721 4.724 1996     2700      1     2.705
## 4509    0942276272 5.608 1996     2700      1     3.688
## 4519   13344293705 4.423 1996     2700      1     2.608
## 4535    0030066467 5.150 1996     2700      1     3.230
## 4536    0030070793 4.738 1996     2700      1     3.023
## 4548    0030020590 4.565 1996     1300      2     2.546
## 4552    0030020630 4.999 1996     2700      1     3.079
## 4567    0030058515 5.136 1996     2700      1     3.216
## 4591    0030042790 5.170 1996     2700      1     3.250
## 4593   13344286314 4.575 1996     2700      1     2.655
## 4596    0030068231 4.602 1996     2700      1     2.682
## 4598    0030058666 4.263 1996     2700      1     2.548
## 4603    0030033367 4.824 1996     2700      1     2.904
## 4622    0030032444 5.228 1996     2700      1     3.308
## 4629    0030049026 4.570 1996     2700      1     2.650
## 4632    0030056373 4.773 1996     2700      1     2.853
## 4784    0030032258 4.399 1996     2700      1     2.684
## 4802    0030044553 4.676 1996     2700      1     2.961
## 4818    0030049823 4.429 1996     2700      1     2.614
## 4819   13344260688 4.602 1996     2700      1     2.682
## 4826    0030048691 4.811 1996     2700      1     3.096
## 4827    0030060343 4.437 1996     2700      1     2.517
## 4833    0030026954 4.778 1996     2700      1     2.963
## 4843    0030032378 4.408 1996     2700      1     2.693
```

```
## 4866    0030043663 4.884 1996     2700      1     2.964
## 4868    0030579588 4.838 1996     2700      1     2.918
## 4869    0030034463 5.009 1996     2700      1     2.990
## 4883    0030034592 5.638 1996     2700      1     3.718
## 4884    0030054309 6.200 1996     2700      1     4.280
## 4890    0030024585 4.558 1996     2700      1     2.638
## 4891    0030046074 4.501 1996     2700      1     2.786
## 4897    0000677401 5.404 1996     2700      1     3.589
## 4904    0001196648 5.048 1996     2700      1     3.128
## 4908    0001854685 4.563 1996     2700      1     2.748
## 4916    0004851872 6.351 1996     2700      1     4.636
## 5482    0029959975 4.583 1996     2700      1     2.564
## 5653    0030034465 5.226 1996     2700      1     3.306
## 5706    0030054203 4.894 1996     2700      1     2.974
## 5729    0030060710 4.744 1996     2700      1     2.725
## 5737    0030061707 4.711 1996     2700      1     2.791
## 5847    0030183205 4.634 1996     2700      1     2.819
## 6074    0030576183 4.548 1996     2700      1     2.628
## 6086    0040419081 4.738 1996     2700      1     3.023
## 6118   16144365807 4.546 1996     2700      1     2.527
## 6164    9244238642 4.854 1996     2700      1     2.934
## 7329    0031435838 4.794 1997     2700      1     3.086
## 7330    2642611953 4.758 1997     2700      1     2.951
## 7352    0031468254 4.620 1997     2700      1     2.708
## 7360    0031472414 5.042 1997     2700      1     3.130
## 7367    0031437624 4.545 1997     2700      1     2.633
## 7375    0031466866 4.796 1997     2700      1     2.884
## 7379    2642597076 5.193 1997     2700      1     3.281
## 7415    0030704598 4.531 1997     2700      1     2.619
## 7534    0031136510 4.780 1997     2700      1     2.868
## 7757    0031454125 4.470 1997     2700      1     2.558
## 7787    0031472452 4.431 1997     2700      1     2.723
## 8304    0030703240 4.771 1997     2700      1     2.759
## 8307    0030665744 4.564 1997     2700      1     2.652
## 8323    0031590631 4.483 1997     2700      1     2.571
## 8332    0030692782 5.405 1997     2700      1     3.493
## 8339    0030666228 4.564 1997     2700      1     2.757
## 8340    0030671045 4.349 1997     2700      1     2.641
## 8357    0030809817 4.318 1997     2700      1     2.610
## 8378    0030695138 5.236 1997     2700      1     3.324
## 8424    0030716498 5.322 1997     2700      1     3.410
## 8442    0030661687 5.440 1997     2700      1     3.528
## 8455    0030671503 4.307 1997     2700      1     2.599
## 8586    0030731486 5.178 1997     2700      1     3.371
## 8589    0030831628 4.379 1997     2700      1     2.671
## 8608    0030720886 5.370 1997     2700      1     3.662
## 8610    0030759072 5.217 1997     2700      1     3.305
## 8625    0030744945 5.294 1997     2700      1     3.382
## 8627    0030762081 4.480 1997     2700      1     2.568
## 8631    0030879325 4.935 1997     2700      1     3.023
```

```
## 8692    0030766770 4.941 1997     2700      1     3.233
## 8713    0030779037 5.224 1997     2700      1     3.212
## 8714    0030803131 4.352 1997     2700      1     2.644
## 8728    0030774499 4.679 1997     2700      1     2.872
## 8849    0030886175 4.489 1997     2700      1     2.577
## 8850    0030954873 5.207 1997     2700      1     3.499
## 8851    0030967953 6.065 1997     2700      1     4.153
## 8861    0006750594 4.483 1997     2700      1     2.571
## 8891    0030756517 4.679 1997     2700      1     2.767
## 8980    0030869269 5.624 1997     2700      1     3.712
## 8983    9844257578 4.711 1997     2700      1     3.003
## 9007    0030763532 5.703 1997     2700      1     3.691
## 9014    0030928107 4.550 1997     2700      1     2.842
## 9045    0030921893 4.869 1997     2700      1     2.957
## 9237    0030803718 4.703 1997     2700      1     2.896
## 9239    0030868293 5.188 1997     2700      1     3.276
## 9245    0030804947 4.885 1997     2700      1     3.177
## 9250    0030792120 4.836 1997     2700      1     2.924
## 9309    8544284052 4.624 1997     2700      1     2.712
## 9329    0030851442 4.840 1997     2700      1     2.928
## 9331    0030872063 4.662 1997     2700      1     2.750
## 9332    1842332651 4.618 1997     2700      1     2.706
## 9378    0030854951 4.445 1997     2700      1     2.533
## 9468    8544252402 4.498 1997     2700      1     2.790
## 9514    0030744599 4.744 1997     2700      1     2.832
## 9515    0030756101 4.907 1997     2700      1     3.100
## 9519    0030854780 4.755 1997     2700      1     2.843
## 9522    0030742587 4.609 1997     2700      1     2.901
## 9539    0030876805 5.124 1997     2700      1     3.212
## 9568    0030803395 4.639 1997     2700      1     2.931
## 9597    0030610461 5.331 1997     2700      1     3.319
## 9603    0038025895 4.877 1997     2700      1     2.965
## 9607    0030957310 5.247 1997     2700      1     3.335
## 9629    0030760729 4.653 1997     2700      1     2.741
## 9652    0030878260 4.429 1997     2700      1     2.517
## 9760    0031005933 5.176 1997     2700      1     3.264
## 9763    1842410169 4.665 1997     2700      1     2.653
## 9767    0030916401 5.117 1997     2700      1     3.205
## 9816    0030908055 5.171 1997     2700      1     3.259
## 9818    0030941817 4.703 1997     2700      1     2.691
## 9820    0030976173 4.495 1997     2700      1     2.787
## 9823    0031009275 4.344 1997     2700      1     2.537
## 9824    8244234470 4.344 1997     2700      1     2.636
## 9833    0343157354 5.144 1997     2700      1     3.232
## 9902    0030974188 4.972 1997     2700      1     3.060
## 10055   0030909527 4.552 1997     2700      1     2.745
## 10056   0030910022 5.719 1997     2700      1     4.011
## 10060   0030939253 4.627 1997     2700      1     2.715
## 10066   0031003334 5.309 1997     2700      1     3.397
## 10071   0343918505 4.869 1997     2700      1     3.161
```

```
## 10072   0346593922 4.853 1997     2700      1     2.941
## 10098   0031009871 4.652 1997     2700      1     2.845
## 10102   0030919511 5.119 1997     2700      1     3.207
## 10103   0030919667 4.467 1997     2700      1     2.555
## 10127   0030977450 4.957 1997     2700      1     3.045
## 10149   0030913316 4.821 1997     2700      1     2.909
## 10156   0030961921 5.444 1997     2700      1     3.736
## 10184   0030896520 5.060 1997     2700      1     3.048
## 10222   0030979720 4.965 1997     2700      1     3.053
## 10347   0030967165 5.834 1997     2700      1     3.822
## 10353   0030955080 4.722 1997     2700      1     3.014
## 10378   0012444519 6.220 1997     2700      1     4.308
## 10386   0030903440 4.583 1997     2700      1     2.875
## 10427   0030896996 4.738 1997     2700      1     3.030
## 10428   0030899940 5.040 1997     2700      1     3.028
## 10454   0030956673 6.413 1997     2700      1     4.501
## 10461   0030951198 5.067 1997     2700      1     3.155
## 10464   0031004940 4.520 1997     2700      1     2.812
## 10467   0030891988 4.695 1997     2700      1     2.683
## 10490   0030946366 4.576 1997     2700      1     2.564
## 10535   0031127072 5.869 1997     2700      1     4.062
## 10598   0030976067 5.207 1997     2700      1     3.195
## 10622   0030898397 4.415 1997     2700      1     2.503
## 10623   0030948376 4.687 1997     2700      1     2.775
## 10646   0030945186 4.445 1997     2700      1     2.533
## 10660   0031047918 4.572 1997     2700      1     2.560
## 10671   0031054674 5.776 1997     2700      1     3.864
## 10694   0031047361 4.710 1997     2700      1     2.798
## 10735   0030948337 4.434 1997     2700      1     2.627
## 10891   0013655172 4.847 1997     2700      1     2.935
## 10895   0031052369 4.993 1997     2700      1     3.081
## 10901   0031018269 4.695 1997     2700      1     2.783
## 10919   0031045652 5.274 1997     2700      1     3.362
## 10920   0031050740 4.653 1997     2700      1     2.741
## 10933   0031052862 5.331 1997     2700      1     3.319
## 10935   0031057614 4.810 1997     2700      1     2.798
## 10972   0031058448 4.616 1997     2700      1     2.809
## 10979   0031019745 4.552 1997     1000      2     2.640
## 10981   0031019923 4.600 1997     2700      1     2.688
## 10983   0031028712 4.450 1997     2700      1     2.538
## 10987   0031028804 5.501 1997     2700      1     3.589
## 10993   0030614495 4.450 1997     2700      1     2.538
## 11012   0030807358 4.886 1997     2700      1     2.874
## 11019   0030868077 4.859 1997     2700      1     2.947
## 11136   0031012726 5.498 1997     2700      1     3.586
## 11137   0031030450 4.891 1997     2700      1     2.979
## 11138   0031032055 5.515 1997     2700      1     3.603
## 11153   0031030620 5.167 1997     2700      1     3.255
## 11162   0031012532 4.615 1997     2700      1     2.703
## 11174   0031028365 4.604 1997     2700      1     2.592
```

```
## 11175   0031029379 4.969 1997     2700      1     2.957
## 11202   0031025101 4.552 1997     2700      1     2.844
## 11213   0031037065 4.605 1997     2700      1     2.593
## 11823   0030938817 4.429 1997     2700      1     2.721
## 11955   0031011193 4.851 1997     2700      1     2.939
## 11971   0031020373 4.539 1997     2700      1     2.527
## 11972   0031022490 4.618 1997     2700      1     2.606
## 13373   0031472453 4.648 1997     2700      1     2.736
## 13487   1842295778 5.282 1997     2700      1     3.370
## 13948   0032585514 4.968 1998     2700      1     3.012
## 13961   0032583492 5.604 1998     2700      1     3.852
## 13975   0032564658 4.473 1998     2700      1     2.517
## 13977   0032564703 4.515 1998     2700      1     2.763
## 13994   0032542385 4.640 1998     2700      1     2.684
## 14001   0032539164 4.551 1998     2700      1     2.700
## 14018   0032506614 5.304 1998     2700      1     3.348
## 14032   0032480930 4.637 1998     2700      1     2.681
## 14040   0032477294 5.316 1998     2700      1     3.564
## 14041   0032477305 5.311 1998     2700      1     3.559
## 14043   0032477331 5.005 1998     2700      1     3.154
## 14065   0031609707 4.571 1998     2700      1     2.615
## 14801   0032567115 5.271 1998     2700      1     3.315
## 14816   0032547938 6.002 1998     2700      1     4.250
## 14828   0032544946 4.306 1998     2700      1     2.554
## 14860   0032512097 4.594 1998     2700      1     2.638
## 14866   0032508952 6.390 1998     2700      1     4.434
## 14881   0032487923 5.141 1998     2700      1     3.185
## 14888   0032483685 5.248 1998     2700      1     3.292
## 14889   0032483690 4.752 1998     2700      1     2.796
## 15041   0032578789 4.449 1998     2700      1     2.697
## 15080   0032556180 5.427 1998     2700      1     3.675
## 15109   0032517251 4.717 1998     2700      1     2.965
## 15110   0032517258 4.889 1998     2700      1     3.137
## 15130   0032494396 4.981 1998     2700      1     2.925
## 15186   0032189897 5.522 1998     2700      1     3.566
## 15322   0032563824 4.832 1998     2700      1     2.876
## 15354   0032541616 5.053 1998     2700      1     3.097
## 15357   0032541672 4.883 1998     2700      1     2.927
## 15362   0032538047 4.487 1998     2700      1     2.531
## 15396   0032504985 4.662 1998     2700      1     2.706
## 15418   0032480346 4.888 1998     2700      1     2.932
## 15530   0032169485 5.282 1998     2700      1     3.326
## 15532   0032169557 4.561 1998     2700      1     2.605
## 15630   0032572929 4.615 1998     2700      1     2.659
## 15666   0032547326 6.369 1998     2700      1     4.413
## 15667   0032547328 4.489 1998     2700      1     2.533
## 15690   0032529103 5.002 1998     2700      1     3.046
## 15700   0008926519 5.466 1998     2700      1     3.510
## 15741   0032486734 4.466 1998     2700      1     2.714
## 15784   0032145836 4.610 1998     2700      1     2.554
```

```
## 15882   0032578110 4.511 1998     2700      1     2.660
## 15906  18744421762 4.597 1998     2700      1     2.641
## 15932   0032528169 4.517 1998     2700      1     2.666
## 15935   0032528180 5.239 1998     2700      1     3.283
## 15947   0032508297 5.453 1998     2700      1     3.497
## 15961   0032496880 4.515 1998     2700      1     2.559
## 15971   0032474694 4.931 1998     2700      1     2.975
## 16151   0032551162 4.753 1998     2700      1     2.797
## 16167   0032543663 5.069 1998     2700      1     3.113
## 16172   7144228604 4.850 1998     2700      1     2.894
## 16174   0009452157 4.306 1998     2700      1     2.554
## 16193   0032525901 4.517 1998     2700      1     2.765
## 16222   0032503637 4.575 1998     2700      1     2.823
## 16251   0032482329 5.835 1998     2700      1     4.083
## 16263   0032479034 4.911 1998     2700      1     2.855
## 16346   0032102292 5.131 1998     2700      1     3.175
## 16483   0032572043 5.088 1998     2700      1     3.336
## 16487   0032572086 6.274 1998     2700      1     4.318
## 16504   0032550626 5.681 1998     2700      1     3.725
## 16505   0032550630 5.027 1998     2700      1     3.275
## 16536   0032516267 4.935 1998     2700      1     3.183
## 16539   0032516296 5.565 1998     2700      1     3.609
## 16568   0032490147 4.494 1998     2700      1     2.538
## 16773   0032580484 4.324 1998     2700      1     2.572
## 16793   0032560037 4.393 1998     2700      1     2.542
## 16800   0032557174 4.358 1998     2700      1     2.507
## 16801   0032557175 4.491 1998     2700      1     2.535
## 16802   2642713359 4.651 1998     2700      1     2.695
## 16816   0032537068 4.638 1998     2700      1     2.682
## 16819   0032537191 4.969 1998     2700      1     3.118
## 16832   0032522736 4.715 1998     2700      1     2.659
## 16839   0032523018 4.548 1998     2700      1     2.592
## 16851   0032523214 4.715 1998     2700      1     2.659
## 16875   0032499126 4.411 1998     2700      1     2.659
## 16876   0032499141 4.532 1998     2700      1     2.576
## 16910   0008947293 4.653 1998     2700      1     2.697
## 16913   0032473922 4.668 1998     2700      1     2.712
## 17047   0032568257 4.810 1998     2700      1     2.854
## 17053   0032565112 5.107 1998     2700      1     3.151
## 17063   0032554552 4.707 1998     2700      1     2.955
## 17064   0032554569 4.610 1998     2700      1     2.554
## 17091   0032542709 5.010 1998     2700      1     3.054
## 17163   0032485350 5.150 1998     2700      1     3.194
## 17177   0032481673 4.592 1998     2700      1     2.536
## 17289   0032032575 4.574 1998     2700      1     2.618
## 17293   0032033112 4.403 1998     2700      1     2.651
## 17386   0032564902 5.256 1998     2700      1     3.200
## 17401   0032545939 4.480 1998     2700      1     2.524
## 17404   0032546036 4.281 1998     2700      1     2.529
## 17408   0032542516 4.484 1998     2700      1     2.528
```

```
## 17410   0032542541 4.339 1998     2700      1     2.587
## 17420   0032519431 4.715 1998     2700      1     2.659
## 17429   0032519925 5.694 1998     2700      1     3.638
## 17433   0032520153 4.772 1998     2700      1     2.816
## 17444   0032510076 5.618 1998     2700      1     3.562
## 17473   0032484953 4.572 1998     2700      1     2.516
## 17481   0032481152 4.867 1998     2700      1     2.911
## 17482   0032481323 4.638 1998     2700      1     2.682
## 17502   0031987322 4.598 1998     2700      1     2.642
## 17541   0032005791 4.628 1998     2700      1     2.672
## 17595   0032573850 4.468 1998     2700      1     2.512
## 17606   0032556992 4.604 1998     2700      1     2.648
## 17613   0032554096 4.586 1998     2700      1     2.630
## 17627   0032518022 6.088 1998     2700      1     4.132
## 17631   0032518290 5.115 1998     2700      1     3.159
## 17633   0032518420 4.457 1998     2700      1     2.501
## 17644   0032518878 4.711 1998     2700      1     2.755
## 17655   0032515386 4.679 1998     2700      1     2.623
## 17680   0032491837 4.969 1998     2700      1     3.118
## 17686   0032477553 4.484 1998     2700      1     2.528
## 17696   0009440989 4.352 1998     2700      1     2.600
## 18344   0031945285 5.058 1998     2700      1     3.102
## 18436   0031975335 5.254 1998     2700      1     3.403
## 18437   0031975401 4.758 1998     2700      1     2.802
## 20189   0033619959 5.403 1999     2700      1     3.607
## 20216   0033582097 4.593 1999     2700      1     2.692
## 20242   0033572972 5.596 1999     2700      1     3.695
## 20262   0033540004 5.090 1999     2700      1     3.189
## 20278   0033518193 4.549 1999     2700      1     2.548
## 20279   0033518231 5.873 1999     2700      1     3.972
## 20319   0032803699 4.811 1999     2700      1     2.910
## 20484   0033485508 4.616 1999     2700      1     2.715
## 20874   0032748385 5.646 1999     2700      1     3.645
## 20925   0033544340 6.198 1999     2700      1     4.502
## 20962   0033520771 4.299 1999     2700      1     2.503
## 20967   0032589826 4.410 1999     2700      1     2.509
## 21115   0032694263 5.196 1999     2700      1     3.295
## 21120   0032695482 5.349 1999     2700      1     3.448
## 21123   0032741518 4.596 1999     2700      1     2.595
## 21124   0032742418 5.935 1999     2700      1     3.934
## 21125   0032742419 4.335 1999     2700      1     2.639
## 21126   0032743289 5.343 1999     2700      1     3.442
## 21127   0032747280 5.109 1999     2700      1     3.208
## 21134   0033598598 5.651 1999     2700      1     3.855
## 21152   0032696770 5.142 1999     2700      1     3.241
## 21155   0032740874 4.729 1999     2700      1     3.033
## 21255   0033517494 4.412 1999     2700      1     2.616
## 21273   0032855570 4.367 1999     2700      1     2.671
## 21436   0033615307 4.467 1999     2700      1     2.566
## 21443   0033575997 4.748 1999     2700      1     2.847
```

```
## 21498   0033536463 4.644 1999     2700      1     2.743
## 21500   0033536483 4.408 1999     2700      1     2.712
## 21501   0033536501 4.406 1999     2700      1     2.710
## 21548   0032841392 4.784 1999     2700      1     2.883
## 21656   0033199833 4.945 1999     2700      1     3.044
## 21763   0033603792 4.795 1999     2700      1     2.999
## 21775   0342961310 4.661 1999     2700      1     2.965
## 21785   0033584440 5.727 1999     2700      1     3.726
## 21826   0033546647 5.256 1999     2700      1     3.255
## 21990   0033595079 5.033 1999     2700      1     3.032
## 21993   0033595120 4.514 1999     2700      1     2.613
## 22017   0033565321 4.419 1999     2700      1     2.518
## 22021   0033565955 5.009 1999     2700      1     3.008
## 22042   0033542870 5.579 1999     2700      1     3.578
## 22168   0033168952 4.690 1999     2700      1     2.894
## 22237   0033600275 4.833 1999     2700      1     3.137
## 22248   0033031252 4.771 1999     2700      1     2.870
## 22286   0033542393 4.685 1999     2700      1     2.989
## 22303   0033538266 4.432 1999     2700      1     2.531
## 22305   0033538314 4.793 1999     2700      1     2.892
## 22486   0033609374 4.633 1999     2700      1     2.732
## 22491   0033606238 4.793 1999     2700      1     2.892
## 22521   0033562593 4.474 1999     2700      1     2.573
## 22535   0033551370 4.720 1999     2700      1     2.819
## 22563   0033526309 5.723 1999     2700      1     3.822
## 22759   0033611975 4.434 1999     2700      1     2.638
## 22779   0033594375 4.872 1999     2700      1     2.971
## 22793   0033591010 4.634 1999     2700      1     2.733
## 22832   0033553188 4.783 1999     2700      1     3.087
## 22834   0033553194 4.617 1999     2700      1     2.716
## 22836   0033553207 4.636 1999     2700      1     2.735
## 22852   0033531690 4.688 1999     2700      1     2.787
## 22855   0033515827 5.255 1999     1000      2     3.354
## 22909   0033119011 4.749 1999     2700      1     2.953
## 23018   0033577290 4.397 1999     2700      1     2.701
## 23019   0033577336 4.245 1999     2700      1     2.549
## 23056   0033545541 5.796 1999     2700      1     3.895
## 23057   0033545542 4.631 1999     2700      1     2.730
## 23064   0033540993 4.594 1999     2700      1     2.693
## 23067   0033541047 4.233 1999     2700      1     2.537
## 23076   0033522146 5.479 1999     2700      1     3.478
## 23077   0033522206 5.337 1999     2700      1     3.436
## 23111   0033037922 4.231 1999     2700      1     2.535
## 23167   0033104847 4.549 1999     2700      1     2.548
## 23271   0033602049 5.262 1999     2700      1     3.361
## 23292   0033580206 4.603 1999     2700      1     2.807
## 23293   0033580211 4.651 1999     2700      1     2.750
## 23304   0033577041 4.749 1999     2700      1     2.748
## 23314   0033557202 4.498 1999     2700      1     2.597
## 23357   0033540644 4.402 1999     2700      1     2.501
```

```
## 23358   0033540646 5.996 1999     2700      1     4.095
## 23359   0033540680 4.688 1999     2700      1     2.687
## 23360   0033540714 4.973 1999     2700      1     3.177
## 23361   0033540719 4.883 1999     2700      1     2.882
## 23541   0033608182 4.940 1999     2700      1     2.939
## 23544   0033608191 4.549 1999     2700      1     2.648
## 23560   0033590516 4.464 1999     2700      1     2.563
## 23561   0033590525 4.644 1999     2700      1     2.743
## 23567   0033585499 4.878 1999     2700      1     2.977
## 23577   0032918414 4.830 1999     2700      1     3.134
## 23584   0032954778 5.005 1999     2700      1     3.104
## 23591   0033552854 4.631 1999     2700      1     2.935
## 23605   0033537343 4.731 1999     2700      1     2.830
## 23621   0033528072 4.427 1999     2700      1     2.526
## 23623   0033528101 4.622 1999     2700      1     2.721
## 23860   0032701521 4.971 1999     2700      1     3.070
## 23861   0032701642 4.548 1999     2700      1     2.647
## 23882   0032713075 5.527 1999     2700      1     3.626
## 23914   0032722149 4.823 1999     2700      1     3.027
## 23925   0032725185 4.746 1999     2700      1     2.845
## 23974   0032742992 5.265 1999     2700      1     3.364
## 23997   0032746174 4.892 1999     2700      1     2.991
## 24023   0032752579 4.462 1999     2700      1     2.561
## 24083   0032824349 4.787 1999     2700      1     2.886
## 24251   0032919402 4.464 1999     2700      1     2.563
## 24424   0033013684 4.775 1999     2700      1     2.774
## 26283   0034055392 0.000 2000     2700      1    -2.554
## 26314  24044533275 0.000 2000     1306      3    -2.759
## 26347  85018888882 0.000 2000     2700      3    -2.554
## 26354   0034735827 5.241 2000     2700      1     2.787
## 26368   0034626988 5.215 2000     2700      1     2.556
## 26410   0034305633 0.000 2000     2700      2    -2.659
## 26509  33746530366 0.000 2000     2700      1    -2.659
## 26516   0033853362 0.000 2000     2700      1    -2.659
## 26526   0034226622 0.000 2000     1300      2    -2.659
## 26692   0034048842 0.000 2000     2700      1    -2.659
## 26889  85008036014 0.000 2000     1704      3    -2.659
## 26890  85048650487 0.000 2000     2700      1    -2.659
## 26910  12944252960 0.000 2000     2700      1    -2.659
## 26948  85047454626 0.000 2000     2700      1    -2.659
## 26964   0034732201 5.487 2000     2700      1     2.828
## 26988   0034655836 0.000 2000     2700      1    -2.659
## 27018   0034102286 0.000 2000     2700      1    -2.659
## 27031   0034682247 5.466 2000     2700      1     2.707
## 27110   0035956483 5.056 2001     2700      1     3.100
## 27142   0035924765 5.181 2001     2700      1     3.225
## 27205   0035857147 4.489 2001     2700      1     2.633
## 27210   0035852034 5.010 2001     2700      1     3.054
## 27211   0035852037 5.008 2001     2700      1     3.152
## 28065   0035944839 4.480 2001     2700      1     2.828
```

```
## 28077   0035936007 4.392 2001     2700      1     2.536
## 28092   0035930101 4.593 2001     2700      1     2.737
## 28118   0035904776 4.828 2001     2700      1     2.972
## 28120   0035892018 5.798 2001     2700      1     3.942
## 28145   0035861037 4.689 2001     2700      1     2.733
## 28177   0035829840 4.821 2001     2700      1     3.169
## 28179   0035829842 6.472 2001     2700      1     4.616
## 28353   0035522330 5.300 2001     2700      1     3.444
## 28468   0035950685 4.432 2001     2700      1     2.576
## 28564   0035846315 4.580 2001     2700      1     2.724
## 28628   0035802216 4.544 2001     2700      1     2.588
## 28791   0035960116 4.693 2001     2700      1     2.837
## 28832   0035913589 5.497 2001     2700      1     3.641
## 28851   0035856017 5.495 2001     2700      1     3.639
## 28900   0035818048 5.736 2001     2700      1     3.880
## 29126   0035975408 4.632 2001     2700      1     2.776
## 29159   0035940039 4.392 2001     2700      1     2.536
## 29160   0035940040 4.567 2001     2700      1     2.711
## 29182   0034898468 4.742 2001     2700      1     2.786
## 29248   0035827865 4.721 2001     2700      1     2.865
## 29267   0034926528 4.251 2001     2700      1     2.599
## 29299   0035413919 4.889 2001     2700      1     3.237
## 29405   0035963919 4.481 2001     2700      1     2.625
## 29417   0035954660 4.486 2001     2700      1     2.630
## 29418   0035954670 4.624 2001     2700      1     2.768
## 29428   0035948630 4.447 2001     2700      1     2.591
## 29439   0034939686 4.476 2001     2700      1     2.620
## 29458   0035928419 4.926 2001     2700      1     3.274
## 29472   0035913228 4.532 2001     2700      1     2.676
## 29523   0035849988 4.543 2001     2700      1     2.687
## 29524   0035850010 5.106 2001     2700      1     3.150
## 29529   0034933989 4.909 2001     2700      1     3.158
## 29532   0034945516 5.330 2001     2700      1     3.474
## 29705   0035963494 4.396 2001     2700      1     2.540
## 29708   0035963529 5.289 2001     2700      1     3.433
## 29762   0035919183 4.463 2001     2700      1     2.507
## 29763   0035919184 4.459 2001     2700      1     2.503
## 29766   0035897882 4.216 2001     2700      1     2.564
## 29767   0035897888 4.789 2001     2700      1     2.933
## 29776   0035859201 4.805 2001     2700      1     3.054
## 29778   0035859233 4.976 2001     2700      1     3.324
## 29787   0035854054 5.981 2001     2700      1     4.329
## 29826   0035816007 5.037 2001     2700      1     3.181
## 29832   0035816032 5.067 2001     2700      1     3.211
## 30038   0035902194 4.674 2001     2700      1     2.818
## 30046   0035897686 4.612 2001     2700      1     2.756
## 30047   0035897688 5.037 2001     2700      1     3.181
## 30084   0035832261 6.117 2001     2700      1     4.261
## 30089   0035832341 4.478 2001     2700      1     2.826
## 30856   0035953646 4.483 2001     2700      1     2.831
```

```
## 30872   0035946664 5.610 2001     2700      1     3.654
## 30947   0035857967 4.501 2001     2700      1     2.645
## 30954   0035849144 4.588 2001     2700      1     2.936
## 30967   0035843629 4.350 2001     2700      1     2.698
## 31007   0035804852 4.734 2001     2700      1     2.878
## 31008   0035804857 4.620 2001     2700      1     2.869
## 31151   0035962077 4.916 2001     2700      1     3.264
## 31177   0035925774 4.765 2001     2700      1     3.014
## 31204   0035869435 4.410 2001     2700      1     2.758
## 31220   0035857786 4.465 2001     2700      1     2.813
## 31237   0035835981 4.789 2001     2700      1     2.933
## 31257   0035826089 5.495 2001     2700      1     3.639
## 31264   0035819910 4.759 2001     2700      1     2.803
## 31434   0035961566 5.156 2001     2700      1     3.300
## 31461   0035931959 4.840 2001     2700      1     2.884
## 31463   0035931973 4.682 2001     2700      1     2.826
## 31477   0035925169 4.717 2001     2700      1     2.761
## 31485   0035901051 5.395 2001     2700      1     3.539
## 31491   0035864943 4.678 2001     2700      1     2.822
## 31492   0035864977 4.470 2001     2700      1     2.614
## 31541   0035825726 5.310 2001     2700      1     3.454
## 31569   0035819481 4.496 2001     2700      1     2.540
## 31715   0035945667 4.738 2001     2700      1     2.882
## 31734   0035915688 4.647 2001     2700      1     2.691
## 31742   0035905889 4.581 2001     2700      1     2.725
## 31744   0035905908 4.896 2001     2700      1     3.040
## 31773   0035843243 5.197 2001     2700      1     3.446
## 31775   0035843254 4.324 2001     2700      1     2.672
## 31782   0035835027 4.362 2001     2700      1     2.611
## 31783   0035835029 4.403 2001     2700      1     2.547
## 31795   0035814656 4.673 2001     2700      1     2.817
## 31799   0035804277 5.895 2001     2700      1     4.039
## 31801   0035804283 4.647 2001     2700      1     2.791
## 31933   0034795654 4.634 2001     2700      1     2.778
## 31937   0034799675 4.729 2001     2700      1     2.873
## 32022   0034857255 4.388 2001     2700      1     2.532
## 32059   0034903710 4.656 2001     2700      1     2.800
## 32248   0035013499 4.468 2001     2700      1     2.512
## 32609   0035174173 4.507 2001     2700      1     2.651
## 33892   0035960866 5.082 2001     2700      1     3.126
## 34423   0037132575 4.351 2002     2700      1     2.973
## 34424   0037132651 3.900 2002     2700      1     2.627
## 34438   0037079352 4.660 2002     2700      1     3.183
## 34445   0037069763 4.294 2002     2700      1     2.916
## 34446   0037069783 4.449 2002     2700      1     3.276
## 34454   0037065306 5.195 2002     2700      1     3.817
## 34457   0037065340 4.209 2002     2700      1     2.831
## 34467   0037028001 4.145 2002     2700      1     2.767
## 34474   0037021538 3.880 2002     2700      1     2.502
## 34477   0037021550 3.922 2002     2700      1     2.544
```

```
## 34553   0036908576 3.908 2002     2700      1     2.530
## 34596   0036950665 4.266 2002     2700      1     2.888
## 34958   0037202854 4.238 2002     2700      1     2.860
## 34973   0037191726 4.195 2002     2700      1     2.817
## 34977   0037184416 4.294 2002     2700      1     2.916
## 34979   0037184418 4.896 2002     2700      1     3.518
## 35001   0037153024 4.533 2002     2700      1     3.155
## 35004   0037153037 4.517 2002     2700      1     3.139
## 35005   0037153042 5.145 2002     2700      1     3.872
## 35013   0037145856 5.083 2002     2700      1     3.705
## 35014   0037145863 5.578 2002     2700      1     4.200
## 35032   0037078969 5.588 2002     2700      1     4.415
## 35039   0037073256 4.972 2002     2700      1     3.699
## 35043   0037073289 4.418 2002     2700      1     2.941
## 35079   0037032401 4.648 2002     2700      1     3.270
## 35130   0036840210 3.984 2002     2700      1     2.606
## 35358   0037206364 5.141 2002     2700      1     3.763
## 35359   0037206368 4.285 2002     2700      1     2.907
## 35370   0037180233 4.412 2002     2700      1     3.034
## 35377   0037167982 4.361 2002     2700      1     2.983
## 35385   0037164045 4.478 2002     2700      1     3.001
## 35386   0037164054 5.560 2002     2700      1     4.083
## 35397   0037126201 4.294 2002     2700      1     2.916
## 35398   0037126332 4.518 2002     2700      1     3.041
## 35400   0037126341 5.773 2002     2700      1     4.395
## 35408   0037120864 4.463 2002     2700      1     3.085
## 35411   0037120907 4.813 2002     2700      1     3.640
## 35446   0037048640 4.161 2002     2700      1     2.783
## 35448   0037048669 6.012 2002     2700      1     4.535
## 35451   0037048695 5.664 2002     2700      1     4.187
## 35452   0037048697 4.170 2002     2700      1     2.897
## 35457   0037027050 4.327 2002     2700      1     3.154
## 35462   0037015543 4.380 2002     2700      1     3.207
## 35481   0036771775 4.268 2002     2700      1     2.791
## 35699   0037179578 4.011 2002     2700      1     2.633
## 35709   0037174346 4.951 2002     2700      1     3.573
## 35730   0037130645 3.986 2002     2700      1     2.608
## 35758   0037063394 3.715 2002     2700      1     2.542
## 35759   0037063403 4.739 2002     2700      1     3.466
## 35760   0037063415 4.234 2002     2700      1     2.856
## 35778  18644385210 4.777 2002     2700      1     3.300
## 35867   0036733312 4.085 2002     2700      1     2.707
## 35911   0036745457 3.816 2002     2700      1     2.643
## 36123   0037190086 4.360 2002     2700      1     2.982
## 36181   0037158614 4.762 2002     2700      1     3.384
## 36244   0037125569 4.079 2002     2700      1     2.602
## 36254   0037103420 4.673 2002     2700      1     3.295
## 36269   0037077489 4.481 2002     2700      1     3.004
## 36270   0037077490 4.958 2002     2700      1     3.785
## 36292   0037043652 4.862 2002     2700      1     3.385
```

```
## 36295   0037043682 4.263 2002     2700      1     2.990
## 36306   0037036787 4.703 2002     2700      1     3.325
## 36318   0037014941 4.189 2002     2700      1     2.712
## 36365   0036681988 5.232 2002     2700      1     3.854
## 36367   0036682163 4.454 2002     2700      1     3.076
## 36369   0036682267 4.857 2002     2700      1     3.479
## 36524   0037178579 4.374 2002     2700      1     2.897
## 36554   0037166985 4.791 2002     2700      1     3.413
## 36555   0037166986 4.893 2002     2700      1     3.515
## 36583   0037130113 4.121 2002     2700      1     2.848
## 36586   0037130183 4.068 2002     2700      1     2.690
## 36593   0037125379 6.631 2002     2700      1     5.458
## 36594   0037125417 3.979 2002     2700      1     2.601
## 36597   0037125454 4.380 2002     2700      1     3.207
## 36619   0037055011 4.224 2002     2700      1     2.846
## 36620   0037055016 4.251 2002     2700      1     2.873
## 36658   0037019322 4.383 2002     2700      1     3.005
## 36666   0037014584 5.135 2002     2700      1     3.862
## 36667   0037014630 4.589 2002     2700      1     3.211
## 36668   0037014644 4.131 2002     2700      1     2.753
## 36692   0035983642 4.125 2002     2700      1     2.747
## 36821   0037182768 4.258 2002     2700      1     2.985
## 36860   0037134824 4.216 2002     2700      1     2.739
## 36864   0037134872 4.163 2002     2700      1     2.686
## 36865   0037134927 3.726 2002     2700      1     2.553
## 36867   0037134933 3.807 2002     2700      1     2.534
## 36901   0037071793 4.430 2002     2700      1     3.157
## 36904   0037071808 5.799 2002     2700      1     4.421
## 36910   0037067128 4.048 2002     2700      1     2.670
## 36912   0037067150 4.184 2002     2700      1     2.806
## 36940   0037030694 3.959 2002     2700      1     2.581
## 36956   0037024313 4.125 2002     2700      1     2.952
## 36996   0036595171 4.043 2002     2700      1     2.566
## 37144   0037198420 4.789 2002     2700      1     3.411
## 37146   0037198432 5.119 2002     2700      1     3.946
## 37175   0036251741 4.327 2002     2700      1     3.154
## 37181   0037162115 4.877 2002     2700      1     3.499
## 37182   0037162116 5.260 2002     2700      1     3.882
## 37188   0037157095 3.984 2002     2700      1     2.606
## 37191   0037157169 3.999 2002     2700      1     2.621
## 37208   0037129633 4.406 2002     2700      1     3.028
## 37241   0037092912 4.464 2002     2700      1     3.086
## 37243   0037092914 4.102 2002     2700      1     2.724
## 37246   0037093012 4.590 2002     2700      1     3.417
## 37286   0037042231 4.892 2002     2700      1     3.514
## 37288   0037042250 3.905 2002     2700      1     2.632
## 37305   0037018763 3.994 2002     2700      1     2.616
## 37309   0037007679 4.195 2002     2700      1     2.817
## 37311   0037007683 4.588 2002     2700      1     3.111
## 37362   0036570099 4.672 2002     2700      1     3.195
```

```
## 37452   0037181697 3.871 2002     2700      1     2.698
## 37453   0037181727 4.139 2002     2700      1     2.662
## 37462   0037171703 4.324 2002     2700      1     2.946
## 37475   0037165667 4.466 2002     2700      1     2.989
## 37476   0037165703 4.734 2002     2700      1     3.356
## 37477   0037165713 4.125 2002     2700      1     2.747
## 37478   0037165728 4.888 2002     2700      1     3.510
## 37479   0037165740 5.429 2002     2700      1     4.156
## 37527   0037123168 4.123 2002     2700      1     2.646
## 37528   0037123206 4.022 2002     2700      1     2.849
## 37570   0037061912 5.703 2002     2700      1     4.325
## 37581   0037051930 4.197 2002     2700      1     2.819
## 37582   0037051975 4.690 2002     2700      1     3.517
## 37626   0037018496 4.615 2002     2700      1     3.342
## 37638   0037012465 4.786 2002     2700      1     3.408
## 37640   0037012476 4.433 2002     2700      1     3.055
## 37744   0037197039 3.880 2002     2700      1     2.502
## 37745   0037197044 4.518 2002     2700      1     3.140
## 37752   0037187907 5.013 2002     2700      1     3.635
## 37753   0037187926 4.094 2002     2700      1     2.716
## 37758   0037181496 4.070 2002     2700      1     2.692
## 37759   0037181511 4.841 2002     2700      1     3.463
## 37760   0037181515 4.728 2002     2700      1     3.350
## 37778   0036119735 4.484 2002     2700      1     3.106
## 37782   0036122659 4.001 2002     2700      1     2.524
## 37790   0037160936 4.064 2002     2700      1     2.686
## 37819   0037149714 4.285 2002     2700      1     2.907
## 37826   0037139438 4.401 2002     2700      1     3.023
## 37868   0037076025 5.024 2002     2700      1     3.751
## 37875   0037070525 4.224 2002     2700      1     2.846
## 37876   0037070528 4.231 2002     2700      1     3.058
## 37900   0037035121 4.647 2002     2700      1     3.374
## 37913   0037029069 3.902 2002     2700      1     2.524
## 38083   0037181148 4.810 2002     2700      1     3.432
## 38109   0037148928 4.712 2002     2700      1     3.235
## 38143   0037116832 5.558 2002     2700      1     4.180
## 38145   0037116836 4.256 2002     2700      1     2.779
## 38161   0037075257 5.488 2002     2700      1     4.011
## 38163   0037075268 3.927 2002     2700      1     2.654
## 38164   0037075269 3.951 2002     2700      1     2.573
## 38165   0037075271 4.109 2002     2700      1     2.632
## 38172   0037070259 4.711 2002     2700      1     3.333
## 38192   0036155211 4.109 2002     2700      1     2.936
## 38200   0037045561 4.075 2002     2700      1     2.697
## 38208   0037034257 6.602 2002     2700      1     5.224
## 38217   0037028761 4.268 2002     2700      1     2.995
## 38218   0037028766 3.946 2002     2700      1     2.773
## 38220   0037028769 4.250 2002     2700      1     2.872
## 38247   0037006397 4.366 2002     2700      1     2.988
## 38328   0037204196 4.718 2002     2700      1     3.340
```

```
## 38357   0037160567 3.981 2002     2700      1     2.603
## 38367   0037132805 3.880 2002     2700      1     2.707
## 38375   0037122950 4.667 2002     2700      1     3.394
## 38377   0037122987 3.981 2002     2700      1     2.603
## 38381   0037116641 6.031 2002     2700      1     4.653
## 38382   0037116642 5.093 2002     2700      1     3.715
## 38383   0037116647 4.485 2002     2700      1     3.212
## 38384   0037116658 3.807 2002     2700      1     2.634
## 38390   0037065519 4.048 2002     2700      1     2.875
## 38398   0037050352 5.876 2002     2700      1     4.399
## 38407   0037045439 4.562 2002     2700      1     3.184
## 38421   0037012134 4.436 2002     2700      1     3.058
## 38430   0037006110 4.025 2002     2700      1     2.548
## 38432   0037006117 3.976 2002     2700      1     2.598
## 38529   0036083829 4.036 2002     2700      1     2.658
## 38535   0036086409 4.229 2002     2700      1     2.752
## 38536   0036086413 3.999 2002     2700      1     2.726
## 38553   0036106054 4.887 2002     2700      1     3.714
## 38572   0036123845 3.948 2002     2700      1     2.570
## 38606   0036143018 4.750 2002     2700      1     3.273
## 38662   0036168150 3.908 2002     2700      1     2.635
## 38755   0036210052 4.381 2002     2700      1     3.108
## 38789   0036239507 4.691 2002     2700      1     3.214
## 39652   0036731990 3.940 2002     2700      1     2.767
## 39758   0036800759 3.956 2002     2700      1     2.578
## 39896   0036894541 4.121 2002     2700      1     2.743
## 40605   0346030568 5.028 2003     2700      1     3.817
## 40608   0347882750 5.233 2003     2700      1     3.917
## 40610   0348011697 4.079 2003     2700      1     2.763
## 40638  10744222633 4.103 2003     2700      1     2.787
## 40646   0344235454 4.984 2003     2700      1     3.873
## 40657   0344738736 4.084 2003     2700      1     2.973
## 40659   0345168925 4.208 2003     2700      1     2.892
## 40660   0345492466 4.735 2003     2700      1     3.419
## 40682   0345690179 5.350 2003     2700      1     4.034
## 40689   0344630326 3.954 2003     2700      1     2.638
## 40731   0346340042 4.157 2003     2700      1     2.841
## 40749   0346969977 4.442 2003     2700      1     3.231
## 40771   0348230958 6.275 2003     2700      1     4.959
## 41119   0345283207 3.840 2003     2700      1     2.524
## 41122   0345714860 5.499 2003     2700      1     4.084
## 41127   0344081179 4.283 2003     2700      1     2.868
## 41128   0344861847 4.621 2003     2700      1     3.305
## 41129   0345293127 3.833 2003     2700      1     2.517
## 41131   0345293130 4.162 2003     2700      1     3.051
## 41139   0345414665 4.343 2003     2700      1     2.928
## 41148   0344926414 5.006 2003     2700      1     3.895
## 41151   0642278662 3.806 2003     2700      1     2.595
## 41156   0344943245 4.973 2003     2700      1     3.657
## 41157   0345374595 4.903 2003     2700      1     3.587
```

```
## 41159   0642272544 4.675 2003     2700      1     3.260
## 41161   0242551979 3.904 2003     2700      1     2.588
## 41176   0242658928 4.458 2003     2700      1     3.142
## 41183   0242410368 4.973 2003     2700      1     3.657
## 41185   0242493774 4.400 2003     2700      1     2.985
## 41210   0242493661 4.273 2003     2700      1     2.957
## 41237   0344444858 4.520 2003     2700      1     3.105
## 41251   0345824713 4.352 2003     2700      1     3.036
## 41258   0346969978 5.002 2003     2700      1     3.587
## 41458   0142025460 3.894 2003     2700      1     2.578
## 41460   0142089171 4.888 2003     2700      1     3.473
## 41463   0142123411 4.222 2003     2700      1     2.906
## 41468   0142087597 4.833 2003     2700      1     3.517
## 41484   0142009533 5.306 2003     2700      1     3.891
## 41493   0142024742 4.567 2003     2700      1     3.152
## 41496   0142088521 4.407 2003     2700      1     3.091
## 41531   0141863194 4.217 2003     2700      1     2.901
## 41575   0141653838 4.222 2003     2700      1     2.906
## 41587   0141705375 4.884 2003     2700      1     3.469
## 41593   0141816759 4.768 2003     2700      1     3.353
## 41594   0141816761 4.287 2003     2700      1     3.076
## 41647   0242285697 4.478 2003     2700      1     3.162
## 41655   0242380323 3.866 2003     2700      1     2.755
## 41771   0141425718 4.211 2003     2700      1     2.796
## 41772   0141484564 4.167 2003     2700      1     2.851
## 41777   0141462439 4.563 2003     2700      1     3.148
## 41778   0141573545 5.008 2003     2700      1     3.593
## 41779   0141573546 3.904 2003     2700      1     2.793
## 41781   0141685592 4.301 2003     2700      1     2.886
## 41782   0141796735 4.455 2003     2700      1     3.244
## 41783   0141796739 4.693 2003     2700      1     3.377
## 41807   0141611906 4.138 2003     2700      1     2.822
## 41813   0141834950 4.316 2003     2700      1     3.000
## 41829   0042387879 4.992 2003     2700      1     3.577
## 41838   0042413423 4.335 2003     2700      1     3.124
## 41850   0042360213 5.273 2003     2700      1     3.957
## 41855   0041381147 3.891 2003     2700      1     2.780
## 41856   0041381153 4.095 2003     2700      1     2.779
## 41858   0042884162 3.758 2003     2700      1     2.647
## 41886   0141612004 3.855 2003     2700      1     2.539
## 41945   0142217904 3.947 2003     2700      1     2.532
## 42124   0041327804 5.015 2003     2700      1     3.699
## 42140   0042422040 4.594 2003     2700      1     3.483
## 42142   0042922806 4.186 2003     2700      1     2.975
## 42150   0041464854 4.417 2003     2700      1     3.101
## 42171   0042195833 4.746 2003     2700      1     3.430
## 42172   0042195862 4.003 2003     2700      1     2.792
## 42175   0042697063 4.958 2003     2700      1     3.642
## 42188   0041912565 4.113 2003     2700      1     2.797
## 42190   0042413567 4.065 2003     2700      1     2.650
```

```
## 42191   0042914547 3.963 2003     2700      1     2.852
## 42203   0042661252 4.135 2003     2700      1     2.720
## 42210   0041735992 4.538 2003     2700      1     3.222
## 42216   0042093742 5.388 2003     2700      1     4.177
## 42217   0042594633 4.172 2003     2700      1     2.856
## 42222   0041708062 4.398 2003     2700      1     3.287
## 42223   0041708064 4.291 2003     2700      1     3.080
## 42228   0042125511 4.751 2003     2700      1     3.336
## 42236   0041385951 4.038 2003     2700      1     2.927
## 42246   0041886776 3.994 2003     2700      1     2.678
## 42267   0042888943 4.611 2003     2700      1     3.295
## 42397   0041842634 4.262 2003     2700      1     2.946
## 42398   0042343801 5.646 2003     2700      1     4.330
## 42404   0041698076 4.135 2003     2700      1     2.720
## 42412   0038497523 4.054 2003     2700      1     2.843
## 42420   0038455694 4.334 2003     2700      1     3.018
## 42438   0038825532 3.870 2003     2700      1     2.554
## 42444   0038601952 4.724 2003     2700      1     3.613
## 42449   0038304776 5.520 2003     2700      1     4.204
## 42452   0038679758 4.587 2003     2700      1     3.172
## 42462   0038013919 4.018 2003     2700      1     2.702
## 42464   0038352047 3.855 2003     2700      1     2.744
## 42466   0038690424 5.414 2003     2700      1     4.098
## 42467   0038690437 4.549 2003     2700      1     3.338
## 42489   0038504054 4.322 2003     2700      1     2.907
## 42515   0041302382 4.281 2003     2700      1     2.965
## 42539   0042304086 4.467 2003     2700      1     3.052
## 42551   0042743965 3.866 2003     2700      1     2.550
## 42564   0043245147 4.497 2003     2700      1     3.181
## 42684   0037830097 3.972 2003     2700      1     2.656
## 42690   0037973279 5.594 2003     2700      1     4.278
## 42691   0038482206 5.915 2003     2700      1     4.704
## 42699   0038167811 4.243 2003     2700      1     3.132
## 42713   0038130715 5.029 2003     2700      1     3.713
## 42719   0037493499 4.842 2003     2700      1     3.526
## 42722   0038507413 4.230 2003     2700      1     3.119
## 42729   0038434032 3.859 2003     2700      1     2.543
## 42736   0037840165 4.892 2003     2700      1     3.576
## 42745   0038242951 4.006 2003     2700      1     2.690
## 42746   0038242968 5.038 2003     2700      1     3.722
## 42754   0037805278 4.410 2003     2700      1     2.995
## 42756   0038142845 4.115 2003     2700      1     2.904
## 42757   0038142850 3.852 2003     2700      1     2.741
## 42772   0037986208 4.721 2003     2700      1     3.405
## 42773   0038323914 4.575 2003     2700      1     3.259
## 42780   0038314212 3.960 2003     2700      1     2.644
## 42893   0042532322 4.480 2003     2700      1     3.164
## 43069   0038526363 4.506 2003     2700      1     3.091
## 43074   0037638884 5.206 2003     2700      1     3.890
## 43077   0038724280 4.848 2003     2700      1     3.737
```

```
## 43085   0037737900 5.165 2003     2700      1     3.849
## 43086   0037737901 3.822 2003     2700      1     2.506
## 43088   0037947320 3.982 2003     2700      1     2.666
## 43089   0038075468 5.272 2003     2700      1     3.956
## 43091   0038751994 4.929 2003     2700      1     3.818
## 43115   0038663167 4.361 2003     2700      1     3.045
## 43116   0038663174 4.006 2003     2700      1     2.895
## 43129   0038316599 4.797 2003     2700      1     3.382
## 43130   0038655478 3.979 2003     2700      1     2.868
## 43132   0037906573 4.466 2003     2700      1     3.051
## 43138   0012868624 4.671 2003     2700      1     3.460
## 43146   0038824056 4.162 2003     2700      1     3.051
## 43153   0038298787 4.370 2003     2700      1     3.054
## 43194   0037988905 4.148 2003     2700      1     2.832
## 43240   0038651907 4.577 2003     2700      1     3.261
## 43241   0038651918 3.988 2003     2700      1     2.672
## 43286   0242600542 4.485 2003     2700      1     3.070
## 43287   0242684416 5.438 2003     2700      1     4.227
## 43416   0037464536 4.084 2003     2700      1     2.873
## 43420   0345636017 4.427 2003     2700      1     3.111
## 43428   0038528237 4.980 2003     2700      1     3.664
## 43435   0037451905 5.245 2003     2700      1     3.929
## 43438   0037451929 6.034 2003     2700      1     4.718
## 43443   0037448932 3.966 2003     2700      1     2.650
## 43479   0037502809 5.082 2003     2700      1     3.971
## 43481   0037840394 4.140 2003     2700      1     3.029
## 43483   0037986313 5.056 2003     2700      1     3.740
## 43485   0038516861 4.525 2003     2700      1     3.314
## 43498   0037417219 4.629 2003     2700      1     3.418
## 43510   0037414165 3.901 2003     2700      1     2.585
## 43511   0037414194 4.269 2003     2700      1     2.953
## 43512   0037414217 4.123 2003     2700      1     2.807
## 43552   0037396734 4.424 2003     2700      1     3.108
## 43561   0037399439 5.125 2003     2700      1     3.809
## 43733   0037468682 3.829 2003     2700      1     2.718
## 43734   0037468691 4.429 2003     2700      1     3.113
## 43739   0037467321 3.688 2003     2700      1     2.577
## 43766   0037454273 4.160 2003     2700      1     2.844
## 43767   0037454282 4.138 2003     2700      1     2.723
## 43768   0037454283 4.138 2003     2700      1     2.822
## 43785   0037434851 3.802 2003     2700      1     2.691
## 43787   0037434895 4.553 2003     2700      1     3.237
## 43807   0037420270 4.012 2003     2700      1     2.696
## 43808   0037420274 5.142 2003     2700      1     3.826
## 43884   0037364369 4.444 2003     2700      1     3.128
## 44099   0037468409 3.866 2003     2700      1     2.550
## 44100   0037468422 4.369 2003     2700      1     3.053
## 44104   0037466917 4.108 2003     2700      1     2.693
## 44120   0037456351 4.578 2003     2700      1     3.262
## 44121   0037456358 4.402 2003     2700      1     3.086
```

```
## 44125   0037453950 4.520 2003     2700      1     3.204
## 44127   0037453976 4.581 2003     2700      1     3.166
## 44133   0037440670 3.950 2003     2700      1     2.634
## 44150   0037434515 4.739 2003     2700      1     3.423
## 44159   0037433178 4.012 2003     2700      1     2.597
## 44160   0037433181 3.866 2003     2700      1     2.550
## 44166   0037425790 3.859 2003     2700      1     2.648
## 44178   0037419908 4.233 2003     2700      1     2.917
## 44182   0037289753 4.297 2003     2700      1     2.981
## 44192   0037308404 5.103 2003     2700      1     3.787
## 44201   0037311377 4.989 2003     2700      1     3.574
## 44212   0037313797 3.966 2003     2700      1     2.650
## 44261   0037330518 3.985 2003     2700      1     2.669
## 44428   0037438792 4.376 2003     2700      1     3.060
## 44429   0037438809 4.264 2003     2700      1     3.053
## 44432   0037438936 3.985 2003     2700      1     2.669
## 44433   0037438998 4.038 2003     2700      1     2.722
## 44454   0037425535 5.533 2003     2700      1     4.118
## 44455   0037425558 4.402 2003     2700      1     3.086
## 44456   0037425564 5.704 2003     2700      1     4.388
## 44457   0037425578 5.336 2003     2700      1     4.020
## 44461   0037417523 5.839 2003     2700      1     4.628
## 44468   0037413484 4.422 2003     2700      1     3.211
## 44469   0037413492 5.094 2003     2700      1     3.883
## 44472   0037413628 3.962 2003     1000      2     2.547
## 44490   0037216353 5.144 2003     2700      1     3.828
## 44501   0037218814 4.737 2003     2700      1     3.421
## 44525   0037237980 3.908 2003     2700      1     2.592
## 44529   0037241112 4.447 2003     2700      1     3.131
## 44541   0037248459 3.918 2003     2700      1     2.707
## 44543   0037249283 3.684 2003     2700      1     2.573
## 44546   0037251810 4.453 2003     2700      1     3.137
## 44690   0037330279 4.140 2003     2700      1     2.824
## 44719   0037382166 3.982 2003     2700      1     2.666
## 44720   0037382807 4.266 2003     2700      1     2.950
## 44745   0037498595 4.430 2003     2700      1     3.219
## 44793   0037685168 4.190 2003     2700      1     2.874
## 44806   0037716856 4.108 2003     2700      1     2.792
## 44889   0038165476 3.954 2003     2700      1     2.539
## 44918   0038303393 3.711 2003     2700      1     2.600
## 45018   0038731023 4.110 2003     2700      1     2.695
## 45080   0041878535 4.148 2003     2700      1     2.832
## 45330   0242515752 3.911 2003     2700      1     2.595
## 45365   0345374590 4.847 2003     2700      1     3.531
## 45366   0345374591 4.435 2003     2700      1     3.324
## 45373   0345690174 3.988 2003     2700      1     2.672
## 45403   0346843100 3.991 2003     2700      1     2.675
## 45422   0347986777 4.410 2003     2700      1     3.094
## 46148   0042737443 4.087 2003     2700      1     2.672
## 46153   0141499228 4.645 2003     2700      1     3.434
```

```
## 46333  11144239923 5.214 2004     2700      1     3.986
## 46334  19944395825 4.649 2004     2700      1     3.421
## 46341  10844262691 4.397 2004     2700      1     3.374
## 46342  10844281210 4.459 2004     2700      1     3.231
## 46367  10344231441 4.676 2004     2700      1     3.348
## 46368  10344239881 4.585 2004     2700      1     3.357
## 46369  10344242939 4.474 2004     2700      1     3.246
## 46375  10344221567 4.188 2004     2700      1     2.960
## 46376  10344229444 4.159 2004     2700      1     3.036
## 46381  11344275794 3.847 2004     2700      1     2.619
## 46424  19744365355 3.936 2004     2700      1     2.913
## 46467  10044271620 4.226 2004     2700      1     2.998
## 46589  15244363856 4.164 2004     2700      1     2.836
## 46593  15744363493 3.986 2004     2700      1     2.863
## 46597  15744399190 4.100 2004     2700      1     2.872
## 46820   9644252909 4.467 2004     2700      1     3.239
## 47011   9244240268 3.964 2004     2700      1     2.736
## 47018   9244240769 4.129 2004     2700      1     2.801
## 47019   9244247612 4.578 2004     2700      1     3.555
## 47020   9244264413 3.803 2004     2700      1     2.575
## 47044   8544252449 4.680 2004     2700      1     3.557
## 47045   8544258102 4.526 2004     2700      1     3.403
## 47053   8444225132 4.176 2004     2700      1     2.948
## 47062  19644400578 4.870 2004     2700      1     3.542
## 47069   7744231805 4.691 2004     2700      1     3.463
## 47070   7744237066 4.382 2004     2700      1     3.154
## 47072   7744239901 4.468 2004     2700      1     3.345
## 47085   7644221218 3.947 2004     2700      1     2.824
## 47140  13844273390 4.129 2004     2700      1     3.006
## 47147  13944251605 4.572 2004     2700      1     3.344
## 47168  15244339164 4.231 2004     2700      1     3.003
## 47393   6944244875 4.867 2004     2700      1     3.639
## 47395   6944248992 4.460 2004     2700      1     3.232
## 47399   6944229468 4.497 2004     2700      1     3.269
## 47407   6944232728 4.311 2004     2700      1     3.083
## 47443   5044248578 3.831 2004     2700      1     2.603
## 47458  19544385747 5.348 2004     2700      1     4.120
## 47601   7444240833 3.982 2004     2700      1     2.859
## 47764   4544305451 4.264 2004     2700      1     3.241
## 47790   4544279029 4.188 2004     2700      1     2.960
## 47806   4544332903 5.079 2004     2700      1     3.851
## 47815   4544222914 3.776 2004     2700      1     2.548
## 47825   4444358501 4.553 2004     2700      1     3.325
## 47843   4544329012 3.814 2004     2700      1     2.586
## 47851   4344690525 4.410 2004     2700      1     3.082
## 47995   4944239035 4.273 2004     2700      1     3.045
## 48006   5444255241 3.855 2004     2700      1     2.627
## 48052   8544244084 4.077 2004     2700      1     2.849
## 48054   9644268242 3.755 2004     2700      1     2.527
## 48197   4143052707 4.043 2004     2700      1     2.815
```

```
## 48207   4143067005 4.774 2004     2700      1     3.546
## 48208   4143069253 3.924 2004     2700      1     2.696
## 48227   4143071570 4.102 2004     2700      1     2.774
## 48228   4143094988 5.453 2004     2700      1     4.225
## 48236   4043076922 3.999 2004     2700      1     2.771
## 48238   4043153049 4.489 2004     2700      1     3.261
## 48239   4043156247 4.890 2004     2700      1     3.662
## 48367   4043055316 4.632 2004     2700      1     3.404
## 48373   4043152981 3.933 2004     2700      1     2.705
## 48385   4344583554 4.318 2004     2700      1     3.090
## 48512   3342892905 4.641 2004     2700      1     3.413
## 48522   3242759883 3.736 2004     2700      1     2.508
## 48553   3142735110 4.217 2004     2700      1     3.194
## 48554   3142745348 4.790 2004     2700      1     3.462
## 48585   3042728480 4.043 2004     2700      1     3.020
## 48665   3042642128 5.510 2004     2700      1     4.282
## 48672   3042731163 4.134 2004     2700      1     2.906
## 48674   3042781155 3.884 2004     2700      1     2.656
## 48679   3042821849 3.670 2004     2700      1     2.647
## 48702   3242749074 4.079 2004     2700      1     2.851
## 48705   3242754344 3.746 2004     2700      1     2.623
## 48718   4043070821 3.779 2004     2700      1     2.551
## 48720   4043082182 3.928 2004     2700      1     2.700
## 48721   4043092238 4.451 2004     2700      1     3.223
## 48951   2942537772 4.149 2004     2700      1     3.026
## 48952   2942554887 5.618 2004     2700      1     4.495
## 48957   2942627194 4.077 2004     2700      1     2.849
## 49006   2542548063 5.360 2004     2700      1     4.337
## 49069   2942709937 3.782 2004     2700      1     2.554
## 49074   2942720933 4.617 2004     2700      1     3.289
## 49281   2442665224 4.280 2004     2700      1     3.052
## 49283   2442715038 5.135 2004     2700      1     3.907
## 49290   2442696436 4.547 2004     2700      1     3.319
## 49306   2342471392 6.269 2004     2700      1     5.041
## 49307   2442479695 5.758 2004     2700      1     4.530
## 49309   2442572117 5.004 2004     2700      1     3.776
## 49313   2442590642 4.673 2004     2700      1     3.445
## 49338   2342464257 4.234 2004     2700      1     3.006
## 49345   2342486731 4.168 2004     2700      1     2.940
## 49369   2342517421 5.147 2004     2700      1     4.124
## 49373   3843094224 4.764 2004     2700      1     3.536
## 49380   2342501856 4.398 2004     2700      1     3.170
## 49381   2342564429 4.008 2004     2700      1     2.780
## 49424   2342488880 3.755 2004     2700      1     2.527
## 49427   2342516208 3.700 2004     2700      1     2.677
## 49428   2342524107 3.852 2004     2700      1     2.624
## 49434   2342646956 4.407 2004     2700      1     3.179
## 49472   2942536636 4.358 2004     2700      1     3.130
## 49475   2942561024 4.094 2004     2700      1     2.866
## 49649   2142649221 4.212 2004     2700      1     3.189
```

```
## 49651   2142758687 4.063 2004     2700      1     2.835
## 49664   1942436013 4.103 2004     2700      1     2.875
## 49674  11144354823 4.207 2004     2700      1     2.979
## 49696  12144291219 3.697 2004     2700      1     2.574
## 49710   3042762336 4.314 2004     2700      1     3.086
## 49718  11144357519 3.986 2004     2700      1     2.658
## 49722   1842864234 5.018 2004     2700      1     3.790
## 49739   1842815777 4.502 2004     2700      1     3.174
## 49743  11144356419 3.866 2004     2700      1     2.638
## 49830   2442509789 3.581 2004     2700      1     2.558
## 49967  12144288049 4.590 2004     2700      1     3.362
## 49970   1642400686 4.214 2004     2700      1     2.986
## 49977   1642296706 3.739 2004     2700      1     2.511
## 50015  11144355828 4.350 2004     2700      1     3.122
## 50019  10744229257 3.730 2004     2700      1     2.502
## 50026   1642369930 4.229 2004     2700      1     3.001
## 50055   1442355581 4.481 2004     2700      1     3.253
## 50062  10744220250 4.551 2004     2700      1     3.223
## 50063  10744225301 5.224 2004     2700      1     3.996
## 50065   1442353066 3.979 2004     2700      1     2.751
## 50071  11144357189 3.839 2004     2700      1     2.511
## 50087   1442357943 3.847 2004     2700      1     2.619
## 50332  10744233940 4.677 2004     2700      1     3.449
## 50335   1342331884 3.850 2004     2700      1     2.622
## 50388   1442314663 3.997 2004     2700      1     2.669
## 50413  10744223871 4.045 2004     2700      1     2.817
## 50414   1342288777 4.142 2004     2700      1     3.119
## 50468   1642327683 4.090 2004     2700      1     2.967
## 50613   1642499234 4.500 2004     2700      1     3.272
## 50622   1642540483 4.321 2004     2700      1     2.993
## 50623   9144252520 4.184 2004     2700      1     2.956
## 50643   0345872128 4.849 2004     2700      1     3.621
## 50657   0347948542 3.871 2004     2700      1     2.643
## 50661   0345824715 4.381 2004     2700      1     3.153
## 50674   0346599193 4.602 2004     2700      1     3.374
## 50675   0347416893 3.842 2004     2700      1     2.614
## 50678   0348047527 4.767 2004     2700      1     3.439
## 50695   0345830478 3.919 2004     2700      1     2.591
## 50712   0346787782 3.666 2004     2700      1     2.543
## 50713   0346787909 3.820 2004     2700      1     2.592
## 50725   0347418277 4.315 2004     2700      1     3.292
## 50732   0742306162 3.544 2004     2700      1     2.521
## 50771  10944261840 3.779 2004     2700      1     2.656
## 50818   1342332130 3.828 2004     2700      1     2.600
## 50848  13744263906 4.057 2004     2700      1     3.034
## 50998   1542571327 3.817 2004     2700      1     2.589
## 51428   2142717404 4.031 2004     2700      1     2.803
## 51821   3242733807 3.931 2004     2700      1     2.703
## 51895   3442899889 4.734 2004     2700      1     3.506
## 51927   4043085036 3.742 2004     2700      1     2.514
```

```
## 52262  85047690933 4.353 2004     2700      1     3.125
## 52323   9644283066 4.178 2004     2700      1     2.950
## 52387   3142514201 3.624 2004     1300      2     2.601
## 52521  29544447206 3.996 2005     2700      1     2.697
## 52524  29544437848 4.305 2005     2700      1     3.006
## 52539  29144451858 4.279 2005     2700      1     3.184
## 52543  29144490030 3.866 2005     2700      1     2.567
## 52545  29144533834 4.497 2005     2700      1     3.297
## 52558  28944434329 4.369 2005     2700      1     3.169
## 52561  28944445445 4.451 2005     2700      1     3.356
## 52565  28944433025 3.552 2005     2700      1     2.557
## 52569  28944447646 4.919 2005     2700      1     3.620
## 52607  28944437578 4.673 2005     2700      1     3.678
## 52614  28844472594 3.996 2005     2700      1     2.697
## 52618  28844501802 4.241 2005     2700      1     3.041
## 52767  30144444279 4.733 2005     2700      1     3.533
## 52796  31044442642 3.817 2005     2700      1     2.617
## 52797  31044444463 4.156 2005     2700      1     2.956
## 52803  31044451759 4.265 2005     2700      1     3.065
## 52804  31044452189 4.001 2005     2700      1     2.702
## 53762  28144433147 4.679 2005     2700      1     3.380
## 53764  28144453057 4.091 2005     2700      1     2.891
## 53771  28144443357 3.807 2005     2700      1     2.712
## 53772  28144451163 4.332 2005     2700      1     3.132
## 53773  28144453509 3.901 2005     2700      1     2.806
## 53794  27744431927 4.342 2005     2700      1     3.142
## 53795  27744494434 4.618 2005     2700      1     3.623
## 53807  27744496592 4.514 2005     2700      1     3.314
## 53820  27744606737 4.954 2005     2700      1     3.754
## 53828  27744477622 3.728 2005     2700      1     2.528
## 53845  27644443333 4.575 2005     2700      1     3.580
## 53863  27644452293 3.749 2005     2700      1     2.549
## 53865  27644461904 3.702 2005     2700      1     2.707
## 53875  27644513772 4.335 2005     2700      1     3.240
## 53883  27644568878 4.218 2005     2700      1     3.223
## 53918  28444478439 4.387 2005     2700      1     3.088
## 54155  27244440305 4.455 2005     2700      1     3.360
## 54170  26844536978 5.830 2005     2700      1     4.630
## 54171  26844552488 5.412 2005     2700      1     4.212
## 54178  26844494491 4.633 2005     2700      1     3.538
## 54179  26844538114 4.966 2005     2700      1     3.766
## 54180  26844544418 3.601 2005     2700      1     2.506
## 54208  26444452073 5.212 2005     2700      1     3.913
## 54210  26444477603 4.006 2005     2700      1     3.011
## 54211  26444506232 4.042 2005     2700      1     2.842
## 54217  26444574824 4.419 2005     2700      1     3.120
## 54243  25844488342 3.971 2005     2700      1     2.771
## 54246  25844509466 4.220 2005     2700      1     2.921
## 54285  25844438380 4.442 2005     2700      1     3.242
## 54287  25844440999 3.660 2005     2700      1     2.665
```

```
## 54305  26444515022 3.926 2005     2700      1     2.627
## 54312  26444551183 3.878 2005     2700      1     2.678
## 54314  26444577543 3.881 2005     2700      1     2.681
## 54319  26444598506 4.478 2005     2700      1     3.278
## 54557  25444461935 3.898 2005     2700      1     2.698
## 54562  25144470719 3.920 2005     2700      1     2.720
## 54575  25144456112 5.883 2005     2700      1     4.683
## 54580  25144502974 3.881 2005     2700      1     2.681
## 54581  25144505285 3.742 2005     2700      1     2.542
## 54582  25144518364 4.074 2005     2700      1     2.874
## 54614  24944531368 3.878 2005     2700      1     2.678
## 54632  24644439190 4.674 2005     2700      1     3.474
## 54633  24644443217 4.833 2005     2700      1     3.633
## 54677  24344483122 3.953 2005     2700      1     2.753
## 54701  24044484768 3.823 2005     2700      1     2.728
## 54742  24744470476 4.035 2005     2700      1     2.835
## 55090  23944436856 3.575 2005     2700      1     2.580
## 55101  23944436934 4.091 2005     2700      1     2.891
## 55113  23844539714 4.976 2005     2700      1     3.776
## 55121  23844454293 3.800 2005     2700      1     2.600
## 55123  23844494864 4.211 2005     2700      1     2.912
## 55133  23844533906 4.026 2005     2700      1     2.826
## 55149  23744500093 4.222 2005     2700      1     2.923
## 55162  23444448145 3.866 2005     2700      1     2.771
## 55163  23444452077 4.300 2005     2700      1     3.100
## 55164  23744459272 3.820 2005     2700      1     2.620
## 55189  23044480500 3.937 2005     2700      1     2.737
## 55193  23044442114 4.460 2005     2700      1     3.260
## 55235  23644435612 3.976 2005     2700      1     2.677
## 55238  23644439061 4.516 2005     2700      1     3.316
## 55240  23644446866 3.989 2005     2700      1     2.789
## 55244  23644456321 3.901 2005     2700      1     2.701
## 55245  23644460325 3.942 2005     2700      1     2.947
## 55428  22844445230 4.774 2005     2700      1     3.475
## 55430  22844447051 4.585 2005     2700      1     3.385
## 55438  22844434877 4.780 2005     2700      1     3.580
## 55442  22844448780 4.222 2005     2700      1     3.022
## 55448  22144468433 3.996 2005     2700      1     2.796
## 55466  22344443114 4.392 2005     2700      1     3.192
## 55468  22344450570 4.102 2005     2700      1     2.902
## 55471  22344457679 4.119 2005     2700      1     2.919
## 55485  22244446183 4.420 2005     2700      1     3.425
## 55486  22244485721 4.183 2005     2700      1     3.088
## 55521  22144499588 3.713 2005     2700      1     2.618
## 55536  21444447949 4.377 2005     2700      1     3.177
## 55569  21444441271 4.299 2005     2700      1     3.000
## 55570  21444447966 4.369 2005     2700      1     3.169
## 55571  21444450527 4.501 2005     2700      1     3.406
## 55591  22144440464 4.362 2005     2700      1     3.162
## 55605  22144492953 4.207 2005     2700      1     2.908
```

```
## 55606  22144495740 3.790 2005     2700      1     2.590
## 55869  20644465409 3.857 2005     2700      1     2.657
## 55886  20444444376 4.187 2005     2700      1     2.987
## 55888  20444501831 4.934 2005     2700      1     3.734
## 55913  20444398121 3.931 2005     2700      1     2.632
## 55914  20544462061 4.452 2005     2700      1     3.357
## 55957  21144448879 4.965 2005     2700      1     3.765
## 55970  18844374582 3.884 2005     2700      1     2.889
## 55986  19644379430 3.989 2005     2700      1     2.789
## 55989  19644391940 4.484 2005     2700      1     3.185
## 56001  20444454585 3.817 2005     2700      1     2.617
## 56002  20444457518 4.596 2005     2700      1     3.601
## 56009  20444499355 4.459 2005     2700      1     3.259
## 56310  19444363154 3.931 2005     2700      1     2.731
## 56315  21144457072 4.436 2005     2700      1     3.236
## 56318  18944367562 4.915 2005     2700      1     3.616
## 56337  18644370841 4.807 2005     2700      1     3.812
## 56339  18744382834 3.953 2005     2700      1     2.958
## 56341  20844438843 4.222 2005     2700      1     3.022
## 56349  18244371651 5.531 2005     2700      1     4.331
## 56350  20844432890 4.096 2005     2700      1     3.001
## 56396  18744401357 3.950 2005     2700      1     2.651
## 56397  18744406319 3.706 2005     2700      1     2.506
## 56407  17944362664 4.751 2005     2700      1     3.452
## 56409  17944380530 3.760 2005     2700      1     2.560
## 56410  20944448998 3.953 2005     2700      1     2.753
## 56491  20944436157 4.708 2005     2700      1     3.508
## 56656  17644392071 4.823 2005     2700      1     3.524
## 56664  17544364774 4.269 2005     2700      1     3.069
## 56684  17844402965 4.271 2005     2700      1     3.071
## 56693  17244379193 4.136 2005     2700      1     2.936
## 56702  17144385790 5.319 2005     2700      1     4.020
## 56705  17144430418 4.592 2005     2700      1     3.392
## 56707  21544449115 4.018 2005     2700      2     2.719
## 56712  17144398838 3.739 2005     2700      1     2.539
## 56719  20244366093 3.760 2005     2700      1     2.560
## 56730  20144386747 4.154 2005     2700      1     2.954
## 56731  20144388867 4.033 2005     2700      1     2.833
## 56738  16444381731 3.829 2005     2700      1     2.734
## 56740  20144387681 4.355 2005     2700      1     3.155
## 56748  15744372810 5.001 2005     2700      1     3.801
## 56757  20144386837 4.994 2005     2700      1     3.999
## 56762  15944387060 3.940 2005     2700      1     2.641
## 56763  15944406488 4.216 2005     2700      1     3.016
## 56824  20144387499 3.895 2005     2700      1     2.596
## 56965  15744378579 5.147 2005     2700      1     4.152
## 56967  20144366696 4.939 2005     2700      1     3.739
## 56992  15044344420 3.717 2005     2700      1     2.517
## 57041  14744275842 4.576 2005     2700      1     3.376
## 57048  14544304095 5.144 2005     2700      1     3.944
```

```
## 57060  20044369898 3.543 2005     2700      1     2.548
## 57105  14644435711 3.499 2005     2700      1     2.504
## 57189  20044368802 4.056 2005     2700      1     2.856
## 57354  13844315270 3.929 2005     2700      1     2.630
## 57359  13944250706 3.610 2005     2700      1     2.615
## 57391  13744253924 3.839 2005     2700      1     2.744
## 57392  13744261333 3.974 2005     2700      1     2.774
## 57400  13844265752 4.225 2005     2700      1     3.025
## 57402  13844274980 3.923 2005     2700      1     2.723
## 57417  13444256401 4.619 2005     2700      1     3.419
## 57420  13444293219 5.349 2005     2700      1     4.149
## 57426  13444309949 4.491 2005     2700      1     3.291
## 57427  13444310598 3.728 2005     2700      1     2.528
## 57437  13444263619 4.684 2005     2700      1     3.484
## 57495  14944348752 3.863 2005     2700      1     2.663
## 57674  19944432606 4.121 2005     2700      1     3.126
## 57678  12544253745 4.628 2005     2700      1     3.428
## 57683  12544253835 4.548 2005     2700      1     3.348
## 57684  12544257550 4.190 2005     2700      1     2.990
## 57685  19944431657 4.004 2005     2700      1     2.804
## 57698  12144284532 3.878 2005     2700      1     2.883
## 57714  19944426812 4.047 2005     2700      1     2.748
## 57732  11844268070 4.117 2005     2700      1     2.917
## 57734  13744249546 3.742 2005     2700      1     2.647
## 57739  11844269164 4.196 2005     2700      1     3.101
## 57784  19944429519 4.900 2005     2700      1     3.700
## 57787  11244285095 3.904 2005     2700      1     2.605
## 57788  11244331742 4.906 2005     2700      1     3.706
## 57831  12744263350 4.085 2005     2700      1     3.090
## 57836  12744278427 3.829 2005     2700      1     2.530
## 57868  13444260835 3.732 2005     2700      1     2.637
## 57870  13444268127 4.350 2005     2700      1     3.355
## 57881  13744254861 3.749 2005     2700      1     2.549
## 57883  13744256812 4.438 2005     2700      1     3.139
## 57919  14144253829 3.857 2005     2700      1     2.657
## 57965  14644425916 3.631 2005     2700      1     2.536
## 58080  16244415880 3.909 2005     2700      1     2.914
## 58139  17644377222 4.115 2005     2700      1     3.120
## 58156  17844391245 3.929 2005     2700      1     2.729
## 58196  18244390229 3.826 2005     2700      1     2.831
## 58197  18244393432 3.533 2005     2700      1     2.538
## 58304  20144388724 3.857 2005     2700      1     2.657
## 58349  20944446791 4.391 2005     2700      1     3.191
## 58709  26944497954 3.884 2005     2700      1     2.684
## 59586  33845866114 3.911 2006     2700      1     2.847
## 59588  33845880922 5.714 2006     2700      1     4.749
## 59600  33845866055 4.159 2006     2700      1     2.990
## 59601  33845890326 4.968 2006     2700      1     3.799
## 59635  33845693164 3.800 2006     2700      1     2.631
## 59636  33845707784 4.606 2006     2700      1     3.437
```

```
## 59643  33845717310 4.910 2006     2700      1     3.846
## 59678  33845483482 3.948 2006     2700      1     2.983
## 59679  33845490014 6.122 2006     2700      1     4.953
## 59680  33845497716 3.958 2006     2700      1     2.689
## 59681  33845500545 4.941 2006     2700      1     3.772
## 59686  33845478495 4.306 2006     2700      1     3.037
## 59687  33845485717 3.595 2006     2700      1     2.531
## 59688  33845491101 3.880 2006     2700      1     2.611
## 59689  33947325340 4.051 2006     2700      1     2.987
## 59705  33845449051 4.316 2006     2700      1     3.147
## 59714  33845450660 4.586 2006     2700      1     3.317
## 59746  33748560432 4.653 2006     2700      1     3.484
## 59787  33845297656 3.725 2006     2700      1     2.556
## 59788  33845299821 3.784 2006     2700      1     2.615
## 59792  33845317604 4.246 2006     2700      1     3.077
## 59793  33845322277 3.595 2006     2700      1     2.630
## 59798  33845354407 3.759 2006     2700      1     2.590
## 59849  33845926387 3.808 2006     2700      1     2.639
## 59883  33846317116 3.734 2006     2700      1     2.769
## 60010  33847190697 3.751 2006     2700      1     2.582
## 61367  33751529506 4.659 2006     2700      1     3.390
## 61370  33751545838 5.100 2006     2700      1     3.831
## 61389  33751185526 3.780 2006     2700      1     2.716
## 61412  33751206860 4.493 2006     2700      1     3.324
## 61413  33751217149 4.658 2006     2700      1     3.489
## 61441  33750983605 5.495 2006     2700      1     4.326
## 61448  33750978404 4.084 2006     2700      1     2.915
## 61491  33750694575 4.082 2006     2700      1     2.913
## 61499  33750726707 4.632 2006     2700      1     3.568
## 61500  33750731595 4.693 2006     2700      1     3.524
## 61521  33750630655 3.751 2006     2700      1     2.786
## 61532  33750500326 4.054 2006     2700      1     2.885
## 61535  33750507268 4.113 2006     2700      1     2.844
## 61564  33750507246 3.788 2006     2700      1     2.619
## 61566  33750518186 4.246 2006     2700      1     3.281
## 61570  33750532312 5.148 2006     2700      1     4.183
## 61574  33750584214 5.548 2006     2700      1     4.279
## 61582  33750597734 4.320 2006     2700      1     3.151
## 62170  33750346251 4.432 2006     2700      1     3.368
## 62188  33750300528 4.116 2006     2700      1     2.947
## 62212  33750035507 4.320 2006     2700      1     3.355
## 62216  33750081777 3.983 2006     2700      1     2.714
## 62223  33750106228 4.641 2006     2700      1     3.472
## 62224  33750121615 4.499 2006     2700      1     3.330
## 62227  33749636443 3.738 2006     2700      1     2.773
## 62260  33749597936 4.523 2006     2700      1     3.254
## 62261  33749599695 5.460 2006     2700      1     4.495
## 62263  33749618085 4.879 2006     2700      1     3.710
## 62273  33749613119 5.139 2006     2700      1     4.174
## 62274  33749659678 3.989 2006     2700      1     2.820
```

```
## 62294  33749436870 4.405 2006     2700      1     3.236
## 62297  33749445317 5.971 2006     2700      1     4.802
## 62319  33749020857 4.306 2006     2700      1     3.137
## 62325  33749430001 3.938 2006     2700      1     2.973
## 62331  33749441325 4.657 2006     2700      1     3.488
## 62335  33749446861 3.772 2006     2700      1     2.603
## 62472  33846664981 3.928 2006     2700      1     2.759
## 62539  39049174296 3.751 2006     2700      1     2.582
## 62724  43749107283 4.870 2006     2700      1     3.701
## 62921  33749040947 3.626 2006     2700      1     2.562
## 62924  33749079234 3.958 2006     2700      1     2.789
## 63010  33748354709 3.584 2006     2700      1     2.619
## 63021  33748316794 3.928 2006     2700      1     2.963
## 63047  33748437811 4.028 2006     2700      1     2.859
## 63048  33748438134 4.637 2006     2700      1     3.468
## 63124  33748292955 4.111 2006     2700      1     2.942
## 63132  33748312093 4.065 2006     2700      1     2.896
## 63137  33748331335 3.890 2006     2700      1     2.621
## 63152  33748747557 3.876 2006     2700      1     2.707
## 63833  33748161812 4.051 2006     2700      1     2.882
## 63860  33747834620 3.716 2006     2700      1     2.547
## 63864  33747870163 5.283 2006     2700      1     4.114
## 63866  33747584933 4.308 2006     2700      1     3.139
## 63887  33747135682 4.969 2006     2700      1     3.800
## 63904  33747343208 5.434 2006     2700      1     4.265
## 63910  33747102040 3.816 2006     2700      1     2.647
## 63948  33746875641 4.520 2006     2700      1     3.351
## 63985  33746768834 3.832 2006     2700      1     2.563
## 63987  33746587393 4.057 2006     2700      1     2.993
## 64006  33746437130 4.877 2006     2700      1     3.608
## 64017  33746667851 4.079 2006     2700      1     2.910
## 64020  33746672599 3.964 2006     2700      1     2.695
## 64021  33746675685 4.031 2006     2700      1     2.862
## 64033  33746724157 3.887 2006     2700      1     2.618
## 64083  33748046126 3.747 2006     2700      1     2.578
## 64148  33750630109 3.948 2006     2700      1     2.679
## 64376  33746387027 3.828 2006     2700      1     2.659
## 64377  33746430561 3.742 2006     2700      1     2.573
## 64378  33746430578 4.001 2006     2700      1     2.832
## 64379  33746452905 4.192 2006     2700      1     3.128
## 64395  33748513704 3.579 2006     2700      1     2.614
## 64421  33746075560 4.642 2006     2700      1     3.473
## 64422  33746088001 5.656 2006     2700      1     4.691
## 64482  33745890309 4.593 2006     2700      1     3.324
## 64484  33745913326 3.707 2006     2700      1     2.643
## 64512  33745698681 4.248 2006     2700      1     3.079
## 64514  33745698685 4.688 2006     2700      1     3.624
## 64520  33745611449 5.100 2006     2700      1     4.135
## 64521  33745612500 3.983 2006     2700      1     2.714
## 64522  33745614361 5.082 2006     2700      1     4.117
```

```
## 64523  33745632422 4.146 2006     2700      1     2.877
## 64530  33745812011 4.129 2006     2700      1     2.960
## 64531  33745817034 3.992 2006     2700      1     3.027
## 64535  33745840137 3.992 2006     2700      1     2.823
## 64537  33745851479 4.385 2006     2700      1     3.216
## 64548  33744914821 3.540 2006     2700      1     2.575
## 64979  33745661633 3.776 2006     2700      1     2.712
## 65008  33745227399 4.917 2006     2700      1     3.748
## 65009  33745227751 3.734 2006     2700      1     2.565
## 65010  33745255382 4.457 2006     2700      1     3.393
## 65030  33646360865 4.352 2006     2700      1     3.183
## 65051  33745081608 5.495 2006     2700      1     4.530
## 65054  33745102555 5.215 2006     2700      1     4.046
## 65066  33745065872 3.734 2006     2700      1     2.769
## 65093  33744718154 3.600 2006     2700      1     2.635
## 65106  33744966030 4.777 2006     2700      1     3.608
## 65135  33646825385 4.208 2006     2700      1     3.039
## 65148  33744472167 3.702 2006     2700      1     2.533
## 65149  33744472168 4.090 2006     2700      1     2.921
## 65151  33744498063 4.280 2006     2700      1     3.111
## 65190  33745210374 3.862 2006     2700      1     2.693
## 65191  33745212925 3.876 2006     2700      1     2.707
## 65192  33745217611 3.764 2006     2700      1     2.595
## 65230  33745685701 3.747 2006     2700      1     2.578
## 65338  33750089034 3.980 2006     2700      1     3.015
## 65365  33846406448 3.854 2006     2700      1     2.685
## 65669  33646859471 4.708 2006     2700      1     3.743
## 65672  33646875293 4.246 2006     2700      1     2.977
## 65680  33646828953 3.862 2006     2700      1     2.897
## 65716  33646676272 4.484 2006     2700      1     3.315
## 65717  33646678189 4.410 2006     2700      1     3.241
## 65764  33646482407 3.887 2006     2700      1     2.618
## 65765  33646488054 4.825 2006     2700      1     3.556
## 65772  33646450281 3.851 2006     2700      1     2.582
## 65773  33646452922 4.459 2006     2700      1     3.290
## 65790  33646475441 3.584 2006     2700      1     2.619
## 65816  30044438368 5.433 2006     2700      1     4.369
## 65837  33646177653 3.505 2006     2700      1     2.540
## 65838  33646178951 5.085 2006     2700      1     3.916
## 65847  33646392755 4.004 2006     2700      1     2.835
## 65850  33646401549 3.977 2006     2700      1     2.808
## 65975  33745489614 3.646 2006     2700      1     2.582
## 66274  33646164892 4.031 2006     2700      1     2.862
## 66300  33645865234 4.803 2006     2700      1     3.634
## 66319  33646052556 5.164 2006     2700      1     3.995
## 66338  33645770270 4.219 2006     2700      1     3.155
## 66344  33645729203 4.579 2006     2700      1     3.410
## 66346  33645741222 4.267 2006     2700      1     2.998
## 66372  33645523067 6.444 2006     2700      1     5.380
## 66392  29944434110 3.964 2006     2700      1     2.795
```

```
## 66393  29944434742 4.119 2006     2700      1     2.850
## 66394  29944442726 4.794 2006     2700      1     3.625
## 66408  33645222620 3.812 2006     2700      1     2.543
## 66409  33645236930 3.684 2006     2700      1     2.515
## 66423  33645530745 4.561 2006     2700      1     3.392
## 66490  33646420626 3.788 2006     2700      1     2.619
## 66492  33646432226 4.312 2006     2700      1     3.248
## 66777  33645390604 3.621 2006     2700      1     2.656
## 66778  33645399959 4.643 2006     2700      1     3.474
## 66806  33645102811 3.897 2006     2700      1     2.833
## 66807  33645103550 5.432 2006     2700      1     4.467
## 66813  33645068471 4.040 2006     2700      1     2.771
## 66834  33745029895 3.832 2006     2700      1     2.768
## 66845  33644979597 4.290 2006     2700      1     3.325
## 66847  33644984085 4.144 2006     2700      1     2.975
## 66856  33644967633 3.995 2006     2700      1     3.030
## 66872  33644957630 4.031 2006     2700      1     2.862
## 66880  33644833147 4.678 2006     2700      1     3.509
## 66883  33644807437 5.185 2006     2700      1     3.916
## 66925  33644584065 4.544 2006     2700      1     3.275
## 66926  33644605043 5.148 2006     2700      1     3.979
## 66930  33644638676 3.816 2006     2700      1     2.647
## 66932  33644642974 3.921 2006     2700      1     2.752
## 66942  33644654777 4.106 2006     2700      1     2.937
## 66944  33644660402 3.689 2006     2700      1     2.625
## 67001  33645355399 4.176 2006     2700      1     3.007
## 67354  33344467770 4.351 2006     2700      1     3.182
## 67363  33644536465 4.565 2006     2700      1     3.396
## 67394  32644457434 4.090 2006     2700      1     2.921
## 67395  32644467389 5.178 2006     2700      1     3.909
## 67432  32144436098 4.713 2006     2700      1     3.444
## 67435  32144455695 4.106 2006     2700      1     2.937
## 67438  32544459770 3.748 2006     1000      2     2.579
## 67445  32144432645 4.141 2006     2700      1     3.077
## 67446  32144441493 4.605 2006     2700      1     3.640
## 67447  32144443648 4.776 2006     2700      1     3.712
## 67455  32544444748 3.983 2006     2700      1     2.714
## 67461  32044452414 4.031 2006     2700      1     2.862
## 67474  31444436871 3.689 2006     2700      1     2.520
## 67506  32444436464 3.716 2006     2700      1     2.547
## 67508  32444438623 3.828 2006     2700      1     2.659
## 67510  32444443951 4.627 2006     2700      1     3.358
## 67517  32444450909 3.534 2006     2700      1     2.569
## 67538  33644842676 4.351 2006     2700      1     3.182
## 67781  31544462628 3.679 2006     1000      2     2.510
## 67786  31344467254 4.933 2006     2700      1     3.968
## 67787  31344470705 5.219 2006     2700      1     3.950
## 67790  31344479326 4.396 2006     2700      1     3.227
## 67795  31344454603 3.772 2006     2700      1     2.603
## 67804  30844432417 3.675 2006     2700      1     2.506
```

```
## 67813  30944452384 4.432 2006     2700      1     3.263
## 67815  30944457531 4.478 2006     2700      1     3.414
## 67876  32244433703 3.636 2006     2700      1     2.671
## 67920  30944433554 4.252 2006     2700      1     3.083
## 67933  31044431832 4.425 2006     2700      1     3.256
## 67938  31044436630 4.688 2006     2700      1     3.519
## 67941  31044442783 3.734 2006     2700      1     2.565
## 67942  31044446194 4.154 2006     2700      1     3.189
## 67944  31044453008 4.376 2006     2700      1     3.207
## 67945  31044453148 4.025 2006     2700      1     2.856
## 67950  31044456529 5.039 2006     2700      1     3.870
## 67972  32044434020 3.847 2006     2700      1     2.678
## 67974  32044435429 4.349 2006     2700      1     3.080
## 67991  32144461605 3.995 2006     2700      1     2.726
## 68363  33748146129 3.511 2006     2700      1     2.546
## 68847  37549036689 4.103 2007     2700      1     3.151
## 68848  37549039095 3.740 2007     2700      1     2.888
## 68868  37349092105 3.950 2007     2700      1     2.893
## 68877  37349011273 3.858 2007     2700      1     2.701
## 68878  37349028816 4.130 2007     2700      1     3.073
## 68923  37249041477 3.789 2007     2700      1     2.732
## 68924  37249061691 4.883 2007     2700      1     3.826
## 68931  37149028276 4.444 2007     2700      1     3.592
## 68952  38449085341 3.654 2007     2700      1     2.597
## 68959  36549078679 4.337 2007     2700      1     3.280
## 68962  37249065591 3.858 2007     2700      1     2.906
## 69125  36048934283 3.807 2007     2700      1     2.955
## 69128  36049048260 3.795 2007     2700      1     2.638
## 69155  36849011221 4.476 2007     2700      1     3.419
## 69168  36849035575 3.973 2007     2700      1     2.916
## 69181  36849068108 4.107 2007     2700      1     2.950
## 69184  36849079348 3.564 2007     2700      1     2.612
## 69237  38049023539 4.325 2007     2700      1     3.373
## 69238  38049038389 3.992 2007     2700      1     2.935
## 69239  38049053061 4.415 2007     2700      1     3.358
## 69241  38049087576 4.711 2007     2700      1     3.654
## 69968  56149101731 3.783 2007     2700      1     2.626
## 70668  36549006511 3.668 2007     2700      1     2.611
## 70677  36549033449 4.134 2007     2700      1     2.977
## 70678  36549066931 4.045 2007     2700      1     3.193
## 70725  36348961929 4.189 2007     2700      1     3.132
## 70727  36348982774 4.230 2007     2700      1     3.173
## 70728  36349035495 4.062 2007     2700      1     2.905
## 70766  36148981435 4.360 2007     2700      1     3.303
## 70798  36148951701 3.556 2007     2700      1     2.604
## 70799  36148966515 4.227 2007     2700      1     3.070
## 70800  36149000149 5.380 2007     2700      1     4.323
## 70832  35848963822 4.711 2007     2700      1     3.654
## 70833  35848964343 4.234 2007     2700      1     3.282
## 70835  35848969753 4.853 2007     2700      1     3.796
```

```
## 70842  35848935201 5.343 2007     2700      1     4.186
## 70844  35848964571 4.402 2007     2700      1     3.450
## 70845  35848968871 6.256 2007     2700      1     5.199
## 70888  35248896228 3.992 2007     2700      1     2.835
## 70910  35748974550 4.877 2007     2700      1     3.820
## 70923  36048931015 4.213 2007     2700      1     3.156
## 70927  36048938340 3.610 2007     2700      1     2.553
## 70929  36048953291 3.747 2007     2700      1     2.590
## 70932  36048958880 3.498 2007     2700      1     2.646
## 70934  36048963838 4.213 2007     2700      1     3.156
## 70935  36048968753 4.402 2007     2700      1     3.345
## 70950  36049037185 3.835 2007     2700      1     2.778
## 70951  36049042525 3.835 2007     2700      1     2.778
## 70959  36248973101 3.783 2007     2700      1     2.626
## 71265  41149161956 4.099 2007     2700      1     3.147
## 71308  48849113878 4.849 2007     2700      1     3.792
## 71611  35348968910 3.761 2007     1000      2     2.704
## 71623  35348981185 3.708 2007     2700      1     2.856
## 71624  35348997893 4.319 2007     2700      1     3.467
## 71653  35148831175 4.279 2007     2700      1     3.427
## 71667  34948830633 3.376 2007     2700      1     2.524
## 71669  35248829599 4.645 2007     2700      1     3.693
## 71673  35248895346 4.762 2007     2700      1     3.705
## 71681  35348866794 3.930 2007     2700      1     2.773
## 71714  34948849235 3.498 2007     2700      1     2.546
## 71717  34948891641 4.829 2007     2700      1     3.772
## 71720  34948845507 4.340 2007     2700      1     3.283
## 71722  34948882262 4.028 2007     2700      1     2.871
## 71753  34548827323 4.070 2007     2700      1     3.013
## 71755  34548835744 4.466 2007     2700      1     3.409
## 71788  34948860191 3.564 2007     2700      1     2.507
## 71791  34948872683 3.695 2007     2700      1     2.638
## 71809  35348828533 4.122 2007     2700      1     2.965
## 71831  35348985396 3.688 2007     2700      1     2.531
## 71842  35448937632 3.617 2007     2700      1     2.765
## 71847  35648946842 3.807 2007     2700      1     2.955
## 71864  35649003880 4.876 2007     2700      1     3.719
## 72101  40649100791 3.868 2007     2700      1     2.811
## 72327  34648832073 4.784 2007     2700      1     3.727
## 72334  34748875855 4.118 2007     2700      1     3.061
## 72398  34548833077 4.091 2007     2700      1     3.034
## 72400  34548849816 4.178 2007     2700      1     3.121
## 72454  34548776934 3.753 2007     2700      1     2.696
## 72465  34548570632 4.484 2007     2700      1     3.427
## 72494  34547426346 3.681 2007     2700      1     2.624
## 72517  34548419640 3.595 2007     2700      1     2.743
## 72546  34548418912 4.594 2007     2700      1     3.537
## 72548  34548433987 4.078 2007     2700      1     3.021
## 72555  34548426480 4.019 2007     2700      1     2.862
## 72556  34548430952 4.304 2007     2700      1     3.147
```

```
## 72568  34848816502 4.152 2007     2700      1     3.200
## 72569  34848818486 3.879 2007     2700      1     2.822
## 72572  34848831462 4.045 2007     2700      1     2.988
## 72575  34848838902 4.019 2007     2700      1     2.962
## 72578  34848858523 4.244 2007     2700      1     3.187
## 72581  34848881905 4.352 2007     2700      1     3.295
## 72582  34848881907 3.681 2007     2700      1     2.524
## 72583  34848888949 3.954 2007     2700      1     2.897
## 72587  34848901974 4.316 2007     2700      1     3.259
## 72588  34848902937 4.337 2007     2700      1     3.280
## 72591  34848920368 3.801 2007     2700      1     2.644
## 72606  34547502731 3.531 2007     2700      1     2.579
## 72608  34547541570 3.406 2007     2700      1     2.554
## 72609  34547545269 4.719 2007     2700      1     3.767
## 72613  34548040666 3.695 2007     2700      1     2.638
## 72619  34548172901 3.625 2007     2700      1     2.568
## 72623  34548262715 4.493 2007     2700      1     3.336
## 72673  34648836403 4.569 2007     2700      1     3.617
## 72681  34748862577 3.830 2007     2700      1     2.673
## 72807  36048990829 3.632 2007     2700      1     2.575
## 72822  36248973631 3.841 2007     2700      1     2.784
## 73355  34548221486 3.438 2007     1704      3     2.586
## 73492  34548066571 3.759 2007     2700      1     2.702
## 73493  34548075820 5.741 2007     2700      1     4.684
## 73494  34548083742 4.058 2007     2700      1     3.001
## 73500  34548092397 4.740 2007     2700      1     3.683
## 73551  34547950500 4.540 2007     2700      1     3.588
## 73558  34547876797 4.649 2007     2700      1     3.592
## 73559  34547906405 4.663 2007     2700      1     3.606
## 73563  34548361951 3.587 2007     2700      1     2.530
## 73594  34547757915 5.372 2007     2700      1     4.215
## 73599  34547730740 4.006 2007     2700      1     2.949
## 73600  34547755634 3.747 2007     2700      1     2.895
## 73649  34547647628 4.354 2007     2700      1     3.297
## 73681  34547614056 4.149 2007     2700      1     3.092
## 73683  34547630852 3.983 2007     2700      1     3.031
## 73684  34547634884 3.728 2007     2700      1     2.671
## 73685  34547643177 3.734 2007     2700      1     2.882
## 73686  34547652884 4.074 2007     2700      1     2.917
## 73697  34547681974 3.954 2007     2700      1     2.897
## 73699  34547686398 4.247 2007     2700      1     3.090
## 73747  34548397297 3.789 2007     2700      1     2.732
## 73760  34548456959 3.595 2007     2700      1     2.538
## 74193  34547148985 4.266 2007     2700      1     3.209
## 74194  34547170053 6.085 2007     2700      1     5.028
## 74199  34547230736 4.019 2007     2700      1     2.962
## 74203  34547159840 4.045 2007     2700      1     2.988
## 74204  34547167603 3.695 2007     2700      1     2.638
## 74226  34447521458 4.639 2007     2700      1     3.482
## 74241  34447544253 3.830 2007     2700      1     2.773
```

```
## 74246  34447300446 4.669 2007     2700      1     3.612
## 74247  34447512756 4.710 2007     2700      1     3.858
## 74249  34447520136 5.342 2007     2700      1     4.285
## 74250  34447558140 5.428 2007     2700      1     4.371
## 74312  34447340945 3.992 2007     2700      1     2.935
## 74314  34548165110 3.506 2007     2700      1     2.554
## 74322  34447257725 3.920 2007     2700      1     2.763
## 74323  34447291144 4.282 2007     2700      1     3.225
## 74369  34447116875 3.695 2007     2700      1     2.638
## 74375  34447121852 3.852 2007     2700      1     2.795
## 74385  34447137331 4.461 2007     2700      1     3.304
## 74398  34250840592 3.920 2007     2700      1     2.863
## 74704  35248881664 4.126 2007     2700      1     3.069
## 75040  34548567516 3.824 2007     2700      1     2.972
## 75044  34347210370 4.041 2007     2700      1     2.984
## 75045  34347237664 4.019 2007     2700      1     2.962
## 75047  34347255039 4.665 2007     2700      1     3.508
## 75057  34347272329 3.753 2007     2700      1     2.901
## 75085  34250696907 4.873 2007     2700      1     3.921
## 75086  34250722018 4.152 2007     2700      1     3.200
## 75089  34250779802 4.066 2007     2700      1     3.009
## 75092  34250835050 3.709 2007     1000      2     2.652
## 75142  34250212715 6.307 2007     2700      1     5.455
## 75144  34250309164 4.292 2007     2700      1     3.235
## 75168  34249899377 4.920 2007     2700      1     3.763
## 75183  34250024383 3.783 2007     2700      1     2.831
## 75202  34249982970 4.196 2007     2700      1     3.139
## 75203  34250027957 3.983 2007     2700      1     2.826
## 75236  34248149838 3.765 2007     2700      1     2.708
## 75239  34248169298 3.795 2007     2700      1     2.738
## 75242  34248201151 4.476 2007     2700      1     3.419
## 75243  34248206205 3.688 2007     2700      1     2.631
## 75256  34249037494 3.356 2007     2700      1     2.504
## 75260  34249662628 3.925 2007     2700      1     2.868
## 75262  34249863368 3.852 2007     2700      1     3.000
## 75274  34249913494 4.647 2007     2700      1     3.490
## 75275  34249915334 4.224 2007     2700      1     3.167
## 75382  34447648357 3.807 2007     2700      1     2.855
## 75793  34547309949 3.854 2007     2700      2     2.797
## 75956  34249302025 3.580 2007     2700      1     2.523
## 75959  34249803312 4.122 2007     2700      1     3.270
## 75961  34249830160 4.754 2007     2700      1     3.902
## 75973  34249027978 3.959 2007     2700      1     2.902
## 75990  34249657825 4.383 2007     2700      1     3.326
## 75993  34249673868 4.807 2007     2700      1     3.750
## 76036  34249047454 4.895 2007     2700      1     3.838
## 76042  34248551882 4.301 2007     2700      1     3.144
## 76043  34248559834 4.710 2007     2700      1     3.653
## 76044  34248564820 3.531 2007     2700      1     2.679
## 76093  34248399089 4.292 2007     2700      1     3.235
```

```
## 76121  34247615981 3.835 2007     2700      1     2.778
## 76133  34247862190 5.281 2007     2700      1     4.224
## 76164  34247268567 3.610 2007     2700      1     2.553
## 76199  34248184445 3.702 2007     2700      1     2.645
## 76200  34248185926 4.203 2007     2700      1     3.046
## 76214  34248214561 3.765 2007     2700      1     2.913
## 76218  34248227278 3.734 2007     2700      1     2.577
## 76835  34247863975 5.154 2007     2700      1     3.997
## 76849  34247476744 5.028 2007     2700      1     3.971
## 76864  34247470278 4.213 2007     2700      1     3.056
## 76865  34247498668 4.800 2007     2700      1     3.643
## 76883  34247167109 3.675 2007     2700      1     2.518
## 76903  34247241630 5.117 2007     2700      1     4.060
## 76945  34247118878 3.813 2007     2700      1     2.961
## 76946  34247144499 4.185 2007     2700      1     3.128
## 76947  34247149303 4.041 2007     2700      1     2.984
## 76950  34247523621 4.713 2007     2700      1     3.656
## 76959  34047096257 4.791 2007     2700      1     3.939
## 76970  34047243811 4.396 2007     2700      1     3.544
## 76972  34047268898 3.900 2007     2700      1     2.843
## 76981  34047237367 5.526 2007     2700      1     4.674
## 76988  34147096980 4.461 2007     2700      1     3.404
## 76993  34147124264 4.538 2007     2700      1     3.381
## 76995  34147139988 3.610 2007     2700      1     2.553
## 76998  34147153376 3.564 2007     2700      1     2.507
## 76999  34147153377 4.107 2007     2700      1     3.050
## 77015  33847667949 4.049 2007     2700      1     3.197
## 77028  33947671432 3.747 2007     2700      1     2.690
## 77077  34247476430 3.747 2007     2700      1     2.590
## 77081  34247498181 3.668 2007     2700      1     2.511
## 77084  34247545653 3.639 2007     2700      1     2.687
## 77570  34247282706 4.091 2007     2700      1     3.139
## 77574  34147113263 3.593 2007     1000      2     2.536
## 77582  34047185285 3.728 2007     2700      1     2.671
## 77604  33947210018 3.884 2007     2700      1     2.827
## 77607  33947595236 5.510 2007     2700      1     4.558
## 77643  33947519044 3.632 2007     2700      1     2.680
## 77661  33947196096 4.621 2007     2700      1     3.564
## 77664  33947279831 4.110 2007     2700      1     3.158
## 77685  33947271991 3.734 2007     2700      1     2.677
## 77712  33847705701 5.514 2007     2700      1     4.357
## 77720  33847723397 4.950 2007     2700      1     4.098
## 77730  33847393654 3.396 2007     2700      1     2.544
## 77752  33846926962 3.715 2007     2700      1     2.763
## 77761  33847253945 4.880 2007     2700      1     3.723
## 77762  33847267289 3.396 2007     2700      1     2.544
## 77772  33847368173 5.051 2007     2700      1     3.994
## 77773  33847369469 4.655 2007     2700      1     3.598
## 77776  33847381116 5.213 2007     2700      1     4.156
## 77777  33847391669 4.160 2007     2700      1     3.103
```

```
## 77779  33847404482 4.431 2007     2700      1     3.374
## 77780  33847406983 3.747 2007     2700      1     2.590
## 77781  33847407101 3.987 2007     2700      1     2.930
## 77786  33847413701 3.920 2007     2700      1     2.863
## 77819  33947356563 3.540 2007     2700      1     2.688
## 77839  33947710793 4.749 2007     2700      1     3.692
## 78473  34547138943 3.721 2007     2700      2     2.664
## 78607  33847353773 4.810 2007     2700      1     3.858
## 78631  33847692398 3.435 2007     2700      1     2.583
## 78651  33847165011 4.126 2007     2700      1     3.069
## 78654  33847103656 4.307 2007     2700      1     3.455
## 78655  33847174197 4.203 2007     2700      1     3.351
## 78704  33846964515 3.819 2007     2700      1     2.762
## 78707  33846996345 5.465 2007     2700      1     4.613
## 78728  33846857305 3.602 2007     2700      1     2.545
## 78737  33846866271 4.955 2007     2700      1     3.898
## 78738  33846875851 5.005 2007     2700      1     3.948
## 78746  33846863146 4.895 2007     2700      1     3.838
## 78749  33846878293 3.715 2007     2700      1     2.558
## 78780  33846571916 4.217 2007     2700      1     3.160
## 78792  33846075235 4.032 2007     2700      1     2.975
## 78810  33846661903 3.959 2007     2700      1     2.902
## 78812  33846673816 4.935 2007     2700      1     3.878
## 78813  33846688887 5.379 2007     2700      1     4.222
## 78816  33846783284 3.715 2007     2700      1     2.658
## 78821  33846787125 3.835 2007     2700      1     2.678
## 78822  33846794896 5.202 2007     2700      1     4.145
## 78825  33846810181 3.564 2007     2700      1     2.507
## 78832  33846817495 4.130 2007     2700      1     3.278
## 78836  33846829811 3.983 2007     2700      1     3.131
## 78839  33846839628 3.915 2007     2700      1     2.758
## 78843  33846850600 3.777 2007     2700      1     2.620
## 78847  33846913993 3.531 2007     2700      1     2.679
## 78899  33847610733 4.122 2007     2700      1     3.065
## 78900  33847617509 3.987 2007     2700      1     2.830
## 79061  34248581736 3.523 2007     2700      1     2.571
## 79359  33846524262 3.396 2007     2700      1     2.544
## 79377  33846294746 5.096 2007     2700      1     4.039
## 79386  33846413027 3.572 2007     2700      1     2.515
## 79416  33846126939 4.946 2007     2700      1     3.789
## 79420  33846175081 3.777 2007     2700      1     2.720
## 79427  33846125284 5.173 2007     2700      1     4.116
## 79428  33846129459 4.066 2007     2700      1     3.114
## 79489  33845970192 5.126 2007     2700      1     4.069
## 79491  33845972969 3.789 2007     2700      1     2.632
## 79492  33845979811 3.889 2007     2700      1     2.732
## 79495  33845994359 4.653 2007     2700      1     3.496
## 79496  33846006533 3.661 2007     2700      1     2.504
## 79498  33846012847 3.841 2007     2700      1     2.684
## 79499  33846017279 4.078 2007     2700      1     2.921
```

```
## 79510  33845984538 4.757 2007     2700      1     3.700
## 79511  33846021262 4.182 2007     2700      1     3.125
## 79512  33846026341 4.599 2007     2700      1     3.542
## 79513  33846027094 4.070 2007     2700      1     2.913
## 79540  33845870422 3.734 2007     2700      1     2.882
## 79560  33846037400 3.935 2007     2700      1     2.983
## 79602  33846668526 3.874 2007     2700      1     2.817
## 79607  33846671273 3.654 2007     2700      1     2.597
## 79609  33846672853 4.091 2007     2700      1     3.034
## 79671  33847184284 3.632 2007     2700      1     2.680
## 79978  34250801440 3.925 2007     2700      1     3.073
## 80011  34347336319 3.646 2007     2700      1     2.589
## 80013  34347343739 3.668 2007     2700      1     2.611
## 80106  35649016098 4.629 2007     2700      1     3.572
## 80111  35848953625 3.728 2007     2700      1     2.671
## 80112  35848954757 4.777 2007     2700      1     3.825
## 80217  42449096529 3.935 2007     2700      1     3.083
## 80636  58149083708 4.778 2008     2700      1     3.724
## 80643  58149218477 3.923 2008     2700      1     3.073
## 80683  58049213696 4.583 2008     2700      1     3.529
## 80684  58049216683 3.839 2008     2700      1     2.785
## 80728  57349192265 3.722 2008     2700      1     2.668
## 80744  57449112519 4.616 2008     2700      1     3.562
## 80797  57349093766 3.663 2008     2700      1     2.609
## 80814  57349133922 3.823 2008     2700      1     2.769
## 80816  57349184260 3.522 2008     2700      1     2.672
## 80943  56149095348 3.843 2008     2700      1     2.893
## 80980  57449084208 4.389 2008     2700      1     3.335
## 80983  57449095883 3.663 2008     2700      1     2.609
## 81062  58149387739 4.794 2008     2700      1     3.740
## 81349  64349102964 3.423 2008     2700      1     2.573
## 82714  57049128384 3.717 2008     2700      1     2.563
## 82785  56649112752 4.452 2008     2700      1     3.398
## 82845  56249128276 4.627 2008     2700      1     3.573
## 82846  56249141853 3.916 2008     2700      1     2.862
## 82880  55849115999 3.659 2008     1000      2     2.605
## 82899  55549095137 3.894 2008     2700      1     2.840
## 82911  55849084700 3.909 2008     2700      1     2.855
## 82920  55849139852 3.698 2008     2700      1     2.748
## 82950  54049103289 3.570 2008     2700      1     2.516
## 83109  58149467012 3.754 2008     2700      1     2.700
## 83609  55249119765 4.514 2008     2700      1     3.460
## 83619  54349091094 3.478 2008     2700      1     2.528
## 83648  55549139481 4.069 2008     2700      1     3.015
## 83661  54949144335 4.216 2008     2700      1     3.162
## 83714  54049121079 3.886 2008     2700      1     2.832
## 83728  54049104424 3.768 2008     2700      1     2.818
## 83729  54049106963 4.294 2008     2700      1     3.240
## 83794  53749095835 4.145 2008     2700      1     3.091
## 83849  54849146500 4.137 2008     2700      1     3.083
```

```
## 83859  54949141082 3.589 2008     1704      3     2.535
## 83861  50249096174 4.078 2008     2700      1     3.128
## 83872  51749084267 3.451 2008     2700      1     2.601
## 83891  52949103244 3.781 2008     2700      1     2.831
## 83892  52949106040 4.278 2008     2700      1     3.224
## 83897  52949144507 4.137 2008     2700      1     3.083
## 83898  52949144863 3.722 2008     2700      1     2.772
## 83901  53149119591 3.401 2008     2700      1     2.551
## 83969  55849112296 3.599 2008     2700      1     2.649
## 83975  55849150592 4.342 2008     2700      1     3.288
## 84213  65549087532 3.839 2008     2700      1     2.889
## 84545  52649151888 3.916 2008     2700      1     2.966
## 84550  52049088955 4.048 2008     2700      1     2.894
## 84553  52049108262 4.242 2008     2700      1     3.188
## 84555  52449113766 3.626 2008     2700      1     2.776
## 84585  52649120886 4.627 2008     2700      1     3.473
## 84598  52249089796 3.843 2008     2700      1     2.789
## 84607  51949095058 4.868 2008     2700      1     3.918
## 84682  51449097277 4.098 2008     2700      1     2.944
## 84728  50949089029 4.489 2008     2700      1     3.435
## 84734  51649088233 4.453 2008     2700      1     3.399
## 84761  51349169034 3.673 2008     2700      1     2.619
## 84831  52749095854 3.593 2008     2700      1     2.539
## 85169  59049103004 3.503 2008     2700      1     2.653
## 85277  65249106599 3.570 2008     2700      1     2.620
## 85688  50149098605 3.746 2008     1000      2     2.692
## 85697  50449091665 3.772 2008     2700      1     2.618
## 85698  50449093260 4.360 2008     2700      1     3.306
## 85709  50449095362 4.693 2008     2700      1     3.639
## 85728  49749143019 3.576 2008     2700      1     2.726
## 85743  50149121231 4.030 2008     1000      2     2.976
## 85752  49949087905 4.118 2008     2700      1     3.168
## 85753  49949091204 3.678 2008     2700      1     2.624
## 85765  49449098216 3.958 2008     2700      1     3.008
## 85828  49449100639 4.042 2008     2700      1     2.988
## 85831  49249087104 4.356 2008     2700      1     3.406
## 85886  49249104701 5.150 2008     2700      1     4.096
## 85962  48149090717 3.855 2008     2700      1     2.801
## 85972  48149110000 3.647 2008     2700      1     2.797
## 86469  47849130092 4.213 2008     2700      1     3.159
## 86470  47849130608 3.731 2008     2700      1     2.677
## 86483  47849100908 3.663 2008     2700      1     2.509
## 86505  47949092540 4.171 2008     2700      1     3.221
## 86507  47949114668 5.232 2008     2700      1     4.078
## 86564  47549096598 3.768 2008     2700      1     2.818
## 86565  47549105407 4.511 2008     2700      1     3.357
## 86620  46949093558 3.886 2008     2700      1     2.732
## 86634  47049126504 3.827 2008     2700      1     2.773
## 86680  46449113799 4.877 2008     2700      1     3.723
## 86684  46449128469 3.546 2008     1000      2     2.596
```

```
## 86729  45749146166 3.855 2008     2700      1     3.005
## 86742  46449098821 3.879 2008     2700      1     2.929
## 86743  46749084688 3.843 2008     2700      1     2.789
## 86746  46749099010 4.045 2008     2700      1     3.195
## 86749  46749117203 3.712 2008     2700      1     2.658
## 86755  46749156706 4.460 2008     2700      1     3.306
## 86780  47949131252 3.387 2008     2700      1     2.537
## 86795  48749129542 4.693 2008     2700      1     3.743
## 87045  55149104425 3.768 2008     2700      1     2.818
## 87540  45949092534 4.637 2008     2700      1     3.583
## 87543  45949107235 3.831 2008     2700      1     2.777
## 87544  46749131798 5.204 2008     2700      1     4.150
## 87592  45549084293 4.115 2008     2700      1     3.061
## 87602  45349084317 4.501 2008     2700      1     3.447
## 87603  45349103702 4.595 2008     2700      1     3.441
## 87663  45149127545 4.115 2008     2700      1     3.061
## 87691  44849103815 4.757 2008     2700      1     3.907
## 87696  44849115945 4.501 2008     2700      1     3.447
## 87735  45749100010 3.951 2008     2700      1     2.897
## 87739  45749114895 4.710 2008     2700      1     3.556
## 87740  45749124735 3.409 2008     2700      1     2.559
## 87745  45749149476 3.642 2008     2700      1     2.588
## 87792  44949119632 3.451 2008     2700      1     2.601
## 87828  46349104324 3.831 2008     2700      1     2.777
## 88491  44349145194 4.455 2008     2700      1     3.401
## 88496  44449106055 3.641 2008     1000      2     2.587
## 88499  44349153009 3.978 2008     2700      1     2.924
## 88500  44349155992 4.235 2008     2700      1     3.285
## 88501  44349165737 5.327 2008     2700      1     4.377
## 88546  44249122799 5.169 2008     2700      1     4.115
## 88634  43549085056 3.948 2008     2700      1     3.098
## 88637  43549110959 4.509 2008     2700      1     3.355
## 88680  43049170245 4.574 2008     2700      1     3.624
## 88682  43049174930 3.968 2008     2700      1     2.814
## 88683  43049176549 4.334 2008     2700      1     3.384
## 88750  42949149810 3.901 2008     2700      1     3.051
## 88757  43049090817 4.174 2008     2700      1     3.120
## 88759  43049094836 3.740 2008     2700      1     2.790
## 88760  43049095682 4.414 2008     2700      1     3.360
## 88784  43049136576 3.570 2008     2700      1     2.516
## 88857  44449096577 3.731 2008     2700      1     2.677
## 88859  44449122269 3.944 2008     2700      1     3.094
## 89498  42449106956 4.533 2008     2700      1     3.683
## 89500  42449141936 3.802 2008     2700      1     2.648
## 89501  44149123099 4.666 2008     2700      1     3.816
## 89566  42249112261 3.621 2008     1000      2     2.567
## 89567  42249098105 4.181 2008     2700      1     3.127
## 89601  42249083269 4.686 2008     2700      1     3.836
## 89602  42249092533 4.619 2008     2700      1     3.769
## 89605  48149094782 3.698 2008     2700      1     2.544
```

```
## 89617  42049123233 4.048 2008     2700      1     3.198
## 89645  40849108663 3.574 2008     1000      2     2.520
## 89663  41649112174 3.663 2008     2700      1     2.509
## 89739  41849105806 3.930 2008     2700      1     2.776
## 89751  41849148289 3.992 2008     2700      1     2.938
## 89783  43249086840 3.683 2008     2700      1     2.529
## 89784  43249087616 3.859 2008     2700      1     2.805
## 89878  46349092759 3.712 2008     2700      1     2.558
## 90277  43049134750 3.647 2008     2700      1     2.593
## 90284  41449104685 3.576 2008     2700      1     2.522
## 90287  41449112372 5.371 2008     2700      1     4.317
## 90289  41449113183 4.795 2008     2700      1     3.741
## 90338  40949100835 3.944 2008     2700      1     2.994
## 90347  40949102607 5.086 2008     2700      1     3.932
## 90361  40949127393 4.633 2008     2700      1     3.783
## 90407  40849124009 4.384 2008     2700      1     3.330
## 90409  40849142102 4.916 2008     2700      1     3.762
## 90412  40749132230 3.356 2008     1000      2     2.506
## 90415  40749141596 4.476 2008     2700      1     3.526
## 90432  41649090186 4.800 2008     2700      1     3.746
## 90458  40449141326 4.620 2008     2700      1     3.566
## 90464  40449104676 4.174 2008     2700      1     3.324
## 90509  39749103691 3.615 2008     2700      1     2.765
## 90521  40449086087 4.445 2008     2700      1     3.291
## 90523  40549084318 3.879 2008     2700      1     2.825
## 90545  40749105777 3.541 2008     2700      1     2.691
## 91382  39549099909 3.693 2008     2700      1     2.639
## 91383  39549100109 3.988 2008     2700      1     2.934
## 91384  39549102834 4.683 2008     2700      1     3.629
## 91386  39549111874 3.958 2008     2700      1     2.904
## 91388  39549096658 4.328 2008     2700      1     3.174
## 91415  39749102149 4.416 2008     2700      1     3.566
## 91426  39349097864 3.575 2008     1000      2     2.521
## 91488  38949105879 4.256 2008     2700      1     3.306
## 91489  38949106566 3.754 2008     2700      1     2.700
## 91490  38949188680 3.871 2008     2700      1     2.817
## 91499  38949152338 4.285 2008     2700      1     3.231
## 91520  38649125875 4.145 2008     2700      1     3.091
## 91534  38649100024 3.604 2008     2700      1     2.754
## 91552  38849130851 3.827 2008     2700      1     2.773
## 91565  38849174979 4.368 2008     2700      1     3.418
## 91567  38849182472 3.972 2008     2700      1     2.918
## 91602  39849097577 4.313 2008     2700      1     3.159
## 91603  39849099704 3.516 2008     2700      1     2.666
## 91605  39849102541 3.615 2008     2700      1     2.665
## 91725  43349085680 3.955 2008     2700      1     2.901
## 91726  43349087466 3.772 2008     2700      1     2.718
## 91730  43349096215 3.707 2008     2700      1     2.653
## 91738  43349107306 3.975 2008     2700      1     2.921
## 91845  38949168818 3.621 2008     1300      2     2.771
```

```
## 91993  38749111066 5.445 2008     2700      1     4.391
## 91995  38749127156 4.616 2008     2700      1     3.562
## 92001  38749133726 3.560 2008     1000      2     2.506
## 92033  38549088096 4.842 2008     2700      1     3.788
## 92037  38549134052 4.324 2008     2700      1     3.270
## 92040  38549177642 4.453 2008     2700      1     3.399
## 92046  38349164176 4.531 2008     2700      1     3.477
## 92068  38349049478 5.357 2008     2700      1     4.303
## 92111  38049146378 4.148 2008     2700      1     3.094
## 92113  38049169559 4.935 2008     2700      1     3.881
## 92134  38049062200 4.629 2008     2700      1     3.575
## 92136  38049077991 3.843 2008     2700      1     2.789
## 92140  38049082149 4.298 2008     2700      1     3.244
## 92148  38149064606 3.631 2008     2700      1     2.681
## 92161  38149141767 3.958 2008     2700      1     2.804
## 92193  37549009899 4.066 2008     2700      1     3.116
## 92197  37549020305 3.582 2008     2700      1     2.528
## 92199  37549027612 4.244 2008     2700      1     3.190
## 92203  37549064327 3.776 2008     2700      1     2.722
## 92219  38149074110 3.934 2008     2700      1     2.880
## 92231  38349085066 3.610 2008     2700      1     2.556
## 92494  41149157013 3.547 2008     2700      1     2.597
## 92576  42149177179 4.216 2008     2700      1     3.162
## 92681  44949253343 4.140 2008     2700      1     2.986
## 92735  46749123025 3.731 2008     2700      1     2.677
## 92922  57149102701 3.912 2008     2700      1     3.062
## 93538  74049108635 3.752 2009     2700      1     2.653
## 93541  73949086506 3.919 2009     2700      1     2.721
## 93592  73549083853 3.805 2009     2700      1     2.706
## 93602  77954601328 4.143 2009     2700      1     3.044
## 93632  70350733579 3.757 2009     2700      1     2.559
## 93633  70449636557 4.842 2009     2700      1     3.644
## 93634  71849097187 3.961 2009     2700      1     2.763
## 93635  71849098875 3.947 2009     2700      1     2.848
## 93652  70350733425 3.985 2009     2700      1     2.991
## 93654  70350738289 4.729 2009     2700      1     3.630
## 93659  72549091935 4.039 2009     2700      1     3.145
## 93707  71849118976 3.968 2009     2700      1     2.869
## 93720  71549150905 3.500 2009     2700      1     2.606
## 93760  73349094732 4.184 2009     2700      1     3.190
## 93764  73349134686 4.036 2009     2700      1     2.937
## 93775  84870055817 4.573 2009     2700      1     3.679
## 93843  65549167028 3.801 2009     2700      1     2.702
## 93889  67650760521 3.826 2009     2700      1     2.727
## 93962  70749114018 3.645 2009     2700      1     2.546
## 93964  70749145982 3.814 2009     2700      1     2.715
## 93975  72049129932 3.889 2009     2700      1     2.790
## 93995  72849125358 3.858 2009     2700      1     2.864
## 94224  77049105843 3.904 2009     2700      1     2.910
## 94281  77649213935 3.838 2009     2700      1     2.844
```

```
## 95604  72049129430 4.683 2009     2700      1     3.584
## 95638  70449652591 3.805 2009     2700      1     2.706
## 95723  70349873173 5.082 2009     2700      1     4.088
## 95726  70449636163 5.302 2009     2700      1     4.308
## 95743  77950669604 5.246 2009     2700      1     4.147
## 95828  70449355084 3.784 2009     2700      1     2.685
## 95831  70449377960 3.743 2009     2700      1     2.644
## 95834  70449394580 4.039 2009     2700      1     2.940
## 95842  70449448312 4.397 2009     2700      1     3.298
## 96014  73949153085 4.118 2009     2700      1     3.019
## 96406  73349109845 3.667 2009     2700      2     2.568
## 96572  84881264341 4.415 2009     2700      1     3.316
## 96673  70350064091 3.705 2009     2700      1     2.711
## 96696  70350458703 4.870 2009     2700      1     3.876
## 96716  70350212398 4.751 2009     2700      1     3.652
## 96717  70350220313 3.629 2009     2700      1     2.530
## 96718  70350227015 3.975 2009     2700      1     2.777
## 96719  70350236228 4.659 2009     2700      1     3.765
## 96753  70349506141 4.798 2009     2700      1     3.904
## 96773  70349943834 3.625 2009     1000      2     2.526
## 96854  65749111503 3.762 2009     2700      1     2.663
## 96867  69949132983 3.549 2009     2700      1     2.655
## 96870  69949144253 3.696 2009     2700      1     2.597
## 96895  70349481938 4.575 2009     2700      1     3.476
## 96897  70349610473 4.853 2009     2700      1     3.754
## 96899  70349625757 5.493 2009     2700      1     4.394
## 96900  70349629019 5.086 2009     2700      1     3.987
## 96919  70349684791 4.187 2009     2700      1     2.989
## 96920  70349687102 4.017 2009     2700      1     2.918
## 96953  70350455774 3.691 2009     2700      1     2.697
## 97145  77952988108 4.600 2009     2700      1     3.501
## 97523  70349437154 3.951 2009     2700      1     2.852
## 97545  70349308645 3.601 2009     2700      1     2.502
## 97546  70349308646 4.248 2009     2700      1     3.149
## 97548  70349319816 3.705 2009     2700      1     2.711
## 97550  70349335796 4.765 2009     2700      1     3.666
## 97633  70349238733 4.938 2009     2700      1     4.044
## 97634  70349244812 4.857 2009     2700      1     3.963
## 97636  70349254065 3.850 2009     2700      1     2.751
## 97666  66849121699 3.710 2009     2700      1     2.611
## 97668  66849128097 3.947 2009     2700      1     2.848
## 97740  70149093912 5.343 2009     2700      1     4.145
## 97741  70149098511 3.748 2009     2700      1     2.649
## 97748  70249111091 3.839 2009     1000      2     2.641
## 97755  70149112350 4.282 2009     2700      1     3.084
## 97852  68949150765 3.725 2009     2700      1     2.731
## 97861  69849097352 3.958 2009     2700      1     2.859
## 97897  70349205537 3.506 2009     2700      1     2.512
## 97906  70349213381 3.691 2009     2700      1     2.797
## 97912  70349220917 3.757 2009     2700      1     2.658
```

```
## 97930  70349636711 4.014 2009     2700      1     2.915
## 98067  73449108112 3.846 2009     2700      1     2.852
## 98073  73449123799 4.061 2009     2700      1     2.962
## 98210  75349114036 3.766 2009     2700      1     2.667
## 98756  69249219296 5.111 2009     2700      1     4.012
## 98759  69249229866 3.850 2009     2700      1     2.751
## 98778  69449098588 4.214 2009     2700      1     3.115
## 98813  68949090752 4.850 2009     2700      1     3.652
## 98814  68949092632 4.197 2009     2700      1     3.303
## 98815  68949102467 3.830 2009     2700      1     2.731
## 98816  68949133346 4.250 2009     2700      1     3.052
## 98817  68949171916 3.775 2009     2700      1     2.881
## 98921  68849084949 4.231 2009     2700      1     3.033
## 98930  68849119553 4.573 2009     2700      1     3.474
## 98931  68849123400 4.335 2009     2700      1     3.236
## 98992  68249154875 5.117 2009     2700      1     4.018
## 99008  68349117223 3.919 2009     2700      1     3.025
## 99032  68849104003 3.613 2009     2700      1     2.514
## 99033  68849112640 3.607 2009     2700      1     2.508
## 99081  67650635227 3.991 2009     2700      1     2.793
## 99104  67651160419 4.452 2009     2700      1     3.558
## 99146  69349094143 4.819 2009     2700      1     3.720
## 99564  68049123594 4.967 2009     2700      1     3.868
## 99616  68149144251 3.826 2009     2700      1     2.727
## 99644  67651218924 4.011 2009     2700      1     2.912
## 99664  67650391945 3.818 2009     2700      1     2.620
## 99679  67650318719 3.926 2009     2700      1     2.728
## 99686  67650899270 4.135 2009     2700      1     3.036
## 99688  67650909125 3.968 2009     2700      1     2.869
## 99748  68149112408 3.440 2009     2700      1     2.546
## 99827  67649664439 3.975 2009     2700      1     3.081
## 99858  67649814915 3.743 2009     2700      1     2.545
## 99874  67649967683 4.202 2009     2700      1     3.308
## 99913  68049122014 3.725 2009     2700      1     2.831
## 99961  68849122050 3.988 2009     2700      1     2.889
## 99962  68849124811 4.132 2009     2700      1     3.238
## 100692 67649390890 3.995 2009     2700      1     2.896
## 100705 67649218781 4.174 2009     2700      1     2.976
## 100727 67149116154 4.250 2009     2700      1     3.052
## 100761 67449110743 5.734 2009     2700      1     4.536
## 100762 67449152290 4.549 2009     2700      1     3.655
## 100764 67449164373 3.944 2009     2700      1     2.950
## 100831 67149142042 5.168 2009     2700      1     4.069
## 100844 67049116139 3.882 2009     2700      1     2.783
## 100938 66649113371 4.055 2009     2700      1     2.857
## 100947 66449100700 4.172 2009     2700      1     2.974
## 101089 67651005390 3.748 2009     2700      1     2.754
## 101093 67651006589 3.433 2009     2700      1     2.539
## 101786 67650169237 4.101 2009     2700      1     3.002
## 101788 65849187703 5.312 2009     2700      1     4.213
```

```
## 101794 66249084112 4.991 2009     2700      1     3.892
## 101808 66149152644 4.179 2009     2700      1     2.981
## 101854 66149118225 3.814 2009     2700      1     2.715
## 101861 66249137759 3.789 2009     1000      2     2.690
## 101904 66149123748 3.968 2009     1000      2     2.869
## 101942 65549163328 4.148 2009     2700      1     3.154
## 101994 65449136284 3.759 2009     1303      7     2.660
## 102096 66449111715 3.766 2009     2700      1     2.667
## 102501 65549092261 3.656 2009     1000      2     2.557
## 102867 65449152185 5.067 2009     2700      1     3.968
## 102988 64749093541 4.983 2009     2700      1     3.884
## 103006 64749086505 4.036 2009     2700      1     2.937
## 103008 64749114159 4.724 2009     2700      1     3.526
## 103061 62549134121 3.909 2009     1000      2     2.810
## 103846 63249124498 3.838 2009     2700      1     2.739
## 103851 63349094964 3.846 2009     2700      1     2.952
## 103903 62749097289 4.481 2009     2700      1     3.587
## 103904 62749131642 4.113 2009     2700      1     2.915
## 103921 62649136134 4.286 2009     2700      1     3.292
## 103922 62649155822 4.030 2009     2700      1     3.036
## 103984 61849137328 3.830 2009     2700      1     2.731
## 103998 62349094336 3.870 2009     2700      1     2.771
## 104003 62249133709 4.143 2009     1000      2     3.044
## 104012 62149120632 4.393 2009     2700      1     3.399
## 104068 61849118968 4.245 2009     2700      1     3.251
## 104126 60049098815 4.857 2009     2700      1     3.758
## 104132 61549094077 5.418 2009     2700      1     4.524
## 104148 62449262646 3.640 2009     2700      1     2.646
## 105045 60649118584 4.654 2009     2700      1     3.660
## 105055 60649090430 4.316 2009     2700      1     3.217
## 105061 61449183682 5.109 2009     2700      1     4.010
## 105074 61449162171 3.752 2009     2700      1     2.653
## 105118 60849115270 5.804 2009     2700      1     4.705
## 105126 60749089639 3.858 2009     2700      1     2.759
## 105140 59849099670 3.904 2009     2700      1     3.010
## 105210 60249096283 3.797 2009     2700      1     2.803
## 105211 60249099307 4.478 2009     2700      1     3.280
## 105240 59249084303 4.172 2009     2700      1     3.073
## 105242 59449101337 4.234 2009     2700      1     3.036
## 105251 59749096464 4.506 2009     2700      1     3.407
## 105271 58949085856 3.525 2009     2700      1     2.631
## 105282 61749094289 4.261 2009     2700      1     3.063
## 105283 61749094852 3.775 2009     2700      1     2.577
## 105852 59449094890 4.255 2009     2700      1     3.261
## 105884 58749094444 5.536 2009     2700      1     4.338
## 105903 58749109459 3.607 2009     2700      1     2.613
## 105947 58349091054 3.908 2009     2700      1     2.710
## 105956 58249110796 3.725 2009     1000      2     2.527
## 105970 58249088833 3.991 2009     2700      1     2.793
## 105990 58149385407 4.564 2009     2700      1     3.465
```

```
## 105992 58149389215 5.900 2009     2700      1     4.801
## 106003 58149385649 4.940 2009     2700      1     3.742
## 106031 61749090233 4.005 2009     2700      1     2.906
## 106037 61749102618 4.145 2009     2700      1     3.046
## 106063 57749092004 3.743 2009     2700      1     2.644
## 106066 57749093663 3.623 2009     2700      1     2.524
## 106107 58249087710 3.985 2009     2700      1     2.886
## 106187 61449122079 3.671 2009     2700      1     2.777
## 106195 61449149483 4.101 2009     2700      1     3.002
## 107838 78650858488 4.531 2010     2700      1     3.388
## 107864 78650664067 3.685 2010     2700      1     2.542
## 107918 78649906060 4.743 2010     2700      1     3.600
## 108011 78649880804 4.124 2010     2700      1     2.981
## 108234 78649822277 3.520 2010     2700      1     2.581
## 108250 78649888666 3.441 2010     2700      1     2.502
## 108252 78649895980 3.904 2010     2700      1     2.761
## 109519 78649429727 4.440 2010     2700      1     3.297
## 109522 78649439268 4.337 2010     2700      1     3.194
## 109593 78549288927 3.760 2010     2700      1     2.821
## 109846 78149330579 3.601 2010     2700      1     2.563
## 109891 78049510428 4.327 2010     2700      1     3.084
## 109892 78049515807 4.522 2010     2700      1     3.379
## 109919 78049407029 4.053 2010     2700      1     2.910
## 110787 77957768543 4.019 2010     2700      1     2.876
## 110790 78049428879 4.101 2010     2700      1     2.958
## 110803 78049337151 3.858 2010     2700      1     2.715
## 110805 78049351929 4.511 2010     2700      1     3.368
## 110864 77958149840 4.229 2010     2700      1     2.986
## 111024 77957934893 3.955 2010     2700      1     2.712
## 111045 77957928301 3.775 2010     2700      1     2.632
## 111115 77957725001 3.841 2010     2700      1     2.698
## 111130 77957671700 3.646 2010     2700      1     2.503
## 111485 80051785529 3.594 2010     2700      1     2.556
## 112042 77957333252 3.808 2010     2700      1     2.665
## 112043 77957341503 4.430 2010     2700      1     3.187
## 112124 77957139539 3.680 2010     1000      2     2.537
## 112173 77956929653 3.485 2010     2700      1     2.546
## 112174 77956942602 3.749 2010     2700      1     2.506
## 112259 77956634601 3.549 2010     2700      1     2.610
## 113861 77955877759 5.447 2010     2700      1     4.204
## 113873 77955760684 3.730 2010     2700      1     2.587
## 114067 77955357246 3.858 2010     2700      1     2.715
## 114203 77955296243 3.591 2010     2700      1     2.553
## 114219 77955286305 3.791 2010     2700      1     2.648
## 114324 77955573674 4.157 2010     2700      1     3.218
## 114349 77956057548 3.989 2010     2700      1     2.846
## 115221 77955066199 5.339 2010     2700      1     4.196
## 115230 77954988618 4.006 2010     2700      1     2.863
## 115361 77954635020 4.231 2010     2700      1     3.088
## 115455 77954421915 4.566 2010     2700      1     3.423
```

```
## 115465 77954374543 4.143 2010     2700      1     2.900
## 115552 77954183132 3.685 2010     2700      1     2.542
## 115604 77954974268 3.717 2010     2700      1     2.574
## 116866 77953956205 4.405 2010     2700      1     3.367
## 117020 77953577951 4.518 2010     2700      1     3.375
## 117076 77952962769 3.929 2010     2700      1     2.686
## 117084 77952977625 3.827 2010     2700      1     2.684
## 117093 77952998648 3.957 2010     2700      1     3.018
## 117101 77953013128 3.801 2010     2700      1     2.658
## 117127 77953173228 4.019 2010     2700      1     2.981
## 117148 77952995935 3.649 2010     2700      1     2.506
## 117177 77952766987 3.851 2010     2700      1     2.708
## 117245 77953360245 3.933 2010     2700      1     2.994
## 118168 77952757705 4.297 2010     2700      1     3.358
## 118169 77952760196 3.757 2010     2700      1     2.514
## 118172 77952784840 4.247 2010     2700      1     3.308
## 118242 77952470787 3.736 2010     2700      1     2.797
## 118247 77952356171 3.714 2010     2700      1     2.571
## 118249 77952389110 3.851 2010     2700      1     2.708
## 118307 77952301687 4.418 2010     2700      1     3.275
## 118351 77951868945 4.370 2010     2700      1     3.227
## 118352 77951874018 4.849 2010     2700      1     3.910
## 118367 77951894040 4.247 2010     2700      1     3.308
## 118392 77951589703 4.220 2010     2700      1     3.077
## 118410 77951861324 3.897 2010     2700      1     2.754
## 119263 77951639385 4.329 2010     2700      1     3.086
## 119325 77951179995 4.040 2010     2700      1     3.101
## 119326 77951195763 4.197 2010     2700      1     3.159
## 119394 77950894766 3.527 2010     2700      1     2.588
## 119487 77950847452 3.722 2010     2700      1     2.579
## 119497 77950477601 4.327 2010     2700      1     3.184
## 119499 77950486825 3.875 2010     2700      1     2.732
## 119582 77950475726 4.241 2010     2700      1     3.098
## 119583 77950478147 4.036 2010     2700      1     3.097
## 119585 77950494853 4.319 2010     2700      1     3.076
## 120218 77950246404 4.319 2010     2700      1     3.176
## 120282 77949511388 3.778 2010     2700      1     2.635
## 120453 77649329576 4.566 2010     2700      1     3.323
## 120619 77950355514 3.621 2010     2700      1     2.682
## 121458 77249162677 4.349 2010     2700      1     3.206
## 121459 77249165024 3.722 2010     2700      1     2.579
## 121517 77149173502 3.864 2010     2700      1     2.826
## 121526 77149120471 3.773 2010     2700      1     2.530
## 121527 77149136099 3.803 2010     2700      1     2.560
## 121596 76649086071 4.375 2010     2700      1     3.436
## 122262 74549178560 4.208 2010     1315      2     3.065
## 122299 74849098405 4.260 2010     2700      1     3.321
## 122328 74549140993 5.722 2010     2700      1     4.479
## 122329 74549159752 5.207 2010     2700      1     4.169
## 122330 74549184608 3.581 2010     2700      1     2.642
```

```
## 122358 74949117631 4.260 2010     2700      1     3.222
## 122403 74149095074 3.682 2010     2700      1     2.743
## 122451 74049098673 4.147 2010     2700      1     3.109
## 122453 74049114688 4.434 2010     2700      1     3.291
## 122489 74949089659 3.781 2010     2700      1     2.638
## 122867 77950651061 4.285 2010     2700      1     3.346
## 122994 77952974496 3.813 2010     2700      1     2.670
## 123000 77952987107 3.798 2010     2700      1     2.655
## 123134 77955050878 3.858 2010     2700      1     2.615
## 125455 84855370035 4.953 2011     2700      1     3.917
## 125459 84855425106 5.298 2011     2700      1     4.361
## 125468 84555187324 3.856 2011     2700      1     2.615
## 125541 84255194001 4.947 2011     2700      1     3.806
## 125573 83755178774 3.993 2011     2700      1     2.957
## 125605 83355163991 3.976 2011     2700      1     2.940
## 125606 83455210809 3.567 2011     2700      1     2.630
## 125608 83455235074 3.980 2011     2700      1     3.043
## 125659 83255182742 3.895 2011     2700      1     2.958
## 125678 83155182848 4.324 2011     2700      1     3.183
## 125681 83055177179 3.952 2011     1000      2     2.811
## 125689 83055168450 3.885 2011     2700      1     2.948
## 125690 83055173186 4.057 2011     2700      1     2.916
## 125692 83155168389 3.830 2011     2700      1     2.893
## 125693 83155177076 3.989 2011     2700      1     2.953
## 125777 79960457303 3.482 2011     2700      1     2.545
## 125944 82555168440 4.771 2011     2700      1     3.630
## 125993 84055200171 3.731 2011     2700      1     2.590
## 125994 84055217017 3.890 2011     2700      1     2.749
## 125995 84055217292 3.856 2011     2700      1     2.715
## 127205 82255175382 5.249 2011     2700      1     4.108
## 127216 81855218115 4.488 2011     2700      1     3.347
## 127225 81855161048 3.950 2011     2700      1     3.013
## 127277 81555213207 3.799 2011     2700      1     2.763
## 127287 81455154874 4.549 2011     2700      1     3.408
## 127289 81455159320 4.127 2011     2700      1     2.986
## 127302 81255208675 4.283 2011     2700      1     3.142
## 127362 80755130317 4.428 2011     2700      1     3.491
## 127364 80755188166 4.426 2011     2700      1     3.285
## 127437 80455162464 4.949 2011     2700      1     3.808
## 127457 80355135138 4.850 2011     2700      1     3.913
## 127458 80355148407 4.102 2011     2700      1     2.861
## 127461 80455179713 3.713 2011     2700      1     2.572
## 127572 80555136815 4.124 2011     2700      1     3.088
## 128312 80054769456 4.528 2011     2700      1     3.387
## 128352 80054714783 3.880 2011     2700      1     2.739
## 128374 80054003232 4.343 2011     2700      1     3.202
## 128462 80053469106 4.076 2011     2700      1     3.139
## 128473 80053510713 3.861 2011     2700      1     2.620
## 128565 80053402552 3.851 2011     2700      1     2.610
## 128577 80053426842 4.065 2011     2700      1     2.824
```

```
## 129148 80053359692 4.030 2011     2700      1     2.889
## 129173 84555215993 3.637 2011     2700      1     2.601
## 129220 80053069652 4.106 2011     2700      1     2.965
## 129236 80053065502 3.945 2011     2700      1     2.804
## 129326 80052834654 5.140 2011     2700      1     3.999
## 129352 80052792378 4.382 2011     2700      1     3.241
## 129501 80052462089 3.945 2011     2700      1     2.804
## 129504 80052515621 3.713 2011     2700      1     2.572
## 129538 80052482300 3.725 2011     2700      1     2.584
## 129610 80052371597 3.972 2011     2700      1     2.831
## 129611 80052376347 3.895 2011     2700      1     2.754
## 129619 80052436335 4.001 2011     2700      1     2.860
## 130467 80052358982 3.650 2011     2700      1     2.713
## 130507 80051720194 5.682 2011     2700      1     4.541
## 130522 80052047215 4.316 2011     2700      1     3.175
## 130523 80052063328 3.904 2011     2700      1     2.763
## 130603 84860389821 4.428 2011     2700      1     3.287
## 130622 80051723862 4.452 2011     2700      1     3.311
## 130627 80051831092 4.268 2011     2700      1     3.127
## 130629 84867338003 4.503 2011     2700      1     3.262
## 130677 80051640525 4.042 2011     2700      1     2.901
## 130688 80051571901 4.590 2011     2700      1     3.554
## 130690 80051775476 5.397 2011     2700      1     4.256
## 130735 80051468836 4.057 2011     2700      1     2.916
## 130778 79961102782 3.743 2011     2700      1     2.707
## 130779 79961108629 6.254 2011     2700      1     5.218
## 130780 79961116424 4.087 2011     2700      1     2.946
## 130798 79961040968 3.804 2011     2700      1     2.663
## 130883 79961006281 3.941 2011     2700      1     2.700
## 131346 79960855656 4.776 2011     2700      1     3.635
## 131363 79960704300 3.825 2011     2700      1     2.584
## 131364 79960709532 3.890 2011     2700      1     2.749
## 131435 79960575079 4.596 2011     2700      1     3.455
## 131438 79960611010 3.670 2011     2700      1     2.529
## 131494 79960230129 3.950 2011     2700      1     2.809
## 131599 79960056173 4.241 2011     2700      1     3.100
## 131600 79960066694 3.552 2011     2700      1     2.516
## 131724 79959997891 3.895 2011     2700      1     2.654
## 131727 79960015327 4.590 2011     2700      1     3.554
## 131728 79960015997 5.714 2011     2700      1     4.777
## 131731 79960027556 4.472 2011     2700      1     3.436
## 132437 79959794787 5.141 2011     2700      1     4.000
## 132474 79959772357 4.719 2011     2700      1     3.578
## 132530 79959478455 5.314 2011     2700      1     4.173
## 132531 79959478969 4.354 2011     2700      1     3.213
## 132542 79959392796 3.748 2011     2700      1     2.507
## 132545 79959426992 4.106 2011     2700      1     2.965
## 132546 79959487586 4.565 2011     2700      1     3.424
## 132606 79959363092 3.574 2011     1000      2     2.637
## 132615 79959297709 3.871 2011     2700      1     2.835
```

```
## 132662 79958724181 4.832 2011     2700      1     3.591
## 132668 79959316645 3.976 2011     2700      1     2.835
## 132736 79958195203 4.483 2011     2700      1     3.342
## 132738 79958226299 4.639 2011     2700      1     3.498
## 132745 79958066836 4.674 2011     2700      1     3.533
## 132790 79957795474 4.944 2011     2700      1     3.803
## 132794 79957877087 4.014 2011     2700      1     3.077
## 132801 79952593642 3.682 2011     2700      1     2.541
## 132832 79957520670 3.959 2011     2700      1     2.718
## 132838 79957805100 4.354 2011     2700      1     3.113
## 132839 79957808383 4.120 2011     2700      1     2.979
## 132850 79957889288 3.777 2011     2700      1     2.636
## 132905 79958782230 3.713 2011     2700      1     2.572
## 133601 79957444504 4.382 2011     2700      1     3.346
## 133602 79957445348 4.662 2011     2700      1     3.725
## 133644 79956204057 3.846 2011     2700      1     2.705
## 133684 79956311926 3.968 2011     1000      2     2.727
## 133688 79956189911 4.080 2011     2700      1     2.939
## 133690 79956218863 5.583 2011     2700      1     4.442
## 133696 79956258982 3.993 2011     2700      1     2.752
## 133759 79955909191 4.421 2011     2700      1     3.280
## 133811 79955823001 3.760 2011     2700      1     2.619
## 133813 79955877895 3.880 2011     2700      1     2.739
## 133897 79955624770 4.409 2011     2700      1     3.268
## 133939 79955517235 4.745 2011     2700      1     3.504
## 133996 79955115145 3.695 2011     2700      1     2.554
## 134076 79958006776 3.885 2011     2700      1     2.948
## 134662 79955376345 3.993 2011     2700      1     2.752
## 134719 79955009821 4.244 2011     2700      1     3.003
## 134722 79955016374 3.663 2011     2700      1     2.726
## 134723 79955035027 3.993 2011     2700      1     2.852
## 134724 79955038968 3.967 2011     2700      1     2.826
## 134759 79954561234 4.639 2011     2700      1     3.498
## 134775 79954487919 5.185 2011     2700      1     4.248
## 134778 79954553689 4.428 2011     2700      1     3.287
## 134783 79953848750 3.799 2011     2700      1     2.763
## 134788 79954542852 3.731 2011     2700      1     2.590
## 134883 79953748230 4.910 2011     2700      1     3.669
## 134959 79953297488 4.005 2011     2700      1     2.864
## 134968 79953307878 4.700 2011     2700      1     3.559
## 134978 79953329970 4.175 2011     2700      1     3.034
## 134979 79953331206 4.492 2011     2700      1     3.351
## 135559 79953225169 4.658 2011     2700      1     3.721
## 135593 79952941296 3.793 2011     2700      1     2.856
## 135629 79953232077 4.567 2011     2700      1     3.426
## 135637 79953118839 5.685 2011     2700      1     4.444
## 135705 79952795093 4.650 2011     2700      1     3.713
## 135762 79952376771 4.188 2011     2700      1     3.047
## 135766 79952424506 4.158 2011     2700      1     3.017
## 135793 79952476762 4.836 2011     2700      1     3.800
```

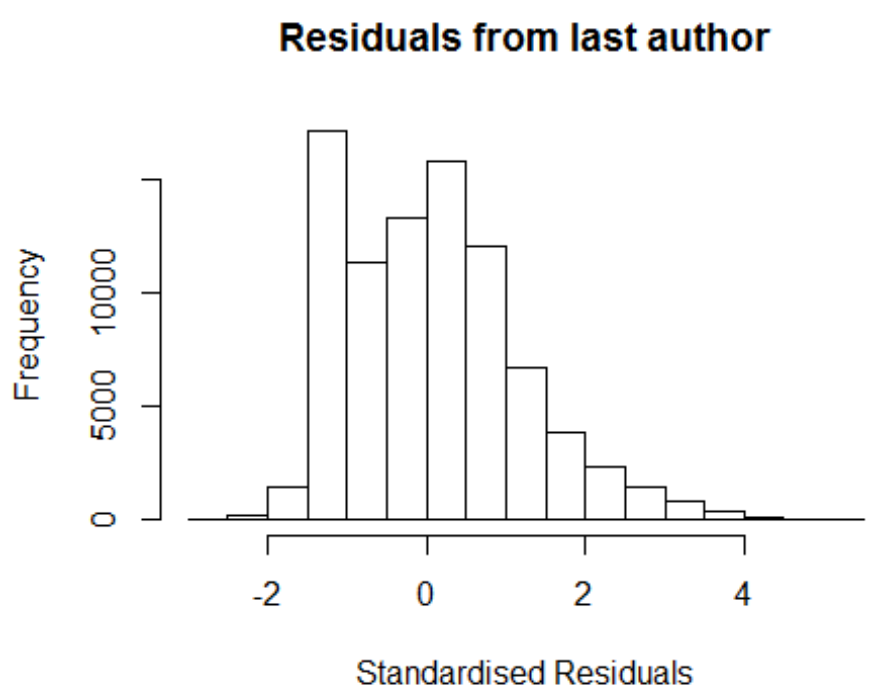
```
## 135911 79952363727 5.013 2011     2700      1     3.872
## 135923 79952267051 4.395 2011     2700      1     3.154
## 135924 79952271716 3.900 2011     2700      1     2.759
## 136005 79952219919 4.001 2011     2700      1     2.860
## 136036 79952570027 3.688 2011     2700      1     2.547
## 136042 79952749048 3.603 2011     2700      1     2.567
## 136941 79951961755 3.835 2011     2700      1     2.694
## 136954 79951978171 3.825 2011     2700      1     2.584
## 136956 79952204045 3.688 2011     2700      1     2.547
## 136959 79951778987 3.914 2011     2700      1     2.773
## 136973 79951850287 4.863 2011     2700      1     3.722
## 136988 79951821800 4.881 2011     2700      1     3.845
## 136997 79951699237 4.608 2011     2700      1     3.367
## 137041 79951514365 3.458 2011     1000      2     2.521
## 137056 79851493086 4.470 2011     2700      1     3.229
## 137064 79751522951 5.633 2011     2700      1     4.696
## 137065 79851475199 3.707 2011     2700      1     2.770
## 137169 79551566712 4.247 2011     2700      1     3.106
## 137251 79251491762 3.799 2011     2700      1     2.658
## 137265 79551523377 4.106 2011     2700      1     3.169
## 137278 79551666741 3.927 2011     2700      1     2.686
## 137775 79251501962 3.820 2011     2700      1     2.679
## 137776 79251525181 3.985 2011     2700      1     2.844
## 137777 79251527865 4.001 2011     2700      1     2.760
## 137786 78751642476 3.804 2011     2700      1     2.663
## 137800 78751610848 4.907 2011     2700      1     3.666
## 137805 79952148734 4.771 2011     2700      1     3.630
## 137813 78751519245 3.731 2011     2700      1     2.695
## 137814 78751532610 3.719 2011     2700      1     2.683
## 137817 78751667739 3.989 2011     2700      1     2.848
## 137857 78651351928 3.959 2011     2700      1     3.022
## 137860 78651378996 3.950 2011     2700      1     2.809
## 137861 78651379500 5.349 2011     2700      1     4.208
## 137865 78651393550 4.594 2011     2700      1     3.453
## 137867 78651505523 3.777 2011     2700      1     2.636
## 137896 78650893810 3.993 2011     2700      1     2.752
## 137905 78650940402 3.574 2011     2700      1     2.538
## 137909 78650949272 4.632 2011     2700      1     3.391
## 137988 78650987230 3.560 2011     2700      1     2.623
## 137995 78651245300 4.046 2011     2700      1     3.010
## 138009 79151481493 3.670 2011     2700      1     2.529
## 138381 79955523146 3.731 2011     2700      1     2.590
## 138610 80051964418 3.980 2011     2700      1     2.839
## 139175 79952183161 3.858 2011     2700      2     2.717
## 139231 79958822104 4.047 2011     2700      2     3.110
## 139857 84871293779 4.110 2012     2700      1     2.991
## 139901 84871314601 3.955 2012     2700      1     2.836
## 140049 84870877380 3.955 2012     2700      1     2.836
## 140050 84870887067 4.610 2012     2700      1     3.491
## 140140 84870549609 4.439 2012     2700      1     3.320
```

```
## 141706 84870012939 4.380 2012     2700      1     3.261
## 141721 84869792699 3.633 2012     2700      1     2.619
## 141725 84870312106 4.184 2012     2700      1     2.965
## 141912 84868691029 4.249 2012     2700      1     3.130
## 142043 84868305308 4.241 2012     2700      1     3.022
## 142202 84868609657 3.465 2012     2700      1     2.551
## 142208 84868613782 3.623 2012     2700      1     2.504
## 142976 84867828250 4.526 2012     2700      1     3.512
## 143180 84867221377 4.243 2012     2700      1     3.024
## 143205 84867437406 3.667 2012     2700      1     2.548
## 143276 84867273101 3.658 2012     2700      1     2.644
## 144201 84866555961 4.254 2012     2700      1     3.135
## 144204 84866628742 3.794 2012     2700      1     2.780
## 144331 84866098277 4.037 2012     2700      1     2.918
## 144357 84866271313 3.691 2012     2700      1     2.677
## 144361 84866395310 4.216 2012     2700      1     3.097
## 144362 84866395319 3.995 2012     2700      1     2.876
## 144505 84865395104 3.895 2012     2700      1     2.776
## 144517 84865987121 3.945 2012     2700      1     2.826
## 145514 84865454335 3.901 2012     2700      1     2.987
## 145517 84865487218 4.045 2012     2700      1     2.826
## 145540 84865701980 4.297 2012     2700      1     3.178
## 145542 84865820981 3.939 2012     2700      1     2.720
## 145641 84865320778 4.511 2012     2700      1     3.392
## 145642 84865323065 4.426 2012     2700      1     3.512
## 145875 84864842736 4.100 2012     2700      1     2.881
## 146034 84864286532 3.526 2012     2700      1     2.512
## 146664 84864242034 4.379 2012     2700      1     3.260
## 146677 84864138716 3.932 2012     2700      1     2.813
## 146785 84864008031 5.104 2012     2700      1     3.985
## 146790 84930197839 3.672 2012     2700      1     2.553
## 146935 84863923855 3.855 2012     2700      1     2.736
## 147014 84863524725 3.731 2012     2700      1     2.817
## 147015 84863524855 5.149 2012     2700      1     3.930
## 147018 84863533887 4.258 2012     2700      1     3.039
## 147075 84863550146 3.806 2012     2700      1     2.587
## 147080 84863554398 3.686 2012     2700      1     2.772
## 147811 84862859820 6.239 2012     2700      1     5.020
## 147814 84862903106 5.924 2012     2700      1     4.705
## 147913 84862511733 3.543 2012     2700      1     2.629
## 147933 84862859054 3.790 2012     2700      1     2.571
## 148038 84862226987 4.247 2012     2700      1     3.333
## 148062 84862489245 3.731 2012     2700      1     2.717
## 148141 84862118837 3.961 2012     2700      1     2.842
## 148146 84862260833 4.332 2012     2700      1     3.113
## 148915 84861689703 3.695 2012     2700      1     2.576
## 149011 84861394764 4.694 2012     2700      1     3.575
## 149012 84861409254 3.735 2012     2700      1     2.616
## 149013 84861409551 4.560 2012     2700      1     3.546
## 149126 84861018546 4.042 2012     2700      1     3.128
```

```
## 149127 84861022041 4.793 2012     2700      1     3.879
## 149136 84862777742 3.759 2012     1000      2     2.540
## 149215 84860914047 3.560 2012     2700      1     2.546
## 149238 84860741191 4.704 2012     2700      1     3.585
## 149259 84860776122 4.355 2012     2700      1     3.236
## 149264 84860910413 3.658 2012     2700      1     2.539
## 149362 84860718270 4.087 2012     2700      1     2.968
## 149363 84860720391 4.234 2012     2700      1     3.115
## 150121 84860344255 3.718 2012     2700      1     2.599
## 150154 84860116634 3.813 2012     2700      1     2.694
## 150155 84860136615 5.159 2012     2700      1     4.040
## 150175 84860200127 4.100 2012     2700      1     2.981
## 150177 84860358233 3.618 2012     2700      1     2.604
## 150256 84859819731 4.273 2012     2700      1     3.154
## 150277 84859834632 4.025 2012     2700      1     2.906
## 150279 84859846433 4.100 2012     2700      1     2.981
## 150384 84859555074 4.313 2012     2700      1     3.194
## 150445 84860383862 4.156 2012     2700      1     3.037
## 150503 84858665432 4.608 2012     2700      1     3.489
## 150508 84859544287 4.701 2012     2700      1     3.582
## 150550 84859092587 3.942 2012     2700      1     2.823
## 150563 84859718265 4.683 2012     2700      1     3.564
## 151282 84859406106 3.740 2012     2700      1     2.621
## 151344 84858638369 4.465 2012     2700      1     3.346
## 151347 84863337617 4.454 2012     2700      1     3.335
## 151349 84858796262 3.993 2012     1000      2     2.874
## 151408 84858331964 4.546 2012     1000      2     3.427
## 151434 84858265970 4.576 2012     2700      1     3.457
## 151459 84858329412 3.813 2012     2700      1     2.694
## 151554 84863229939 4.007 2012     2700      1     3.093
## 151641 84858019974 4.876 2012     2700      1     3.757
## 151895 84863267488 3.633 2012     2700      1     2.514
## 152472 84857848576 3.905 2012     2700      1     2.686
## 152564 84857427752 4.330 2012     2700      1     3.211
## 152567 84863115197 3.667 2012     2700      1     2.548
## 152654 84857132739 4.189 2012     2700      1     3.070
## 152684 84857065859 4.573 2012     2700      1     3.354
## 152856 84856433589 5.714 2012     2700      1     4.700
## 152866 84856487711 5.900 2012     2700      1     4.886
## 152870 84856552278 3.704 2012     2700      1     2.585
## 152908 84857643783 4.842 2012     2700      1     3.723
## 153052 84863012865 4.087 2012     2700      1     3.073
## 153390 84856245226 4.422 2012     2700      1     3.303
## 153392 84856249964 3.802 2012     2700      1     2.683
## 153393 84856253589 4.184 2012     2700      1     3.065
## 153404 84856133161 3.633 2012     2700      1     2.619
## 153473 84856159009 4.542 2012     2700      1     3.323
## 153475 84862908900 3.581 2012     2700      1     2.567
## 153490 84855993116 3.748 2012     2700      1     2.734
## 153532 84855843640 3.790 2012     2700      1     2.876
```

```
## 153533 84855854046 4.736 2012     2700      1     3.617
## 153548 84856176782 4.001 2012     2700      1     3.087
## 153549 84862909100 4.062 2012     2700      1     2.943
## 153663 84855426656 3.662 2012     2700      1     2.543
## 153672 84855459760 4.667 2012     2700      1     3.548
## 153732 84055199809 4.449 2012     2700      1     3.330
## 153835 84856628070 3.840 2012     2700      1     2.721
## 153838 84856657579 4.089 2012     2700      1     2.870
## 154252 84863718620 3.870 2012     2700      1     2.956
## 154329 84864861863 4.382 2012     2700      1     3.263
## 154561 84870478439 4.172 2012     2700      1     3.053
## 154562 84870494510 4.054 2012     2700      1     2.935
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -2.65059 -1.02291  0.00443  0.71396  5.29667
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.89937    0.03106   61.16  < 2e-16 ***
## FirstAuthorFemale1  0.00655    0.00757    0.86     0.39
## Year1997           -0.00691    0.04244   -0.16     0.87
## Year1998            0.03628    0.04236    0.86     0.39
## Year1999           -0.02201    0.04213   -0.52     0.60
## Year2000            0.74467    0.09452    7.88  3.3e-15 ***
## Year2001           -0.06970    0.04279   -1.63     0.10
## Year2002           -0.54910    0.03666  -14.98  < 2e-16 ***
## Year2003           -0.61226    0.03769  -16.25  < 2e-16 ***
## Year2004           -0.69562    0.03571  -19.48  < 2e-16 ***
## Year2005           -0.72934    0.03543  -20.58  < 2e-16 ***
## Year2006           -0.75859    0.03403  -22.29  < 2e-16 ***
## Year2007           -0.87470    0.03336  -26.22  < 2e-16 ***
## Year2008           -0.87646    0.03286  -26.67  < 2e-16 ***
## Year2009           -0.83115    0.03272  -25.40  < 2e-16 ***
## Year2010           -0.78254    0.03218  -24.31  < 2e-16 ***
## Year2011           -0.78540    0.03298  -23.82  < 2e-16 ***
## Year2012           -0.80888    0.03238  -24.98  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 1.08
## Multiple R-squared:  0.0694, Adjusted R-squared:  0.0692
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
```

```
##  9 observations c(14272,15323,22486,33665,36647,38880,39540,73247,82361)
##   are outliers with |weight| = 0 ( < 1.2e-06);
##  6832 weights are ~= 1. The remaining 79817 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0004  0.8990  0.9380  0.9070  0.9830  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.15e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##                   GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.002  1           1.001
## Year             1.002 16           1.000
```



```
## [1] "List of  2719 outliers with residuals above  2.5"
##           ScopusId  NLCS Year OneField Fields residuals
## 28      0029855881 4.801 1996     2700      1     2.881
## 30      0029906954 4.525 1996     2700      1     2.605
## 31      0141469033 5.783 1996     2700      1     3.863
## 47      0030453430 5.115 1996     2700      1     3.195
## 79     10544241545 4.738 1996     2700      1     2.719
## 87     10544235697 5.572 1996     2700      1     3.652
## 95      0029911547 5.746 1996     2700      1     4.031
## 113     0030445619 5.158 1996     2700      1     3.139
## 153     0029683640 4.469 1996     2700      1     2.549
## 179     0029844930 4.941 1996     2700      1     3.021
## 197     0029979690 4.533 1996     2700      1     2.514
## 1288    0029827650 4.623 1996     2700      1     2.703
## 1305    0030599533 4.661 1996     2700      1     2.642
## 1309    0029828983 4.411 1996     2700      1     2.596
## 1355    0030590745 4.426 1996     2700      1     2.506
## 1359    0029808003 5.038 1996     2700      1     3.019
## 1369    0029845013 4.799 1996     2700      1     2.879
## 1394    0029859887 4.512 1996     2700      1     2.797
## 1397    0029919459 5.204 1996     2700      1     3.185
## 1399    3042828081 4.528 1996     2700      1     2.509
## 1414   10544220023 4.705 1996     2700      1     2.785
## 1424    0029907556 5.146 1996     2700      1     3.331
## 1425    0029908290 4.742 1996     2700      1     2.822
## 1649    0029849462 4.874 1996     2700      1     2.954
## 1672    0030603334 4.697 1996     2700      1     2.678
## 1677    0029862114 4.329 1996     2700      1     2.614
## 1708    0029823520 4.701 1996     2700      1     2.781
## 1717   10144258656 4.587 1996     2700      1     2.667
## 1725    0029861097 4.833 1996     2700      1     3.118
## 1728    0029861579 4.669 1996     2700      1     2.650
## 1755    0030581586 5.042 1996     2700      1     3.327
## 1756    0030581587 4.964 1996     2700      1     2.945
## 1757    0030581590 4.755 1996     2700      1     2.736
## 1758    0242409954 4.538 1996     2700      1     2.618
## 1763    0005150171 4.797 1996     2700      1     2.877
## 1764    0007410805 5.440 1996     2700      1     3.520
## 1766   10144244674 5.218 1996     2700      1     3.298
## 1772    0029758899 4.682 1996     2700      1     2.867
## 1773    0029759549 4.361 1996     2700      1     2.646
## 1787    0010496839 5.103 1996     2700      1     3.183
## 1788    0029795224 5.088 1996     2700      1     3.069
## 1798    0029773862 4.819 1996     2700      1     2.899
## 1958   10144241022 5.066 1996     2700      1     3.047
## 1977    0029798819 4.729 1996     1300      2     2.809
## 1979    0029842830 5.287 1996     1300      2     3.367
## 2023    0030604909 4.869 1996     2700      1     2.850
## 2032    0029790755 4.602 1996     2700      1     2.583
## 2033    0029840653 5.036 1996     2700      1     3.116
```

```
## 2070    9544253883 4.485 1996     2700      1     2.770
## 2071    0029740950 5.202 1996     2700      1     3.282
## 2072    0029741921 5.309 1996     2700      1     3.389
## 2074    9544228424 5.880 1996     2700      1     3.960
## 2107    0029835263 5.098 1996     2700      1     3.079
## 2108    0029835392 5.259 1996     2700      1     3.339
## 2109    0029838499 4.691 1996     2700      1     2.672
## 2110    0242435873 5.375 1996     2700      1     3.660
## 2112    9544222721 4.784 1996     2700      1     3.069
## 2113    0029739543 4.452 1996     2700      1     2.532
## 2115    0029771618 4.842 1996     2700      1     3.027
## 2130    0029818344 4.305 1996     2700      1     2.590
## 2131    0029822091 4.840 1996     2700      1     2.920
## 2159    0029737324 4.674 1996     2700      1     2.754
## 2162    0029739970 5.663 1996     2700      1     3.743
## 2167    0029758146 4.466 1996     2700      1     2.546
## 2354    0030250166 4.259 1996     2700      1     2.544
## 2485    0029741822 5.190 1996     2700      1     3.171
## 2486    0029746491 4.535 1996     2700      1     2.615
## 2489    9444268678 5.105 1996     2700      1     3.185
## 2490    0029740582 4.658 1996     2700      1     2.738
## 2500    0030017720 4.990 1996     2700      1     3.070
## 2503    0029783235 4.667 1996     2700      1     2.747
## 2510    0029739053 4.836 1996     2700      1     2.916
## 2513    0029759025 5.617 1996     2700      1     3.598
## 2526    0029830177 4.802 1996     2700      1     2.987
## 2537    0029737310 4.288 1996     2700      1     2.573
## 2543   15844420661 5.335 1996     2700      1     3.415
## 2574    0029830551 4.585 1996     2700      1     2.665
## 2582    0029862164 4.517 1996     2700      1     2.702
## 2588    0029916051 4.499 1996     2700      1     2.684
## 2741    0030594830 5.438 1996     2700      1     3.518
## 2743    0030000230 4.826 1996     2700      1     2.906
## 2744    8944228913 5.408 1996     2700      1     3.389
## 2748    0030056083 5.607 1996     2700      1     3.687
## 2750    0030057106 5.671 1996     2700      1     3.751
## 2797    8944233864 5.465 1996     2700      1     3.545
## 2806    0029902170 4.900 1996     2700      1     3.085
## 2830    0030015522 4.848 1996     2700      1     2.928
## 3004    0029944290 4.689 1996     1300      2     2.769
## 3042    0029786164 4.533 1996     2700      1     2.613
## 3058    0029900294 5.422 1996     2700      1     3.502
## 3059    0029942003 5.780 1996     2700      1     3.860
## 3085    0030594604 4.947 1996     2700      1     3.027
## 3109    0029948212 5.514 1996     2700      1     3.594
## 3111    0029948967 4.647 1996     2700      1     2.727
## 3154    9344267138 4.432 1996     2700      1     2.717
## 3166   15844364006 4.499 1996     2700      1     2.579
## 3185    0029888972 4.889 1996     2700      1     3.174
## 3214    0030000088 4.916 1996     2700      1     2.996
```

```
## 3462    0029892519 4.994 1996     2700      1     3.074
## 3468    0029939573 4.570 1996     2700      1     2.650
## 3471    0029878679 4.693 1996     2700      1     2.773
## 3474    0029897909 5.220 1996     2700      1     3.201
## 3479    0029886630 5.409 1996     2700      1     3.489
## 3487    0029889718 4.606 1996     2700      1     2.686
## 3498    0029870954 4.842 1996     2700      1     2.922
## 3503    0029932125 4.845 1996     2700      1     2.826
## 3521    0029925567 5.575 1996     2700      1     3.655
## 3537    0029871858 4.847 1996     2700      1     2.828
## 3539    0030008301 4.452 1996     2700      1     2.532
## 3552    9244262406 6.270 1996     2700      1     4.350
## 3560    0029863684 5.262 1996     2700      1     3.342
## 3572    0029913381 4.769 1996     2700      1     2.750
## 3573    0029913643 4.471 1996     2700      1     2.551
## 3578    0029929617 5.256 1996     2700      1     3.541
## 3579    0029929618 4.964 1996     2700      1     3.249
## 3795    0029875770 5.139 1996     2700      1     3.219
## 3797    0029982081 4.757 1996     2700      1     2.837
## 3806    0029988316 4.987 1996     2700      1     3.067
## 3830    0029920903 4.358 1996     2700      1     2.643
## 3866    0029881352 4.449 1996     2700      1     2.529
## 3877    0029869817 4.507 1996     2700      1     2.792
## 3890    0029993565 5.060 1996     2700      1     3.140
## 3892   13344285352 4.993 1996     2700      1     2.974
## 4094    0029915834 4.443 1996     2700      1     2.523
## 4105    0029970342 5.515 1996     2700      1     3.595
## 4111    0029917496 5.333 1996     2700      1     3.314
## 4150    0029993531 4.882 1996     2700      1     2.962
## 4157    0030009316 4.530 1996     2700      1     2.610
## 4168    0029873844 4.915 1996     2700      1     2.896
## 4169    0029987128 4.757 1996     2700      1     2.738
## 4171    0029870311 4.393 1996     2700      1     2.678
## 4178    9044236527 5.250 1996     2700      1     3.535
## 4195    0030070405 4.558 1996     2700      1     2.743
## 4217    0029866721 4.724 1996     2700      1     2.705
## 4509    0942276272 5.608 1996     2700      1     3.688
## 4519   13344293705 4.423 1996     2700      1     2.608
## 4535    0030066467 5.150 1996     2700      1     3.230
## 4536    0030070793 4.738 1996     2700      1     3.023
## 4548    0030020590 4.565 1996     1300      2     2.546
## 4552    0030020630 4.999 1996     2700      1     3.079
## 4567    0030058515 5.136 1996     2700      1     3.216
## 4591    0030042790 5.170 1996     2700      1     3.250
## 4593   13344286314 4.575 1996     2700      1     2.655
## 4596    0030068231 4.602 1996     2700      1     2.682
## 4598    0030058666 4.263 1996     2700      1     2.548
## 4603    0030033367 4.824 1996     2700      1     2.904
## 4622    0030032444 5.228 1996     2700      1     3.308
## 4629    0030049026 4.570 1996     2700      1     2.650
```

```
## 4632    0030056373 4.773 1996     2700      1     2.853
## 4784    0030032258 4.399 1996     2700      1     2.684
## 4802    0030044553 4.676 1996     2700      1     2.961
## 4818    0030049823 4.429 1996     2700      1     2.614
## 4819   13344260688 4.602 1996     2700      1     2.682
## 4826    0030048691 4.811 1996     2700      1     3.096
## 4827    0030060343 4.437 1996     2700      1     2.517
## 4833    0030026954 4.778 1996     2700      1     2.963
## 4843    0030032378 4.408 1996     2700      1     2.693
## 4866    0030043663 4.884 1996     2700      1     2.964
## 4868    0030579588 4.838 1996     2700      1     2.918
## 4869    0030034463 5.009 1996     2700      1     2.990
## 4883    0030034592 5.638 1996     2700      1     3.718
## 4884    0030054309 6.200 1996     2700      1     4.280
## 4890    0030024585 4.558 1996     2700      1     2.638
## 4891    0030046074 4.501 1996     2700      1     2.786
## 4897    0000677401 5.404 1996     2700      1     3.589
## 4904    0001196648 5.048 1996     2700      1     3.128
## 4908    0001854685 4.563 1996     2700      1     2.748
## 4916    0004851872 6.351 1996     2700      1     4.636
## 5482    0029959975 4.583 1996     2700      1     2.564
## 5653    0030034465 5.226 1996     2700      1     3.306
## 5706    0030054203 4.894 1996     2700      1     2.974
## 5729    0030060710 4.744 1996     2700      1     2.725
## 5737    0030061707 4.711 1996     2700      1     2.791
## 5847    0030183205 4.634 1996     2700      1     2.819
## 6074    0030576183 4.548 1996     2700      1     2.628
## 6086    0040419081 4.738 1996     2700      1     3.023
## 6118   16144365807 4.546 1996     2700      1     2.527
## 6164    9244238642 4.854 1996     2700      1     2.934
## 7329    0031435838 4.794 1997     2700      1     3.086
## 7330    2642611953 4.758 1997     2700      1     2.951
## 7352    0031468254 4.620 1997     2700      1     2.708
## 7360    0031472414 5.042 1997     2700      1     3.130
## 7367    0031437624 4.545 1997     2700      1     2.633
## 7375    0031466866 4.796 1997     2700      1     2.884
## 7379    2642597076 5.193 1997     2700      1     3.281
## 7415    0030704598 4.531 1997     2700      1     2.619
## 7534    0031136510 4.780 1997     2700      1     2.868
## 7757    0031454125 4.470 1997     2700      1     2.558
## 7787    0031472452 4.431 1997     2700      1     2.723
## 8304    0030703240 4.771 1997     2700      1     2.759
## 8307    0030665744 4.564 1997     2700      1     2.652
## 8323    0031590631 4.483 1997     2700      1     2.571
## 8332    0030692782 5.405 1997     2700      1     3.493
## 8339    0030666228 4.564 1997     2700      1     2.757
## 8340    0030671045 4.349 1997     2700      1     2.641
## 8357    0030809817 4.318 1997     2700      1     2.610
## 8378    0030695138 5.236 1997     2700      1     3.324
## 8424    0030716498 5.322 1997     2700      1     3.410
```

```
## 8442    0030661687 5.440 1997     2700      1     3.528
## 8455    0030671503 4.307 1997     2700      1     2.599
## 8586    0030731486 5.178 1997     2700      1     3.371
## 8589    0030831628 4.379 1997     2700      1     2.671
## 8608    0030720886 5.370 1997     2700      1     3.662
## 8610    0030759072 5.217 1997     2700      1     3.305
## 8625    0030744945 5.294 1997     2700      1     3.382
## 8627    0030762081 4.480 1997     2700      1     2.568
## 8631    0030879325 4.935 1997     2700      1     3.023
## 8692    0030766770 4.941 1997     2700      1     3.233
## 8713    0030779037 5.224 1997     2700      1     3.212
## 8714    0030803131 4.352 1997     2700      1     2.644
## 8728    0030774499 4.679 1997     2700      1     2.872
## 8849    0030886175 4.489 1997     2700      1     2.577
## 8850    0030954873 5.207 1997     2700      1     3.499
## 8851    0030967953 6.065 1997     2700      1     4.153
## 8861    0006750594 4.483 1997     2700      1     2.571
## 8891    0030756517 4.679 1997     2700      1     2.767
## 8980    0030869269 5.624 1997     2700      1     3.712
## 8983    9844257578 4.711 1997     2700      1     3.003
## 9007    0030763532 5.703 1997     2700      1     3.691
## 9014    0030928107 4.550 1997     2700      1     2.842
## 9045    0030921893 4.869 1997     2700      1     2.957
## 9237    0030803718 4.703 1997     2700      1     2.896
## 9239    0030868293 5.188 1997     2700      1     3.276
## 9245    0030804947 4.885 1997     2700      1     3.177
## 9250    0030792120 4.836 1997     2700      1     2.924
## 9309    8544284052 4.624 1997     2700      1     2.712
## 9329    0030851442 4.840 1997     2700      1     2.928
## 9331    0030872063 4.662 1997     2700      1     2.750
## 9332    1842332651 4.618 1997     2700      1     2.706
## 9378    0030854951 4.445 1997     2700      1     2.533
## 9468    8544252402 4.498 1997     2700      1     2.790
## 9514    0030744599 4.744 1997     2700      1     2.832
## 9515    0030756101 4.907 1997     2700      1     3.100
## 9519    0030854780 4.755 1997     2700      1     2.843
## 9522    0030742587 4.609 1997     2700      1     2.901
## 9539    0030876805 5.124 1997     2700      1     3.212
## 9568    0030803395 4.639 1997     2700      1     2.931
## 9597    0030610461 5.331 1997     2700      1     3.319
## 9603    0038025895 4.877 1997     2700      1     2.965
## 9607    0030957310 5.247 1997     2700      1     3.335
## 9629    0030760729 4.653 1997     2700      1     2.741
## 9652    0030878260 4.429 1997     2700      1     2.517
## 9760    0031005933 5.176 1997     2700      1     3.264
## 9763    1842410169 4.665 1997     2700      1     2.653
## 9767    0030916401 5.117 1997     2700      1     3.205
## 9816    0030908055 5.171 1997     2700      1     3.259
## 9818    0030941817 4.703 1997     2700      1     2.691
## 9820    0030976173 4.495 1997     2700      1     2.787
```

```
## 9823    0031009275 4.344 1997     2700      1     2.537
## 9824    8244234470 4.344 1997     2700      1     2.636
## 9833    0343157354 5.144 1997     2700      1     3.232
## 9902    0030974188 4.972 1997     2700      1     3.060
## 10055   0030909527 4.552 1997     2700      1     2.745
## 10056   0030910022 5.719 1997     2700      1     4.011
## 10060   0030939253 4.627 1997     2700      1     2.715
## 10066   0031003334 5.309 1997     2700      1     3.397
## 10071   0343918505 4.869 1997     2700      1     3.161
## 10072   0346593922 4.853 1997     2700      1     2.941
## 10098   0031009871 4.652 1997     2700      1     2.845
## 10102   0030919511 5.119 1997     2700      1     3.207
## 10103   0030919667 4.467 1997     2700      1     2.555
## 10127   0030977450 4.957 1997     2700      1     3.045
## 10149   0030913316 4.821 1997     2700      1     2.909
## 10156   0030961921 5.444 1997     2700      1     3.736
## 10184   0030896520 5.060 1997     2700      1     3.048
## 10222   0030979720 4.965 1997     2700      1     3.053
## 10347   0030967165 5.834 1997     2700      1     3.822
## 10353   0030955080 4.722 1997     2700      1     3.014
## 10378   0012444519 6.220 1997     2700      1     4.308
## 10386   0030903440 4.583 1997     2700      1     2.875
## 10427   0030896996 4.738 1997     2700      1     3.030
## 10428   0030899940 5.040 1997     2700      1     3.028
## 10454   0030956673 6.413 1997     2700      1     4.501
## 10461   0030951198 5.067 1997     2700      1     3.155
## 10464   0031004940 4.520 1997     2700      1     2.812
## 10467   0030891988 4.695 1997     2700      1     2.683
## 10490   0030946366 4.576 1997     2700      1     2.564
## 10535   0031127072 5.869 1997     2700      1     4.062
## 10598   0030976067 5.207 1997     2700      1     3.195
## 10622   0030898397 4.415 1997     2700      1     2.503
## 10623   0030948376 4.687 1997     2700      1     2.775
## 10646   0030945186 4.445 1997     2700      1     2.533
## 10660   0031047918 4.572 1997     2700      1     2.560
## 10671   0031054674 5.776 1997     2700      1     3.864
## 10694   0031047361 4.710 1997     2700      1     2.798
## 10735   0030948337 4.434 1997     2700      1     2.627
## 10891   0013655172 4.847 1997     2700      1     2.935
## 10895   0031052369 4.993 1997     2700      1     3.081
## 10901   0031018269 4.695 1997     2700      1     2.783
## 10919   0031045652 5.274 1997     2700      1     3.362
## 10920   0031050740 4.653 1997     2700      1     2.741
## 10933   0031052862 5.331 1997     2700      1     3.319
## 10935   0031057614 4.810 1997     2700      1     2.798
## 10972   0031058448 4.616 1997     2700      1     2.809
## 10979   0031019745 4.552 1997     1000      2     2.640
## 10981   0031019923 4.600 1997     2700      1     2.688
## 10983   0031028712 4.450 1997     2700      1     2.538
## 10987   0031028804 5.501 1997     2700      1     3.589
```

```
## 10993   0030614495 4.450 1997     2700      1     2.538
## 11012   0030807358 4.886 1997     2700      1     2.874
## 11019   0030868077 4.859 1997     2700      1     2.947
## 11136   0031012726 5.498 1997     2700      1     3.586
## 11137   0031030450 4.891 1997     2700      1     2.979
## 11138   0031032055 5.515 1997     2700      1     3.603
## 11153   0031030620 5.167 1997     2700      1     3.255
## 11162   0031012532 4.615 1997     2700      1     2.703
## 11174   0031028365 4.604 1997     2700      1     2.592
## 11175   0031029379 4.969 1997     2700      1     2.957
## 11202   0031025101 4.552 1997     2700      1     2.844
## 11213   0031037065 4.605 1997     2700      1     2.593
## 11823   0030938817 4.429 1997     2700      1     2.721
## 11955   0031011193 4.851 1997     2700      1     2.939
## 11971   0031020373 4.539 1997     2700      1     2.527
## 11972   0031022490 4.618 1997     2700      1     2.606
## 13373   0031472453 4.648 1997     2700      1     2.736
## 13487   1842295778 5.282 1997     2700      1     3.370
## 13948   0032585514 4.968 1998     2700      1     3.012
## 13961   0032583492 5.604 1998     2700      1     3.852
## 13975   0032564658 4.473 1998     2700      1     2.517
## 13977   0032564703 4.515 1998     2700      1     2.763
## 13994   0032542385 4.640 1998     2700      1     2.684
## 14001   0032539164 4.551 1998     2700      1     2.700
## 14018   0032506614 5.304 1998     2700      1     3.348
## 14032   0032480930 4.637 1998     2700      1     2.681
## 14040   0032477294 5.316 1998     2700      1     3.564
## 14041   0032477305 5.311 1998     2700      1     3.559
## 14043   0032477331 5.005 1998     2700      1     3.154
## 14065   0031609707 4.571 1998     2700      1     2.615
## 14801   0032567115 5.271 1998     2700      1     3.315
## 14816   0032547938 6.002 1998     2700      1     4.250
## 14828   0032544946 4.306 1998     2700      1     2.554
## 14860   0032512097 4.594 1998     2700      1     2.638
## 14866   0032508952 6.390 1998     2700      1     4.434
## 14881   0032487923 5.141 1998     2700      1     3.185
## 14888   0032483685 5.248 1998     2700      1     3.292
## 14889   0032483690 4.752 1998     2700      1     2.796
## 15041   0032578789 4.449 1998     2700      1     2.697
## 15080   0032556180 5.427 1998     2700      1     3.675
## 15109   0032517251 4.717 1998     2700      1     2.965
## 15110   0032517258 4.889 1998     2700      1     3.137
## 15130   0032494396 4.981 1998     2700      1     2.925
## 15186   0032189897 5.522 1998     2700      1     3.566
## 15322   0032563824 4.832 1998     2700      1     2.876
## 15354   0032541616 5.053 1998     2700      1     3.097
## 15357   0032541672 4.883 1998     2700      1     2.927
## 15362   0032538047 4.487 1998     2700      1     2.531
## 15396   0032504985 4.662 1998     2700      1     2.706
## 15418   0032480346 4.888 1998     2700      1     2.932
```

```
## 15530   0032169485 5.282 1998     2700      1     3.326
## 15532   0032169557 4.561 1998     2700      1     2.605
## 15630   0032572929 4.615 1998     2700      1     2.659
## 15666   0032547326 6.369 1998     2700      1     4.413
## 15667   0032547328 4.489 1998     2700      1     2.533
## 15690   0032529103 5.002 1998     2700      1     3.046
## 15700   0008926519 5.466 1998     2700      1     3.510
## 15741   0032486734 4.466 1998     2700      1     2.714
## 15784   0032145836 4.610 1998     2700      1     2.554
## 15882   0032578110 4.511 1998     2700      1     2.660
## 15906  18744421762 4.597 1998     2700      1     2.641
## 15932   0032528169 4.517 1998     2700      1     2.666
## 15935   0032528180 5.239 1998     2700      1     3.283
## 15947   0032508297 5.453 1998     2700      1     3.497
## 15961   0032496880 4.515 1998     2700      1     2.559
## 15971   0032474694 4.931 1998     2700      1     2.975
## 16151   0032551162 4.753 1998     2700      1     2.797
## 16167   0032543663 5.069 1998     2700      1     3.113
## 16172   7144228604 4.850 1998     2700      1     2.894
## 16174   0009452157 4.306 1998     2700      1     2.554
## 16193   0032525901 4.517 1998     2700      1     2.765
## 16222   0032503637 4.575 1998     2700      1     2.823
## 16251   0032482329 5.835 1998     2700      1     4.083
## 16263   0032479034 4.911 1998     2700      1     2.855
## 16346   0032102292 5.131 1998     2700      1     3.175
## 16483   0032572043 5.088 1998     2700      1     3.336
## 16487   0032572086 6.274 1998     2700      1     4.318
## 16504   0032550626 5.681 1998     2700      1     3.725
## 16505   0032550630 5.027 1998     2700      1     3.275
## 16536   0032516267 4.935 1998     2700      1     3.183
## 16539   0032516296 5.565 1998     2700      1     3.609
## 16568   0032490147 4.494 1998     2700      1     2.538
## 16773   0032580484 4.324 1998     2700      1     2.572
## 16793   0032560037 4.393 1998     2700      1     2.542
## 16800   0032557174 4.358 1998     2700      1     2.507
## 16801   0032557175 4.491 1998     2700      1     2.535
## 16802   2642713359 4.651 1998     2700      1     2.695
## 16816   0032537068 4.638 1998     2700      1     2.682
## 16819   0032537191 4.969 1998     2700      1     3.118
## 16832   0032522736 4.715 1998     2700      1     2.659
## 16839   0032523018 4.548 1998     2700      1     2.592
## 16851   0032523214 4.715 1998     2700      1     2.659
## 16875   0032499126 4.411 1998     2700      1     2.659
## 16876   0032499141 4.532 1998     2700      1     2.576
## 16910   0008947293 4.653 1998     2700      1     2.697
## 16913   0032473922 4.668 1998     2700      1     2.712
## 17047   0032568257 4.810 1998     2700      1     2.854
## 17053   0032565112 5.107 1998     2700      1     3.151
## 17063   0032554552 4.707 1998     2700      1     2.955
## 17064   0032554569 4.610 1998     2700      1     2.554
```

```
## 17091   0032542709 5.010 1998     2700      1     3.054
## 17163   0032485350 5.150 1998     2700      1     3.194
## 17177   0032481673 4.592 1998     2700      1     2.536
## 17289   0032032575 4.574 1998     2700      1     2.618
## 17293   0032033112 4.403 1998     2700      1     2.651
## 17386   0032564902 5.256 1998     2700      1     3.200
## 17401   0032545939 4.480 1998     2700      1     2.524
## 17404   0032546036 4.281 1998     2700      1     2.529
## 17408   0032542516 4.484 1998     2700      1     2.528
## 17410   0032542541 4.339 1998     2700      1     2.587
## 17420   0032519431 4.715 1998     2700      1     2.659
## 17429   0032519925 5.694 1998     2700      1     3.638
## 17433   0032520153 4.772 1998     2700      1     2.816
## 17444   0032510076 5.618 1998     2700      1     3.562
## 17473   0032484953 4.572 1998     2700      1     2.516
## 17481   0032481152 4.867 1998     2700      1     2.911
## 17482   0032481323 4.638 1998     2700      1     2.682
## 17502   0031987322 4.598 1998     2700      1     2.642
## 17541   0032005791 4.628 1998     2700      1     2.672
## 17595   0032573850 4.468 1998     2700      1     2.512
## 17606   0032556992 4.604 1998     2700      1     2.648
## 17613   0032554096 4.586 1998     2700      1     2.630
## 17627   0032518022 6.088 1998     2700      1     4.132
## 17631   0032518290 5.115 1998     2700      1     3.159
## 17633   0032518420 4.457 1998     2700      1     2.501
## 17644   0032518878 4.711 1998     2700      1     2.755
## 17655   0032515386 4.679 1998     2700      1     2.623
## 17680   0032491837 4.969 1998     2700      1     3.118
## 17686   0032477553 4.484 1998     2700      1     2.528
## 17696   0009440989 4.352 1998     2700      1     2.600
## 18344   0031945285 5.058 1998     2700      1     3.102
## 18436   0031975335 5.254 1998     2700      1     3.403
## 18437   0031975401 4.758 1998     2700      1     2.802
## 20189   0033619959 5.403 1999     2700      1     3.607
## 20216   0033582097 4.593 1999     2700      1     2.692
## 20242   0033572972 5.596 1999     2700      1     3.695
## 20262   0033540004 5.090 1999     2700      1     3.189
## 20278   0033518193 4.549 1999     2700      1     2.548
## 20279   0033518231 5.873 1999     2700      1     3.972
## 20319   0032803699 4.811 1999     2700      1     2.910
## 20484   0033485508 4.616 1999     2700      1     2.715
## 20874   0032748385 5.646 1999     2700      1     3.645
## 20925   0033544340 6.198 1999     2700      1     4.502
## 20962   0033520771 4.299 1999     2700      1     2.503
## 20967   0032589826 4.410 1999     2700      1     2.509
## 21115   0032694263 5.196 1999     2700      1     3.295
## 21120   0032695482 5.349 1999     2700      1     3.448
## 21123   0032741518 4.596 1999     2700      1     2.595
## 21124   0032742418 5.935 1999     2700      1     3.934
## 21125   0032742419 4.335 1999     2700      1     2.639
```

```
## 21126   0032743289 5.343 1999     2700      1     3.442
## 21127   0032747280 5.109 1999     2700      1     3.208
## 21134   0033598598 5.651 1999     2700      1     3.855
## 21152   0032696770 5.142 1999     2700      1     3.241
## 21155   0032740874 4.729 1999     2700      1     3.033
## 21255   0033517494 4.412 1999     2700      1     2.616
## 21273   0032855570 4.367 1999     2700      1     2.671
## 21436   0033615307 4.467 1999     2700      1     2.566
## 21443   0033575997 4.748 1999     2700      1     2.847
## 21498   0033536463 4.644 1999     2700      1     2.743
## 21500   0033536483 4.408 1999     2700      1     2.712
## 21501   0033536501 4.406 1999     2700      1     2.710
## 21548   0032841392 4.784 1999     2700      1     2.883
## 21656   0033199833 4.945 1999     2700      1     3.044
## 21763   0033603792 4.795 1999     2700      1     2.999
## 21775   0342961310 4.661 1999     2700      1     2.965
## 21785   0033584440 5.727 1999     2700      1     3.726
## 21826   0033546647 5.256 1999     2700      1     3.255
## 21990   0033595079 5.033 1999     2700      1     3.032
## 21993   0033595120 4.514 1999     2700      1     2.613
## 22017   0033565321 4.419 1999     2700      1     2.518
## 22021   0033565955 5.009 1999     2700      1     3.008
## 22042   0033542870 5.579 1999     2700      1     3.578
## 22168   0033168952 4.690 1999     2700      1     2.894
## 22237   0033600275 4.833 1999     2700      1     3.137
## 22248   0033031252 4.771 1999     2700      1     2.870
## 22286   0033542393 4.685 1999     2700      1     2.989
## 22303   0033538266 4.432 1999     2700      1     2.531
## 22305   0033538314 4.793 1999     2700      1     2.892
## 22486   0033609374 4.633 1999     2700      1     2.732
## 22491   0033606238 4.793 1999     2700      1     2.892
## 22521   0033562593 4.474 1999     2700      1     2.573
## 22535   0033551370 4.720 1999     2700      1     2.819
## 22563   0033526309 5.723 1999     2700      1     3.822
## 22759   0033611975 4.434 1999     2700      1     2.638
## 22779   0033594375 4.872 1999     2700      1     2.971
## 22793   0033591010 4.634 1999     2700      1     2.733
## 22832   0033553188 4.783 1999     2700      1     3.087
## 22834   0033553194 4.617 1999     2700      1     2.716
## 22836   0033553207 4.636 1999     2700      1     2.735
## 22852   0033531690 4.688 1999     2700      1     2.787
## 22855   0033515827 5.255 1999     1000      2     3.354
## 22909   0033119011 4.749 1999     2700      1     2.953
## 23018   0033577290 4.397 1999     2700      1     2.701
## 23019   0033577336 4.245 1999     2700      1     2.549
## 23056   0033545541 5.796 1999     2700      1     3.895
## 23057   0033545542 4.631 1999     2700      1     2.730
## 23064   0033540993 4.594 1999     2700      1     2.693
## 23067   0033541047 4.233 1999     2700      1     2.537
## 23076   0033522146 5.479 1999     2700      1     3.478
```

```
## 23077   0033522206 5.337 1999     2700      1     3.436
## 23111   0033037922 4.231 1999     2700      1     2.535
## 23167   0033104847 4.549 1999     2700      1     2.548
## 23271   0033602049 5.262 1999     2700      1     3.361
## 23292   0033580206 4.603 1999     2700      1     2.807
## 23293   0033580211 4.651 1999     2700      1     2.750
## 23304   0033577041 4.749 1999     2700      1     2.748
## 23314   0033557202 4.498 1999     2700      1     2.597
## 23357   0033540644 4.402 1999     2700      1     2.501
## 23358   0033540646 5.996 1999     2700      1     4.095
## 23359   0033540680 4.688 1999     2700      1     2.687
## 23360   0033540714 4.973 1999     2700      1     3.177
## 23361   0033540719 4.883 1999     2700      1     2.882
## 23541   0033608182 4.940 1999     2700      1     2.939
## 23544   0033608191 4.549 1999     2700      1     2.648
## 23560   0033590516 4.464 1999     2700      1     2.563
## 23561   0033590525 4.644 1999     2700      1     2.743
## 23567   0033585499 4.878 1999     2700      1     2.977
## 23577   0032918414 4.830 1999     2700      1     3.134
## 23584   0032954778 5.005 1999     2700      1     3.104
## 23591   0033552854 4.631 1999     2700      1     2.935
## 23605   0033537343 4.731 1999     2700      1     2.830
## 23621   0033528072 4.427 1999     2700      1     2.526
## 23623   0033528101 4.622 1999     2700      1     2.721
## 23860   0032701521 4.971 1999     2700      1     3.070
## 23861   0032701642 4.548 1999     2700      1     2.647
## 23882   0032713075 5.527 1999     2700      1     3.626
## 23914   0032722149 4.823 1999     2700      1     3.027
## 23925   0032725185 4.746 1999     2700      1     2.845
## 23974   0032742992 5.265 1999     2700      1     3.364
## 23997   0032746174 4.892 1999     2700      1     2.991
## 24023   0032752579 4.462 1999     2700      1     2.561
## 24083   0032824349 4.787 1999     2700      1     2.886
## 24251   0032919402 4.464 1999     2700      1     2.563
## 24424   0033013684 4.775 1999     2700      1     2.774
## 26283   0034055392 0.000 2000     2700      1    -2.554
## 26314  24044533275 0.000 2000     1306      3    -2.759
## 26347  85018888882 0.000 2000     2700      3    -2.554
## 26354   0034735827 5.241 2000     2700      1     2.787
## 26368   0034626988 5.215 2000     2700      1     2.556
## 26410   0034305633 0.000 2000     2700      2    -2.659
## 26509  33746530366 0.000 2000     2700      1    -2.659
## 26516   0033853362 0.000 2000     2700      1    -2.659
## 26526   0034226622 0.000 2000     1300      2    -2.659
## 26692   0034048842 0.000 2000     2700      1    -2.659
## 26889  85008036014 0.000 2000     1704      3    -2.659
## 26890  85048650487 0.000 2000     2700      1    -2.659
## 26910  12944252960 0.000 2000     2700      1    -2.659
## 26948  85047454626 0.000 2000     2700      1    -2.659
## 26964   0034732201 5.487 2000     2700      1     2.828
```

```
## 26988   0034655836 0.000 2000     2700      1    -2.659
## 27018   0034102286 0.000 2000     2700      1    -2.659
## 27031   0034682247 5.466 2000     2700      1     2.707
## 27110   0035956483 5.056 2001     2700      1     3.100
## 27142   0035924765 5.181 2001     2700      1     3.225
## 27205   0035857147 4.489 2001     2700      1     2.633
## 27210   0035852034 5.010 2001     2700      1     3.054
## 27211   0035852037 5.008 2001     2700      1     3.152
## 28065   0035944839 4.480 2001     2700      1     2.828
## 28077   0035936007 4.392 2001     2700      1     2.536
## 28092   0035930101 4.593 2001     2700      1     2.737
## 28118   0035904776 4.828 2001     2700      1     2.972
## 28120   0035892018 5.798 2001     2700      1     3.942
## 28145   0035861037 4.689 2001     2700      1     2.733
## 28177   0035829840 4.821 2001     2700      1     3.169
## 28179   0035829842 6.472 2001     2700      1     4.616
## 28353   0035522330 5.300 2001     2700      1     3.444
## 28468   0035950685 4.432 2001     2700      1     2.576
## 28564   0035846315 4.580 2001     2700      1     2.724
## 28628   0035802216 4.544 2001     2700      1     2.588
## 28791   0035960116 4.693 2001     2700      1     2.837
## 28832   0035913589 5.497 2001     2700      1     3.641
## 28851   0035856017 5.495 2001     2700      1     3.639
## 28900   0035818048 5.736 2001     2700      1     3.880
## 29126   0035975408 4.632 2001     2700      1     2.776
## 29159   0035940039 4.392 2001     2700      1     2.536
## 29160   0035940040 4.567 2001     2700      1     2.711
## 29182   0034898468 4.742 2001     2700      1     2.786
## 29248   0035827865 4.721 2001     2700      1     2.865
## 29267   0034926528 4.251 2001     2700      1     2.599
## 29299   0035413919 4.889 2001     2700      1     3.237
## 29405   0035963919 4.481 2001     2700      1     2.625
## 29417   0035954660 4.486 2001     2700      1     2.630
## 29418   0035954670 4.624 2001     2700      1     2.768
## 29428   0035948630 4.447 2001     2700      1     2.591
## 29439   0034939686 4.476 2001     2700      1     2.620
## 29458   0035928419 4.926 2001     2700      1     3.274
## 29472   0035913228 4.532 2001     2700      1     2.676
## 29523   0035849988 4.543 2001     2700      1     2.687
## 29524   0035850010 5.106 2001     2700      1     3.150
## 29529   0034933989 4.909 2001     2700      1     3.158
## 29532   0034945516 5.330 2001     2700      1     3.474
## 29705   0035963494 4.396 2001     2700      1     2.540
## 29708   0035963529 5.289 2001     2700      1     3.433
## 29762   0035919183 4.463 2001     2700      1     2.507
## 29763   0035919184 4.459 2001     2700      1     2.503
## 29766   0035897882 4.216 2001     2700      1     2.564
## 29767   0035897888 4.789 2001     2700      1     2.933
## 29776   0035859201 4.805 2001     2700      1     3.054
## 29778   0035859233 4.976 2001     2700      1     3.324
```

```
## 29787   0035854054 5.981 2001     2700      1     4.329
## 29826   0035816007 5.037 2001     2700      1     3.181
## 29832   0035816032 5.067 2001     2700      1     3.211
## 30038   0035902194 4.674 2001     2700      1     2.818
## 30046   0035897686 4.612 2001     2700      1     2.756
## 30047   0035897688 5.037 2001     2700      1     3.181
## 30084   0035832261 6.117 2001     2700      1     4.261
## 30089   0035832341 4.478 2001     2700      1     2.826
## 30856   0035953646 4.483 2001     2700      1     2.831
## 30872   0035946664 5.610 2001     2700      1     3.654
## 30947   0035857967 4.501 2001     2700      1     2.645
## 30954   0035849144 4.588 2001     2700      1     2.936
## 30967   0035843629 4.350 2001     2700      1     2.698
## 31007   0035804852 4.734 2001     2700      1     2.878
## 31008   0035804857 4.620 2001     2700      1     2.869
## 31151   0035962077 4.916 2001     2700      1     3.264
## 31177   0035925774 4.765 2001     2700      1     3.014
## 31204   0035869435 4.410 2001     2700      1     2.758
## 31220   0035857786 4.465 2001     2700      1     2.813
## 31237   0035835981 4.789 2001     2700      1     2.933
## 31257   0035826089 5.495 2001     2700      1     3.639
## 31264   0035819910 4.759 2001     2700      1     2.803
## 31434   0035961566 5.156 2001     2700      1     3.300
## 31461   0035931959 4.840 2001     2700      1     2.884
## 31463   0035931973 4.682 2001     2700      1     2.826
## 31477   0035925169 4.717 2001     2700      1     2.761
## 31485   0035901051 5.395 2001     2700      1     3.539
## 31491   0035864943 4.678 2001     2700      1     2.822
## 31492   0035864977 4.470 2001     2700      1     2.614
## 31541   0035825726 5.310 2001     2700      1     3.454
## 31569   0035819481 4.496 2001     2700      1     2.540
## 31715   0035945667 4.738 2001     2700      1     2.882
## 31734   0035915688 4.647 2001     2700      1     2.691
## 31742   0035905889 4.581 2001     2700      1     2.725
## 31744   0035905908 4.896 2001     2700      1     3.040
## 31773   0035843243 5.197 2001     2700      1     3.446
## 31775   0035843254 4.324 2001     2700      1     2.672
## 31782   0035835027 4.362 2001     2700      1     2.611
## 31783   0035835029 4.403 2001     2700      1     2.547
## 31795   0035814656 4.673 2001     2700      1     2.817
## 31799   0035804277 5.895 2001     2700      1     4.039
## 31801   0035804283 4.647 2001     2700      1     2.791
## 31933   0034795654 4.634 2001     2700      1     2.778
## 31937   0034799675 4.729 2001     2700      1     2.873
## 32022   0034857255 4.388 2001     2700      1     2.532
## 32059   0034903710 4.656 2001     2700      1     2.800
## 32248   0035013499 4.468 2001     2700      1     2.512
## 32609   0035174173 4.507 2001     2700      1     2.651
## 33892   0035960866 5.082 2001     2700      1     3.126
## 34423   0037132575 4.351 2002     2700      1     2.973
```

```
## 34424   0037132651 3.900 2002     2700      1     2.627
## 34438   0037079352 4.660 2002     2700      1     3.183
## 34445   0037069763 4.294 2002     2700      1     2.916
## 34446   0037069783 4.449 2002     2700      1     3.276
## 34454   0037065306 5.195 2002     2700      1     3.817
## 34457   0037065340 4.209 2002     2700      1     2.831
## 34467   0037028001 4.145 2002     2700      1     2.767
## 34474   0037021538 3.880 2002     2700      1     2.502
## 34477   0037021550 3.922 2002     2700      1     2.544
## 34553   0036908576 3.908 2002     2700      1     2.530
## 34596   0036950665 4.266 2002     2700      1     2.888
## 34958   0037202854 4.238 2002     2700      1     2.860
## 34973   0037191726 4.195 2002     2700      1     2.817
## 34977   0037184416 4.294 2002     2700      1     2.916
## 34979   0037184418 4.896 2002     2700      1     3.518
## 35001   0037153024 4.533 2002     2700      1     3.155
## 35004   0037153037 4.517 2002     2700      1     3.139
## 35005   0037153042 5.145 2002     2700      1     3.872
## 35013   0037145856 5.083 2002     2700      1     3.705
## 35014   0037145863 5.578 2002     2700      1     4.200
## 35032   0037078969 5.588 2002     2700      1     4.415
## 35039   0037073256 4.972 2002     2700      1     3.699
## 35043   0037073289 4.418 2002     2700      1     2.941
## 35079   0037032401 4.648 2002     2700      1     3.270
## 35130   0036840210 3.984 2002     2700      1     2.606
## 35358   0037206364 5.141 2002     2700      1     3.763
## 35359   0037206368 4.285 2002     2700      1     2.907
## 35370   0037180233 4.412 2002     2700      1     3.034
## 35377   0037167982 4.361 2002     2700      1     2.983
## 35385   0037164045 4.478 2002     2700      1     3.001
## 35386   0037164054 5.560 2002     2700      1     4.083
## 35397   0037126201 4.294 2002     2700      1     2.916
## 35398   0037126332 4.518 2002     2700      1     3.041
## 35400   0037126341 5.773 2002     2700      1     4.395
## 35408   0037120864 4.463 2002     2700      1     3.085
## 35411   0037120907 4.813 2002     2700      1     3.640
## 35446   0037048640 4.161 2002     2700      1     2.783
## 35448   0037048669 6.012 2002     2700      1     4.535
## 35451   0037048695 5.664 2002     2700      1     4.187
## 35452   0037048697 4.170 2002     2700      1     2.897
## 35457   0037027050 4.327 2002     2700      1     3.154
## 35462   0037015543 4.380 2002     2700      1     3.207
## 35481   0036771775 4.268 2002     2700      1     2.791
## 35699   0037179578 4.011 2002     2700      1     2.633
## 35709   0037174346 4.951 2002     2700      1     3.573
## 35730   0037130645 3.986 2002     2700      1     2.608
## 35758   0037063394 3.715 2002     2700      1     2.542
## 35759   0037063403 4.739 2002     2700      1     3.466
## 35760   0037063415 4.234 2002     2700      1     2.856
## 35778  18644385210 4.777 2002     2700      1     3.300
```

```
## 35867   0036733312 4.085 2002     2700      1     2.707
## 35911   0036745457 3.816 2002     2700      1     2.643
## 36123   0037190086 4.360 2002     2700      1     2.982
## 36181   0037158614 4.762 2002     2700      1     3.384
## 36244   0037125569 4.079 2002     2700      1     2.602
## 36254   0037103420 4.673 2002     2700      1     3.295
## 36269   0037077489 4.481 2002     2700      1     3.004
## 36270   0037077490 4.958 2002     2700      1     3.785
## 36292   0037043652 4.862 2002     2700      1     3.385
## 36295   0037043682 4.263 2002     2700      1     2.990
## 36306   0037036787 4.703 2002     2700      1     3.325
## 36318   0037014941 4.189 2002     2700      1     2.712
## 36365   0036681988 5.232 2002     2700      1     3.854
## 36367   0036682163 4.454 2002     2700      1     3.076
## 36369   0036682267 4.857 2002     2700      1     3.479
## 36524   0037178579 4.374 2002     2700      1     2.897
## 36554   0037166985 4.791 2002     2700      1     3.413
## 36555   0037166986 4.893 2002     2700      1     3.515
## 36583   0037130113 4.121 2002     2700      1     2.848
## 36586   0037130183 4.068 2002     2700      1     2.690
## 36593   0037125379 6.631 2002     2700      1     5.458
## 36594   0037125417 3.979 2002     2700      1     2.601
## 36597   0037125454 4.380 2002     2700      1     3.207
## 36619   0037055011 4.224 2002     2700      1     2.846
## 36620   0037055016 4.251 2002     2700      1     2.873
## 36658   0037019322 4.383 2002     2700      1     3.005
## 36666   0037014584 5.135 2002     2700      1     3.862
## 36667   0037014630 4.589 2002     2700      1     3.211
## 36668   0037014644 4.131 2002     2700      1     2.753
## 36692   0035983642 4.125 2002     2700      1     2.747
## 36821   0037182768 4.258 2002     2700      1     2.985
## 36860   0037134824 4.216 2002     2700      1     2.739
## 36864   0037134872 4.163 2002     2700      1     2.686
## 36865   0037134927 3.726 2002     2700      1     2.553
## 36867   0037134933 3.807 2002     2700      1     2.534
## 36901   0037071793 4.430 2002     2700      1     3.157
## 36904   0037071808 5.799 2002     2700      1     4.421
## 36910   0037067128 4.048 2002     2700      1     2.670
## 36912   0037067150 4.184 2002     2700      1     2.806
## 36940   0037030694 3.959 2002     2700      1     2.581
## 36956   0037024313 4.125 2002     2700      1     2.952
## 36996   0036595171 4.043 2002     2700      1     2.566
## 37144   0037198420 4.789 2002     2700      1     3.411
## 37146   0037198432 5.119 2002     2700      1     3.946
## 37175   0036251741 4.327 2002     2700      1     3.154
## 37181   0037162115 4.877 2002     2700      1     3.499
## 37182   0037162116 5.260 2002     2700      1     3.882
## 37188   0037157095 3.984 2002     2700      1     2.606
## 37191   0037157169 3.999 2002     2700      1     2.621
## 37208   0037129633 4.406 2002     2700      1     3.028
```

```
## 37241   0037092912 4.464 2002     2700      1     3.086
## 37243   0037092914 4.102 2002     2700      1     2.724
## 37246   0037093012 4.590 2002     2700      1     3.417
## 37286   0037042231 4.892 2002     2700      1     3.514
## 37288   0037042250 3.905 2002     2700      1     2.632
## 37305   0037018763 3.994 2002     2700      1     2.616
## 37309   0037007679 4.195 2002     2700      1     2.817
## 37311   0037007683 4.588 2002     2700      1     3.111
## 37362   0036570099 4.672 2002     2700      1     3.195
## 37452   0037181697 3.871 2002     2700      1     2.698
## 37453   0037181727 4.139 2002     2700      1     2.662
## 37462   0037171703 4.324 2002     2700      1     2.946
## 37475   0037165667 4.466 2002     2700      1     2.989
## 37476   0037165703 4.734 2002     2700      1     3.356
## 37477   0037165713 4.125 2002     2700      1     2.747
## 37478   0037165728 4.888 2002     2700      1     3.510
## 37479   0037165740 5.429 2002     2700      1     4.156
## 37527   0037123168 4.123 2002     2700      1     2.646
## 37528   0037123206 4.022 2002     2700      1     2.849
## 37570   0037061912 5.703 2002     2700      1     4.325
## 37581   0037051930 4.197 2002     2700      1     2.819
## 37582   0037051975 4.690 2002     2700      1     3.517
## 37626   0037018496 4.615 2002     2700      1     3.342
## 37638   0037012465 4.786 2002     2700      1     3.408
## 37640   0037012476 4.433 2002     2700      1     3.055
## 37744   0037197039 3.880 2002     2700      1     2.502
## 37745   0037197044 4.518 2002     2700      1     3.140
## 37752   0037187907 5.013 2002     2700      1     3.635
## 37753   0037187926 4.094 2002     2700      1     2.716
## 37758   0037181496 4.070 2002     2700      1     2.692
## 37759   0037181511 4.841 2002     2700      1     3.463
## 37760   0037181515 4.728 2002     2700      1     3.350
## 37778   0036119735 4.484 2002     2700      1     3.106
## 37782   0036122659 4.001 2002     2700      1     2.524
## 37790   0037160936 4.064 2002     2700      1     2.686
## 37819   0037149714 4.285 2002     2700      1     2.907
## 37826   0037139438 4.401 2002     2700      1     3.023
## 37868   0037076025 5.024 2002     2700      1     3.751
## 37875   0037070525 4.224 2002     2700      1     2.846
## 37876   0037070528 4.231 2002     2700      1     3.058
## 37900   0037035121 4.647 2002     2700      1     3.374
## 37913   0037029069 3.902 2002     2700      1     2.524
## 38083   0037181148 4.810 2002     2700      1     3.432
## 38109   0037148928 4.712 2002     2700      1     3.235
## 38143   0037116832 5.558 2002     2700      1     4.180
## 38145   0037116836 4.256 2002     2700      1     2.779
## 38161   0037075257 5.488 2002     2700      1     4.011
## 38163   0037075268 3.927 2002     2700      1     2.654
## 38164   0037075269 3.951 2002     2700      1     2.573
## 38165   0037075271 4.109 2002     2700      1     2.632
```

```
## 38172   0037070259 4.711 2002     2700      1     3.333
## 38192   0036155211 4.109 2002     2700      1     2.936
## 38200   0037045561 4.075 2002     2700      1     2.697
## 38208   0037034257 6.602 2002     2700      1     5.224
## 38217   0037028761 4.268 2002     2700      1     2.995
## 38218   0037028766 3.946 2002     2700      1     2.773
## 38220   0037028769 4.250 2002     2700      1     2.872
## 38247   0037006397 4.366 2002     2700      1     2.988
## 38328   0037204196 4.718 2002     2700      1     3.340
## 38357   0037160567 3.981 2002     2700      1     2.603
## 38367   0037132805 3.880 2002     2700      1     2.707
## 38375   0037122950 4.667 2002     2700      1     3.394
## 38377   0037122987 3.981 2002     2700      1     2.603
## 38381   0037116641 6.031 2002     2700      1     4.653
## 38382   0037116642 5.093 2002     2700      1     3.715
## 38383   0037116647 4.485 2002     2700      1     3.212
## 38384   0037116658 3.807 2002     2700      1     2.634
## 38390   0037065519 4.048 2002     2700      1     2.875
## 38398   0037050352 5.876 2002     2700      1     4.399
## 38407   0037045439 4.562 2002     2700      1     3.184
## 38421   0037012134 4.436 2002     2700      1     3.058
## 38430   0037006110 4.025 2002     2700      1     2.548
## 38432   0037006117 3.976 2002     2700      1     2.598
## 38529   0036083829 4.036 2002     2700      1     2.658
## 38535   0036086409 4.229 2002     2700      1     2.752
## 38536   0036086413 3.999 2002     2700      1     2.726
## 38553   0036106054 4.887 2002     2700      1     3.714
## 38572   0036123845 3.948 2002     2700      1     2.570
## 38606   0036143018 4.750 2002     2700      1     3.273
## 38662   0036168150 3.908 2002     2700      1     2.635
## 38755   0036210052 4.381 2002     2700      1     3.108
## 38789   0036239507 4.691 2002     2700      1     3.214
## 39652   0036731990 3.940 2002     2700      1     2.767
## 39758   0036800759 3.956 2002     2700      1     2.578
## 39896   0036894541 4.121 2002     2700      1     2.743
## 40605   0346030568 5.028 2003     2700      1     3.817
## 40608   0347882750 5.233 2003     2700      1     3.917
## 40610   0348011697 4.079 2003     2700      1     2.763
## 40638  10744222633 4.103 2003     2700      1     2.787
## 40646   0344235454 4.984 2003     2700      1     3.873
## 40657   0344738736 4.084 2003     2700      1     2.973
## 40659   0345168925 4.208 2003     2700      1     2.892
## 40660   0345492466 4.735 2003     2700      1     3.419
## 40682   0345690179 5.350 2003     2700      1     4.034
## 40689   0344630326 3.954 2003     2700      1     2.638
## 40731   0346340042 4.157 2003     2700      1     2.841
## 40749   0346969977 4.442 2003     2700      1     3.231
## 40771   0348230958 6.275 2003     2700      1     4.959
## 41119   0345283207 3.840 2003     2700      1     2.524
## 41122   0345714860 5.499 2003     2700      1     4.084
```

```
## 41127   0344081179 4.283 2003     2700      1     2.868
## 41128   0344861847 4.621 2003     2700      1     3.305
## 41129   0345293127 3.833 2003     2700      1     2.517
## 41131   0345293130 4.162 2003     2700      1     3.051
## 41139   0345414665 4.343 2003     2700      1     2.928
## 41148   0344926414 5.006 2003     2700      1     3.895
## 41151   0642278662 3.806 2003     2700      1     2.595
## 41156   0344943245 4.973 2003     2700      1     3.657
## 41157   0345374595 4.903 2003     2700      1     3.587
## 41159   0642272544 4.675 2003     2700      1     3.260
## 41161   0242551979 3.904 2003     2700      1     2.588
## 41176   0242658928 4.458 2003     2700      1     3.142
## 41183   0242410368 4.973 2003     2700      1     3.657
## 41185   0242493774 4.400 2003     2700      1     2.985
## 41210   0242493661 4.273 2003     2700      1     2.957
## 41237   0344444858 4.520 2003     2700      1     3.105
## 41251   0345824713 4.352 2003     2700      1     3.036
## 41258   0346969978 5.002 2003     2700      1     3.587
## 41458   0142025460 3.894 2003     2700      1     2.578
## 41460   0142089171 4.888 2003     2700      1     3.473
## 41463   0142123411 4.222 2003     2700      1     2.906
## 41468   0142087597 4.833 2003     2700      1     3.517
## 41484   0142009533 5.306 2003     2700      1     3.891
## 41493   0142024742 4.567 2003     2700      1     3.152
## 41496   0142088521 4.407 2003     2700      1     3.091
## 41531   0141863194 4.217 2003     2700      1     2.901
## 41575   0141653838 4.222 2003     2700      1     2.906
## 41587   0141705375 4.884 2003     2700      1     3.469
## 41593   0141816759 4.768 2003     2700      1     3.353
## 41594   0141816761 4.287 2003     2700      1     3.076
## 41647   0242285697 4.478 2003     2700      1     3.162
## 41655   0242380323 3.866 2003     2700      1     2.755
## 41771   0141425718 4.211 2003     2700      1     2.796
## 41772   0141484564 4.167 2003     2700      1     2.851
## 41777   0141462439 4.563 2003     2700      1     3.148
## 41778   0141573545 5.008 2003     2700      1     3.593
## 41779   0141573546 3.904 2003     2700      1     2.793
## 41781   0141685592 4.301 2003     2700      1     2.886
## 41782   0141796735 4.455 2003     2700      1     3.244
## 41783   0141796739 4.693 2003     2700      1     3.377
## 41807   0141611906 4.138 2003     2700      1     2.822
## 41813   0141834950 4.316 2003     2700      1     3.000
## 41829   0042387879 4.992 2003     2700      1     3.577
## 41838   0042413423 4.335 2003     2700      1     3.124
## 41850   0042360213 5.273 2003     2700      1     3.957
## 41855   0041381147 3.891 2003     2700      1     2.780
## 41856   0041381153 4.095 2003     2700      1     2.779
## 41858   0042884162 3.758 2003     2700      1     2.647
## 41886   0141612004 3.855 2003     2700      1     2.539
## 41945   0142217904 3.947 2003     2700      1     2.532
```

```
## 42124   0041327804 5.015 2003     2700      1     3.699
## 42140   0042422040 4.594 2003     2700      1     3.483
## 42142   0042922806 4.186 2003     2700      1     2.975
## 42150   0041464854 4.417 2003     2700      1     3.101
## 42171   0042195833 4.746 2003     2700      1     3.430
## 42172   0042195862 4.003 2003     2700      1     2.792
## 42175   0042697063 4.958 2003     2700      1     3.642
## 42188   0041912565 4.113 2003     2700      1     2.797
## 42190   0042413567 4.065 2003     2700      1     2.650
## 42191   0042914547 3.963 2003     2700      1     2.852
## 42203   0042661252 4.135 2003     2700      1     2.720
## 42210   0041735992 4.538 2003     2700      1     3.222
## 42216   0042093742 5.388 2003     2700      1     4.177
## 42217   0042594633 4.172 2003     2700      1     2.856
## 42222   0041708062 4.398 2003     2700      1     3.287
## 42223   0041708064 4.291 2003     2700      1     3.080
## 42228   0042125511 4.751 2003     2700      1     3.336
## 42236   0041385951 4.038 2003     2700      1     2.927
## 42246   0041886776 3.994 2003     2700      1     2.678
## 42267   0042888943 4.611 2003     2700      1     3.295
## 42397   0041842634 4.262 2003     2700      1     2.946
## 42398   0042343801 5.646 2003     2700      1     4.330
## 42404   0041698076 4.135 2003     2700      1     2.720
## 42412   0038497523 4.054 2003     2700      1     2.843
## 42420   0038455694 4.334 2003     2700      1     3.018
## 42438   0038825532 3.870 2003     2700      1     2.554
## 42444   0038601952 4.724 2003     2700      1     3.613
## 42449   0038304776 5.520 2003     2700      1     4.204
## 42452   0038679758 4.587 2003     2700      1     3.172
## 42462   0038013919 4.018 2003     2700      1     2.702
## 42464   0038352047 3.855 2003     2700      1     2.744
## 42466   0038690424 5.414 2003     2700      1     4.098
## 42467   0038690437 4.549 2003     2700      1     3.338
## 42489   0038504054 4.322 2003     2700      1     2.907
## 42515   0041302382 4.281 2003     2700      1     2.965
## 42539   0042304086 4.467 2003     2700      1     3.052
## 42551   0042743965 3.866 2003     2700      1     2.550
## 42564   0043245147 4.497 2003     2700      1     3.181
## 42684   0037830097 3.972 2003     2700      1     2.656
## 42690   0037973279 5.594 2003     2700      1     4.278
## 42691   0038482206 5.915 2003     2700      1     4.704
## 42699   0038167811 4.243 2003     2700      1     3.132
## 42713   0038130715 5.029 2003     2700      1     3.713
## 42719   0037493499 4.842 2003     2700      1     3.526
## 42722   0038507413 4.230 2003     2700      1     3.119
## 42729   0038434032 3.859 2003     2700      1     2.543
## 42736   0037840165 4.892 2003     2700      1     3.576
## 42745   0038242951 4.006 2003     2700      1     2.690
## 42746   0038242968 5.038 2003     2700      1     3.722
## 42754   0037805278 4.410 2003     2700      1     2.995
```

```
## 42756   0038142845 4.115 2003     2700      1     2.904
## 42757   0038142850 3.852 2003     2700      1     2.741
## 42772   0037986208 4.721 2003     2700      1     3.405
## 42773   0038323914 4.575 2003     2700      1     3.259
## 42780   0038314212 3.960 2003     2700      1     2.644
## 42893   0042532322 4.480 2003     2700      1     3.164
## 43069   0038526363 4.506 2003     2700      1     3.091
## 43074   0037638884 5.206 2003     2700      1     3.890
## 43077   0038724280 4.848 2003     2700      1     3.737
## 43085   0037737900 5.165 2003     2700      1     3.849
## 43086   0037737901 3.822 2003     2700      1     2.506
## 43088   0037947320 3.982 2003     2700      1     2.666
## 43089   0038075468 5.272 2003     2700      1     3.956
## 43091   0038751994 4.929 2003     2700      1     3.818
## 43115   0038663167 4.361 2003     2700      1     3.045
## 43116   0038663174 4.006 2003     2700      1     2.895
## 43129   0038316599 4.797 2003     2700      1     3.382
## 43130   0038655478 3.979 2003     2700      1     2.868
## 43132   0037906573 4.466 2003     2700      1     3.051
## 43138   0012868624 4.671 2003     2700      1     3.460
## 43146   0038824056 4.162 2003     2700      1     3.051
## 43153   0038298787 4.370 2003     2700      1     3.054
## 43194   0037988905 4.148 2003     2700      1     2.832
## 43240   0038651907 4.577 2003     2700      1     3.261
## 43241   0038651918 3.988 2003     2700      1     2.672
## 43286   0242600542 4.485 2003     2700      1     3.070
## 43287   0242684416 5.438 2003     2700      1     4.227
## 43416   0037464536 4.084 2003     2700      1     2.873
## 43420   0345636017 4.427 2003     2700      1     3.111
## 43428   0038528237 4.980 2003     2700      1     3.664
## 43435   0037451905 5.245 2003     2700      1     3.929
## 43438   0037451929 6.034 2003     2700      1     4.718
## 43443   0037448932 3.966 2003     2700      1     2.650
## 43479   0037502809 5.082 2003     2700      1     3.971
## 43481   0037840394 4.140 2003     2700      1     3.029
## 43483   0037986313 5.056 2003     2700      1     3.740
## 43485   0038516861 4.525 2003     2700      1     3.314
## 43498   0037417219 4.629 2003     2700      1     3.418
## 43510   0037414165 3.901 2003     2700      1     2.585
## 43511   0037414194 4.269 2003     2700      1     2.953
## 43512   0037414217 4.123 2003     2700      1     2.807
## 43552   0037396734 4.424 2003     2700      1     3.108
## 43561   0037399439 5.125 2003     2700      1     3.809
## 43733   0037468682 3.829 2003     2700      1     2.718
## 43734   0037468691 4.429 2003     2700      1     3.113
## 43739   0037467321 3.688 2003     2700      1     2.577
## 43766   0037454273 4.160 2003     2700      1     2.844
## 43767   0037454282 4.138 2003     2700      1     2.723
## 43768   0037454283 4.138 2003     2700      1     2.822
## 43785   0037434851 3.802 2003     2700      1     2.691
```

```
## 43787   0037434895 4.553 2003     2700      1     3.237
## 43807   0037420270 4.012 2003     2700      1     2.696
## 43808   0037420274 5.142 2003     2700      1     3.826
## 43884   0037364369 4.444 2003     2700      1     3.128
## 44099   0037468409 3.866 2003     2700      1     2.550
## 44100   0037468422 4.369 2003     2700      1     3.053
## 44104   0037466917 4.108 2003     2700      1     2.693
## 44120   0037456351 4.578 2003     2700      1     3.262
## 44121   0037456358 4.402 2003     2700      1     3.086
## 44125   0037453950 4.520 2003     2700      1     3.204
## 44127   0037453976 4.581 2003     2700      1     3.166
## 44133   0037440670 3.950 2003     2700      1     2.634
## 44150   0037434515 4.739 2003     2700      1     3.423
## 44159   0037433178 4.012 2003     2700      1     2.597
## 44160   0037433181 3.866 2003     2700      1     2.550
## 44166   0037425790 3.859 2003     2700      1     2.648
## 44178   0037419908 4.233 2003     2700      1     2.917
## 44182   0037289753 4.297 2003     2700      1     2.981
## 44192   0037308404 5.103 2003     2700      1     3.787
## 44201   0037311377 4.989 2003     2700      1     3.574
## 44212   0037313797 3.966 2003     2700      1     2.650
## 44261   0037330518 3.985 2003     2700      1     2.669
## 44428   0037438792 4.376 2003     2700      1     3.060
## 44429   0037438809 4.264 2003     2700      1     3.053
## 44432   0037438936 3.985 2003     2700      1     2.669
## 44433   0037438998 4.038 2003     2700      1     2.722
## 44454   0037425535 5.533 2003     2700      1     4.118
## 44455   0037425558 4.402 2003     2700      1     3.086
## 44456   0037425564 5.704 2003     2700      1     4.388
## 44457   0037425578 5.336 2003     2700      1     4.020
## 44461   0037417523 5.839 2003     2700      1     4.628
## 44468   0037413484 4.422 2003     2700      1     3.211
## 44469   0037413492 5.094 2003     2700      1     3.883
## 44472   0037413628 3.962 2003     1000      2     2.547
## 44490   0037216353 5.144 2003     2700      1     3.828
## 44501   0037218814 4.737 2003     2700      1     3.421
## 44525   0037237980 3.908 2003     2700      1     2.592
## 44529   0037241112 4.447 2003     2700      1     3.131
## 44541   0037248459 3.918 2003     2700      1     2.707
## 44543   0037249283 3.684 2003     2700      1     2.573
## 44546   0037251810 4.453 2003     2700      1     3.137
## 44690   0037330279 4.140 2003     2700      1     2.824
## 44719   0037382166 3.982 2003     2700      1     2.666
## 44720   0037382807 4.266 2003     2700      1     2.950
## 44745   0037498595 4.430 2003     2700      1     3.219
## 44793   0037685168 4.190 2003     2700      1     2.874
## 44806   0037716856 4.108 2003     2700      1     2.792
## 44889   0038165476 3.954 2003     2700      1     2.539
## 44918   0038303393 3.711 2003     2700      1     2.600
## 45018   0038731023 4.110 2003     2700      1     2.695
```

```
## 45080   0041878535 4.148 2003     2700      1     2.832
## 45330   0242515752 3.911 2003     2700      1     2.595
## 45365   0345374590 4.847 2003     2700      1     3.531
## 45366   0345374591 4.435 2003     2700      1     3.324
## 45373   0345690174 3.988 2003     2700      1     2.672
## 45403   0346843100 3.991 2003     2700      1     2.675
## 45422   0347986777 4.410 2003     2700      1     3.094
## 46148   0042737443 4.087 2003     2700      1     2.672
## 46153   0141499228 4.645 2003     2700      1     3.434
## 46333  11144239923 5.214 2004     2700      1     3.986
## 46334  19944395825 4.649 2004     2700      1     3.421
## 46341  10844262691 4.397 2004     2700      1     3.374
## 46342  10844281210 4.459 2004     2700      1     3.231
## 46367  10344231441 4.676 2004     2700      1     3.348
## 46368  10344239881 4.585 2004     2700      1     3.357
## 46369  10344242939 4.474 2004     2700      1     3.246
## 46375  10344221567 4.188 2004     2700      1     2.960
## 46376  10344229444 4.159 2004     2700      1     3.036
## 46381  11344275794 3.847 2004     2700      1     2.619
## 46424  19744365355 3.936 2004     2700      1     2.913
## 46467  10044271620 4.226 2004     2700      1     2.998
## 46589  15244363856 4.164 2004     2700      1     2.836
## 46593  15744363493 3.986 2004     2700      1     2.863
## 46597  15744399190 4.100 2004     2700      1     2.872
## 46820   9644252909 4.467 2004     2700      1     3.239
## 47011   9244240268 3.964 2004     2700      1     2.736
## 47018   9244240769 4.129 2004     2700      1     2.801
## 47019   9244247612 4.578 2004     2700      1     3.555
## 47020   9244264413 3.803 2004     2700      1     2.575
## 47044   8544252449 4.680 2004     2700      1     3.557
## 47045   8544258102 4.526 2004     2700      1     3.403
## 47053   8444225132 4.176 2004     2700      1     2.948
## 47062  19644400578 4.870 2004     2700      1     3.542
## 47069   7744231805 4.691 2004     2700      1     3.463
## 47070   7744237066 4.382 2004     2700      1     3.154
## 47072   7744239901 4.468 2004     2700      1     3.345
## 47085   7644221218 3.947 2004     2700      1     2.824
## 47140  13844273390 4.129 2004     2700      1     3.006
## 47147  13944251605 4.572 2004     2700      1     3.344
## 47168  15244339164 4.231 2004     2700      1     3.003
## 47393   6944244875 4.867 2004     2700      1     3.639
## 47395   6944248992 4.460 2004     2700      1     3.232
## 47399   6944229468 4.497 2004     2700      1     3.269
## 47407   6944232728 4.311 2004     2700      1     3.083
## 47443   5044248578 3.831 2004     2700      1     2.603
## 47458  19544385747 5.348 2004     2700      1     4.120
## 47601   7444240833 3.982 2004     2700      1     2.859
## 47764   4544305451 4.264 2004     2700      1     3.241
## 47790   4544279029 4.188 2004     2700      1     2.960
## 47806   4544332903 5.079 2004     2700      1     3.851
```

```
## 47815   4544222914 3.776 2004     2700      1     2.548
## 47825   4444358501 4.553 2004     2700      1     3.325
## 47843   4544329012 3.814 2004     2700      1     2.586
## 47851   4344690525 4.410 2004     2700      1     3.082
## 47995   4944239035 4.273 2004     2700      1     3.045
## 48006   5444255241 3.855 2004     2700      1     2.627
## 48052   8544244084 4.077 2004     2700      1     2.849
## 48054   9644268242 3.755 2004     2700      1     2.527
## 48197   4143052707 4.043 2004     2700      1     2.815
## 48207   4143067005 4.774 2004     2700      1     3.546
## 48208   4143069253 3.924 2004     2700      1     2.696
## 48227   4143071570 4.102 2004     2700      1     2.774
## 48228   4143094988 5.453 2004     2700      1     4.225
## 48236   4043076922 3.999 2004     2700      1     2.771
## 48238   4043153049 4.489 2004     2700      1     3.261
## 48239   4043156247 4.890 2004     2700      1     3.662
## 48367   4043055316 4.632 2004     2700      1     3.404
## 48373   4043152981 3.933 2004     2700      1     2.705
## 48385   4344583554 4.318 2004     2700      1     3.090
## 48512   3342892905 4.641 2004     2700      1     3.413
## 48522   3242759883 3.736 2004     2700      1     2.508
## 48553   3142735110 4.217 2004     2700      1     3.194
## 48554   3142745348 4.790 2004     2700      1     3.462
## 48585   3042728480 4.043 2004     2700      1     3.020
## 48665   3042642128 5.510 2004     2700      1     4.282
## 48672   3042731163 4.134 2004     2700      1     2.906
## 48674   3042781155 3.884 2004     2700      1     2.656
## 48679   3042821849 3.670 2004     2700      1     2.647
## 48702   3242749074 4.079 2004     2700      1     2.851
## 48705   3242754344 3.746 2004     2700      1     2.623
## 48718   4043070821 3.779 2004     2700      1     2.551
## 48720   4043082182 3.928 2004     2700      1     2.700
## 48721   4043092238 4.451 2004     2700      1     3.223
## 48951   2942537772 4.149 2004     2700      1     3.026
## 48952   2942554887 5.618 2004     2700      1     4.495
## 48957   2942627194 4.077 2004     2700      1     2.849
## 49006   2542548063 5.360 2004     2700      1     4.337
## 49069   2942709937 3.782 2004     2700      1     2.554
## 49074   2942720933 4.617 2004     2700      1     3.289
## 49281   2442665224 4.280 2004     2700      1     3.052
## 49283   2442715038 5.135 2004     2700      1     3.907
## 49290   2442696436 4.547 2004     2700      1     3.319
## 49306   2342471392 6.269 2004     2700      1     5.041
## 49307   2442479695 5.758 2004     2700      1     4.530
## 49309   2442572117 5.004 2004     2700      1     3.776
## 49313   2442590642 4.673 2004     2700      1     3.445
## 49338   2342464257 4.234 2004     2700      1     3.006
## 49345   2342486731 4.168 2004     2700      1     2.940
## 49369   2342517421 5.147 2004     2700      1     4.124
## 49373   3843094224 4.764 2004     2700      1     3.536
```

```
## 49380   2342501856 4.398 2004     2700      1     3.170
## 49381   2342564429 4.008 2004     2700      1     2.780
## 49424   2342488880 3.755 2004     2700      1     2.527
## 49427   2342516208 3.700 2004     2700      1     2.677
## 49428   2342524107 3.852 2004     2700      1     2.624
## 49434   2342646956 4.407 2004     2700      1     3.179
## 49472   2942536636 4.358 2004     2700      1     3.130
## 49475   2942561024 4.094 2004     2700      1     2.866
## 49649   2142649221 4.212 2004     2700      1     3.189
## 49651   2142758687 4.063 2004     2700      1     2.835
## 49664   1942436013 4.103 2004     2700      1     2.875
## 49674  11144354823 4.207 2004     2700      1     2.979
## 49696  12144291219 3.697 2004     2700      1     2.574
## 49710   3042762336 4.314 2004     2700      1     3.086
## 49718  11144357519 3.986 2004     2700      1     2.658
## 49722   1842864234 5.018 2004     2700      1     3.790
## 49739   1842815777 4.502 2004     2700      1     3.174
## 49743  11144356419 3.866 2004     2700      1     2.638
## 49830   2442509789 3.581 2004     2700      1     2.558
## 49967  12144288049 4.590 2004     2700      1     3.362
## 49970   1642400686 4.214 2004     2700      1     2.986
## 49977   1642296706 3.739 2004     2700      1     2.511
## 50015  11144355828 4.350 2004     2700      1     3.122
## 50019  10744229257 3.730 2004     2700      1     2.502
## 50026   1642369930 4.229 2004     2700      1     3.001
## 50055   1442355581 4.481 2004     2700      1     3.253
## 50062  10744220250 4.551 2004     2700      1     3.223
## 50063  10744225301 5.224 2004     2700      1     3.996
## 50065   1442353066 3.979 2004     2700      1     2.751
## 50071  11144357189 3.839 2004     2700      1     2.511
## 50087   1442357943 3.847 2004     2700      1     2.619
## 50332  10744233940 4.677 2004     2700      1     3.449
## 50335   1342331884 3.850 2004     2700      1     2.622
## 50388   1442314663 3.997 2004     2700      1     2.669
## 50413  10744223871 4.045 2004     2700      1     2.817
## 50414   1342288777 4.142 2004     2700      1     3.119
## 50468   1642327683 4.090 2004     2700      1     2.967
## 50613   1642499234 4.500 2004     2700      1     3.272
## 50622   1642540483 4.321 2004     2700      1     2.993
## 50623   9144252520 4.184 2004     2700      1     2.956
## 50643   0345872128 4.849 2004     2700      1     3.621
## 50657   0347948542 3.871 2004     2700      1     2.643
## 50661   0345824715 4.381 2004     2700      1     3.153
## 50674   0346599193 4.602 2004     2700      1     3.374
## 50675   0347416893 3.842 2004     2700      1     2.614
## 50678   0348047527 4.767 2004     2700      1     3.439
## 50695   0345830478 3.919 2004     2700      1     2.591
## 50712   0346787782 3.666 2004     2700      1     2.543
## 50713   0346787909 3.820 2004     2700      1     2.592
## 50725   0347418277 4.315 2004     2700      1     3.292
```

```
## 50732   0742306162 3.544 2004     2700      1     2.521
## 50771  10944261840 3.779 2004     2700      1     2.656
## 50818   1342332130 3.828 2004     2700      1     2.600
## 50848  13744263906 4.057 2004     2700      1     3.034
## 50998   1542571327 3.817 2004     2700      1     2.589
## 51428   2142717404 4.031 2004     2700      1     2.803
## 51821   3242733807 3.931 2004     2700      1     2.703
## 51895   3442899889 4.734 2004     2700      1     3.506
## 51927   4043085036 3.742 2004     2700      1     2.514
## 52262  85047690933 4.353 2004     2700      1     3.125
## 52323   9644283066 4.178 2004     2700      1     2.950
## 52387   3142514201 3.624 2004     1300      2     2.601
## 52521  29544447206 3.996 2005     2700      1     2.697
## 52524  29544437848 4.305 2005     2700      1     3.006
## 52539  29144451858 4.279 2005     2700      1     3.184
## 52543  29144490030 3.866 2005     2700      1     2.567
## 52545  29144533834 4.497 2005     2700      1     3.297
## 52558  28944434329 4.369 2005     2700      1     3.169
## 52561  28944445445 4.451 2005     2700      1     3.356
## 52565  28944433025 3.552 2005     2700      1     2.557
## 52569  28944447646 4.919 2005     2700      1     3.620
## 52607  28944437578 4.673 2005     2700      1     3.678
## 52614  28844472594 3.996 2005     2700      1     2.697
## 52618  28844501802 4.241 2005     2700      1     3.041
## 52767  30144444279 4.733 2005     2700      1     3.533
## 52796  31044442642 3.817 2005     2700      1     2.617
## 52797  31044444463 4.156 2005     2700      1     2.956
## 52803  31044451759 4.265 2005     2700      1     3.065
## 52804  31044452189 4.001 2005     2700      1     2.702
## 53762  28144433147 4.679 2005     2700      1     3.380
## 53764  28144453057 4.091 2005     2700      1     2.891
## 53771  28144443357 3.807 2005     2700      1     2.712
## 53772  28144451163 4.332 2005     2700      1     3.132
## 53773  28144453509 3.901 2005     2700      1     2.806
## 53794  27744431927 4.342 2005     2700      1     3.142
## 53795  27744494434 4.618 2005     2700      1     3.623
## 53807  27744496592 4.514 2005     2700      1     3.314
## 53820  27744606737 4.954 2005     2700      1     3.754
## 53828  27744477622 3.728 2005     2700      1     2.528
## 53845  27644443333 4.575 2005     2700      1     3.580
## 53863  27644452293 3.749 2005     2700      1     2.549
## 53865  27644461904 3.702 2005     2700      1     2.707
## 53875  27644513772 4.335 2005     2700      1     3.240
## 53883  27644568878 4.218 2005     2700      1     3.223
## 53918  28444478439 4.387 2005     2700      1     3.088
## 54155  27244440305 4.455 2005     2700      1     3.360
## 54170  26844536978 5.830 2005     2700      1     4.630
## 54171  26844552488 5.412 2005     2700      1     4.212
## 54178  26844494491 4.633 2005     2700      1     3.538
## 54179  26844538114 4.966 2005     2700      1     3.766
```

```
## 54180  26844544418 3.601 2005     2700      1     2.506
## 54208  26444452073 5.212 2005     2700      1     3.913
## 54210  26444477603 4.006 2005     2700      1     3.011
## 54211  26444506232 4.042 2005     2700      1     2.842
## 54217  26444574824 4.419 2005     2700      1     3.120
## 54243  25844488342 3.971 2005     2700      1     2.771
## 54246  25844509466 4.220 2005     2700      1     2.921
## 54285  25844438380 4.442 2005     2700      1     3.242
## 54287  25844440999 3.660 2005     2700      1     2.665
## 54305  26444515022 3.926 2005     2700      1     2.627
## 54312  26444551183 3.878 2005     2700      1     2.678
## 54314  26444577543 3.881 2005     2700      1     2.681
## 54319  26444598506 4.478 2005     2700      1     3.278
## 54557  25444461935 3.898 2005     2700      1     2.698
## 54562  25144470719 3.920 2005     2700      1     2.720
## 54575  25144456112 5.883 2005     2700      1     4.683
## 54580  25144502974 3.881 2005     2700      1     2.681
## 54581  25144505285 3.742 2005     2700      1     2.542
## 54582  25144518364 4.074 2005     2700      1     2.874
## 54614  24944531368 3.878 2005     2700      1     2.678
## 54632  24644439190 4.674 2005     2700      1     3.474
## 54633  24644443217 4.833 2005     2700      1     3.633
## 54677  24344483122 3.953 2005     2700      1     2.753
## 54701  24044484768 3.823 2005     2700      1     2.728
## 54742  24744470476 4.035 2005     2700      1     2.835
## 55090  23944436856 3.575 2005     2700      1     2.580
## 55101  23944436934 4.091 2005     2700      1     2.891
## 55113  23844539714 4.976 2005     2700      1     3.776
## 55121  23844454293 3.800 2005     2700      1     2.600
## 55123  23844494864 4.211 2005     2700      1     2.912
## 55133  23844533906 4.026 2005     2700      1     2.826
## 55149  23744500093 4.222 2005     2700      1     2.923
## 55162  23444448145 3.866 2005     2700      1     2.771
## 55163  23444452077 4.300 2005     2700      1     3.100
## 55164  23744459272 3.820 2005     2700      1     2.620
## 55189  23044480500 3.937 2005     2700      1     2.737
## 55193  23044442114 4.460 2005     2700      1     3.260
## 55235  23644435612 3.976 2005     2700      1     2.677
## 55238  23644439061 4.516 2005     2700      1     3.316
## 55240  23644446866 3.989 2005     2700      1     2.789
## 55244  23644456321 3.901 2005     2700      1     2.701
## 55245  23644460325 3.942 2005     2700      1     2.947
## 55428  22844445230 4.774 2005     2700      1     3.475
## 55430  22844447051 4.585 2005     2700      1     3.385
## 55438  22844434877 4.780 2005     2700      1     3.580
## 55442  22844448780 4.222 2005     2700      1     3.022
## 55448  22144468433 3.996 2005     2700      1     2.796
## 55466  22344443114 4.392 2005     2700      1     3.192
## 55468  22344450570 4.102 2005     2700      1     2.902
## 55471  22344457679 4.119 2005     2700      1     2.919
```

```
## 55485  22244446183 4.420 2005     2700      1     3.425
## 55486  22244485721 4.183 2005     2700      1     3.088
## 55521  22144499588 3.713 2005     2700      1     2.618
## 55536  21444447949 4.377 2005     2700      1     3.177
## 55569  21444441271 4.299 2005     2700      1     3.000
## 55570  21444447966 4.369 2005     2700      1     3.169
## 55571  21444450527 4.501 2005     2700      1     3.406
## 55591  22144440464 4.362 2005     2700      1     3.162
## 55605  22144492953 4.207 2005     2700      1     2.908
## 55606  22144495740 3.790 2005     2700      1     2.590
## 55869  20644465409 3.857 2005     2700      1     2.657
## 55886  20444444376 4.187 2005     2700      1     2.987
## 55888  20444501831 4.934 2005     2700      1     3.734
## 55913  20444398121 3.931 2005     2700      1     2.632
## 55914  20544462061 4.452 2005     2700      1     3.357
## 55957  21144448879 4.965 2005     2700      1     3.765
## 55970  18844374582 3.884 2005     2700      1     2.889
## 55986  19644379430 3.989 2005     2700      1     2.789
## 55989  19644391940 4.484 2005     2700      1     3.185
## 56001  20444454585 3.817 2005     2700      1     2.617
## 56002  20444457518 4.596 2005     2700      1     3.601
## 56009  20444499355 4.459 2005     2700      1     3.259
## 56310  19444363154 3.931 2005     2700      1     2.731
## 56315  21144457072 4.436 2005     2700      1     3.236
## 56318  18944367562 4.915 2005     2700      1     3.616
## 56337  18644370841 4.807 2005     2700      1     3.812
## 56339  18744382834 3.953 2005     2700      1     2.958
## 56341  20844438843 4.222 2005     2700      1     3.022
## 56349  18244371651 5.531 2005     2700      1     4.331
## 56350  20844432890 4.096 2005     2700      1     3.001
## 56396  18744401357 3.950 2005     2700      1     2.651
## 56397  18744406319 3.706 2005     2700      1     2.506
## 56407  17944362664 4.751 2005     2700      1     3.452
## 56409  17944380530 3.760 2005     2700      1     2.560
## 56410  20944448998 3.953 2005     2700      1     2.753
## 56491  20944436157 4.708 2005     2700      1     3.508
## 56656  17644392071 4.823 2005     2700      1     3.524
## 56664  17544364774 4.269 2005     2700      1     3.069
## 56684  17844402965 4.271 2005     2700      1     3.071
## 56693  17244379193 4.136 2005     2700      1     2.936
## 56702  17144385790 5.319 2005     2700      1     4.020
## 56705  17144430418 4.592 2005     2700      1     3.392
## 56707  21544449115 4.018 2005     2700      2     2.719
## 56712  17144398838 3.739 2005     2700      1     2.539
## 56719  20244366093 3.760 2005     2700      1     2.560
## 56730  20144386747 4.154 2005     2700      1     2.954
## 56731  20144388867 4.033 2005     2700      1     2.833
## 56738  16444381731 3.829 2005     2700      1     2.734
## 56740  20144387681 4.355 2005     2700      1     3.155
## 56748  15744372810 5.001 2005     2700      1     3.801
```

```
## 56757  20144386837 4.994 2005     2700      1     3.999
## 56762  15944387060 3.940 2005     2700      1     2.641
## 56763  15944406488 4.216 2005     2700      1     3.016
## 56824  20144387499 3.895 2005     2700      1     2.596
## 56965  15744378579 5.147 2005     2700      1     4.152
## 56967  20144366696 4.939 2005     2700      1     3.739
## 56992  15044344420 3.717 2005     2700      1     2.517
## 57041  14744275842 4.576 2005     2700      1     3.376
## 57048  14544304095 5.144 2005     2700      1     3.944
## 57060  20044369898 3.543 2005     2700      1     2.548
## 57105  14644435711 3.499 2005     2700      1     2.504
## 57189  20044368802 4.056 2005     2700      1     2.856
## 57354  13844315270 3.929 2005     2700      1     2.630
## 57359  13944250706 3.610 2005     2700      1     2.615
## 57391  13744253924 3.839 2005     2700      1     2.744
## 57392  13744261333 3.974 2005     2700      1     2.774
## 57400  13844265752 4.225 2005     2700      1     3.025
## 57402  13844274980 3.923 2005     2700      1     2.723
## 57417  13444256401 4.619 2005     2700      1     3.419
## 57420  13444293219 5.349 2005     2700      1     4.149
## 57426  13444309949 4.491 2005     2700      1     3.291
## 57427  13444310598 3.728 2005     2700      1     2.528
## 57437  13444263619 4.684 2005     2700      1     3.484
## 57495  14944348752 3.863 2005     2700      1     2.663
## 57674  19944432606 4.121 2005     2700      1     3.126
## 57678  12544253745 4.628 2005     2700      1     3.428
## 57683  12544253835 4.548 2005     2700      1     3.348
## 57684  12544257550 4.190 2005     2700      1     2.990
## 57685  19944431657 4.004 2005     2700      1     2.804
## 57698  12144284532 3.878 2005     2700      1     2.883
## 57714  19944426812 4.047 2005     2700      1     2.748
## 57732  11844268070 4.117 2005     2700      1     2.917
## 57734  13744249546 3.742 2005     2700      1     2.647
## 57739  11844269164 4.196 2005     2700      1     3.101
## 57784  19944429519 4.900 2005     2700      1     3.700
## 57787  11244285095 3.904 2005     2700      1     2.605
## 57788  11244331742 4.906 2005     2700      1     3.706
## 57831  12744263350 4.085 2005     2700      1     3.090
## 57836  12744278427 3.829 2005     2700      1     2.530
## 57868  13444260835 3.732 2005     2700      1     2.637
## 57870  13444268127 4.350 2005     2700      1     3.355
## 57881  13744254861 3.749 2005     2700      1     2.549
## 57883  13744256812 4.438 2005     2700      1     3.139
## 57919  14144253829 3.857 2005     2700      1     2.657
## 57965  14644425916 3.631 2005     2700      1     2.536
## 58080  16244415880 3.909 2005     2700      1     2.914
## 58139  17644377222 4.115 2005     2700      1     3.120
## 58156  17844391245 3.929 2005     2700      1     2.729
## 58196  18244390229 3.826 2005     2700      1     2.831
## 58197  18244393432 3.533 2005     2700      1     2.538
```

```
## 58304  20144388724 3.857 2005     2700      1     2.657
## 58349  20944446791 4.391 2005     2700      1     3.191
## 58709  26944497954 3.884 2005     2700      1     2.684
## 59586  33845866114 3.911 2006     2700      1     2.847
## 59588  33845880922 5.714 2006     2700      1     4.749
## 59600  33845866055 4.159 2006     2700      1     2.990
## 59601  33845890326 4.968 2006     2700      1     3.799
## 59635  33845693164 3.800 2006     2700      1     2.631
## 59636  33845707784 4.606 2006     2700      1     3.437
## 59643  33845717310 4.910 2006     2700      1     3.846
## 59678  33845483482 3.948 2006     2700      1     2.983
## 59679  33845490014 6.122 2006     2700      1     4.953
## 59680  33845497716 3.958 2006     2700      1     2.689
## 59681  33845500545 4.941 2006     2700      1     3.772
## 59686  33845478495 4.306 2006     2700      1     3.037
## 59687  33845485717 3.595 2006     2700      1     2.531
## 59688  33845491101 3.880 2006     2700      1     2.611
## 59689  33947325340 4.051 2006     2700      1     2.987
## 59705  33845449051 4.316 2006     2700      1     3.147
## 59714  33845450660 4.586 2006     2700      1     3.317
## 59746  33748560432 4.653 2006     2700      1     3.484
## 59787  33845297656 3.725 2006     2700      1     2.556
## 59788  33845299821 3.784 2006     2700      1     2.615
## 59792  33845317604 4.246 2006     2700      1     3.077
## 59793  33845322277 3.595 2006     2700      1     2.630
## 59798  33845354407 3.759 2006     2700      1     2.590
## 59849  33845926387 3.808 2006     2700      1     2.639
## 59883  33846317116 3.734 2006     2700      1     2.769
## 60010  33847190697 3.751 2006     2700      1     2.582
## 61367  33751529506 4.659 2006     2700      1     3.390
## 61370  33751545838 5.100 2006     2700      1     3.831
## 61389  33751185526 3.780 2006     2700      1     2.716
## 61412  33751206860 4.493 2006     2700      1     3.324
## 61413  33751217149 4.658 2006     2700      1     3.489
## 61441  33750983605 5.495 2006     2700      1     4.326
## 61448  33750978404 4.084 2006     2700      1     2.915
## 61491  33750694575 4.082 2006     2700      1     2.913
## 61499  33750726707 4.632 2006     2700      1     3.568
## 61500  33750731595 4.693 2006     2700      1     3.524
## 61521  33750630655 3.751 2006     2700      1     2.786
## 61532  33750500326 4.054 2006     2700      1     2.885
## 61535  33750507268 4.113 2006     2700      1     2.844
## 61564  33750507246 3.788 2006     2700      1     2.619
## 61566  33750518186 4.246 2006     2700      1     3.281
## 61570  33750532312 5.148 2006     2700      1     4.183
## 61574  33750584214 5.548 2006     2700      1     4.279
## 61582  33750597734 4.320 2006     2700      1     3.151
## 62170  33750346251 4.432 2006     2700      1     3.368
## 62188  33750300528 4.116 2006     2700      1     2.947
## 62212  33750035507 4.320 2006     2700      1     3.355
```

```
## 62216  33750081777 3.983 2006     2700      1     2.714
## 62223  33750106228 4.641 2006     2700      1     3.472
## 62224  33750121615 4.499 2006     2700      1     3.330
## 62227  33749636443 3.738 2006     2700      1     2.773
## 62260  33749597936 4.523 2006     2700      1     3.254
## 62261  33749599695 5.460 2006     2700      1     4.495
## 62263  33749618085 4.879 2006     2700      1     3.710
## 62273  33749613119 5.139 2006     2700      1     4.174
## 62274  33749659678 3.989 2006     2700      1     2.820
## 62294  33749436870 4.405 2006     2700      1     3.236
## 62297  33749445317 5.971 2006     2700      1     4.802
## 62319  33749020857 4.306 2006     2700      1     3.137
## 62325  33749430001 3.938 2006     2700      1     2.973
## 62331  33749441325 4.657 2006     2700      1     3.488
## 62335  33749446861 3.772 2006     2700      1     2.603
## 62472  33846664981 3.928 2006     2700      1     2.759
## 62539  39049174296 3.751 2006     2700      1     2.582
## 62724  43749107283 4.870 2006     2700      1     3.701
## 62921  33749040947 3.626 2006     2700      1     2.562
## 62924  33749079234 3.958 2006     2700      1     2.789
## 63010  33748354709 3.584 2006     2700      1     2.619
## 63021  33748316794 3.928 2006     2700      1     2.963
## 63047  33748437811 4.028 2006     2700      1     2.859
## 63048  33748438134 4.637 2006     2700      1     3.468
## 63124  33748292955 4.111 2006     2700      1     2.942
## 63132  33748312093 4.065 2006     2700      1     2.896
## 63137  33748331335 3.890 2006     2700      1     2.621
## 63152  33748747557 3.876 2006     2700      1     2.707
## 63833  33748161812 4.051 2006     2700      1     2.882
## 63860  33747834620 3.716 2006     2700      1     2.547
## 63864  33747870163 5.283 2006     2700      1     4.114
## 63866  33747584933 4.308 2006     2700      1     3.139
## 63887  33747135682 4.969 2006     2700      1     3.800
## 63904  33747343208 5.434 2006     2700      1     4.265
## 63910  33747102040 3.816 2006     2700      1     2.647
## 63948  33746875641 4.520 2006     2700      1     3.351
## 63985  33746768834 3.832 2006     2700      1     2.563
## 63987  33746587393 4.057 2006     2700      1     2.993
## 64006  33746437130 4.877 2006     2700      1     3.608
## 64017  33746667851 4.079 2006     2700      1     2.910
## 64020  33746672599 3.964 2006     2700      1     2.695
## 64021  33746675685 4.031 2006     2700      1     2.862
## 64033  33746724157 3.887 2006     2700      1     2.618
## 64083  33748046126 3.747 2006     2700      1     2.578
## 64148  33750630109 3.948 2006     2700      1     2.679
## 64376  33746387027 3.828 2006     2700      1     2.659
## 64377  33746430561 3.742 2006     2700      1     2.573
## 64378  33746430578 4.001 2006     2700      1     2.832
## 64379  33746452905 4.192 2006     2700      1     3.128
## 64395  33748513704 3.579 2006     2700      1     2.614
```

```
## 64421  33746075560 4.642 2006     2700      1     3.473
## 64422  33746088001 5.656 2006     2700      1     4.691
## 64482  33745890309 4.593 2006     2700      1     3.324
## 64484  33745913326 3.707 2006     2700      1     2.643
## 64512  33745698681 4.248 2006     2700      1     3.079
## 64514  33745698685 4.688 2006     2700      1     3.624
## 64520  33745611449 5.100 2006     2700      1     4.135
## 64521  33745612500 3.983 2006     2700      1     2.714
## 64522  33745614361 5.082 2006     2700      1     4.117
## 64523  33745632422 4.146 2006     2700      1     2.877
## 64530  33745812011 4.129 2006     2700      1     2.960
## 64531  33745817034 3.992 2006     2700      1     3.027
## 64535  33745840137 3.992 2006     2700      1     2.823
## 64537  33745851479 4.385 2006     2700      1     3.216
## 64548  33744914821 3.540 2006     2700      1     2.575
## 64979  33745661633 3.776 2006     2700      1     2.712
## 65008  33745227399 4.917 2006     2700      1     3.748
## 65009  33745227751 3.734 2006     2700      1     2.565
## 65010  33745255382 4.457 2006     2700      1     3.393
## 65030  33646360865 4.352 2006     2700      1     3.183
## 65051  33745081608 5.495 2006     2700      1     4.530
## 65054  33745102555 5.215 2006     2700      1     4.046
## 65066  33745065872 3.734 2006     2700      1     2.769
## 65093  33744718154 3.600 2006     2700      1     2.635
## 65106  33744966030 4.777 2006     2700      1     3.608
## 65135  33646825385 4.208 2006     2700      1     3.039
## 65148  33744472167 3.702 2006     2700      1     2.533
## 65149  33744472168 4.090 2006     2700      1     2.921
## 65151  33744498063 4.280 2006     2700      1     3.111
## 65190  33745210374 3.862 2006     2700      1     2.693
## 65191  33745212925 3.876 2006     2700      1     2.707
## 65192  33745217611 3.764 2006     2700      1     2.595
## 65230  33745685701 3.747 2006     2700      1     2.578
## 65338  33750089034 3.980 2006     2700      1     3.015
## 65365  33846406448 3.854 2006     2700      1     2.685
## 65669  33646859471 4.708 2006     2700      1     3.743
## 65672  33646875293 4.246 2006     2700      1     2.977
## 65680  33646828953 3.862 2006     2700      1     2.897
## 65716  33646676272 4.484 2006     2700      1     3.315
## 65717  33646678189 4.410 2006     2700      1     3.241
## 65764  33646482407 3.887 2006     2700      1     2.618
## 65765  33646488054 4.825 2006     2700      1     3.556
## 65772  33646450281 3.851 2006     2700      1     2.582
## 65773  33646452922 4.459 2006     2700      1     3.290
## 65790  33646475441 3.584 2006     2700      1     2.619
## 65816  30044438368 5.433 2006     2700      1     4.369
## 65837  33646177653 3.505 2006     2700      1     2.540
## 65838  33646178951 5.085 2006     2700      1     3.916
## 65847  33646392755 4.004 2006     2700      1     2.835
## 65850  33646401549 3.977 2006     2700      1     2.808
```

```
## 65975  33745489614 3.646 2006     2700      1     2.582
## 66274  33646164892 4.031 2006     2700      1     2.862
## 66300  33645865234 4.803 2006     2700      1     3.634
## 66319  33646052556 5.164 2006     2700      1     3.995
## 66338  33645770270 4.219 2006     2700      1     3.155
## 66344  33645729203 4.579 2006     2700      1     3.410
## 66346  33645741222 4.267 2006     2700      1     2.998
## 66372  33645523067 6.444 2006     2700      1     5.380
## 66392  29944434110 3.964 2006     2700      1     2.795
## 66393  29944434742 4.119 2006     2700      1     2.850
## 66394  29944442726 4.794 2006     2700      1     3.625
## 66408  33645222620 3.812 2006     2700      1     2.543
## 66409  33645236930 3.684 2006     2700      1     2.515
## 66423  33645530745 4.561 2006     2700      1     3.392
## 66490  33646420626 3.788 2006     2700      1     2.619
## 66492  33646432226 4.312 2006     2700      1     3.248
## 66777  33645390604 3.621 2006     2700      1     2.656
## 66778  33645399959 4.643 2006     2700      1     3.474
## 66806  33645102811 3.897 2006     2700      1     2.833
## 66807  33645103550 5.432 2006     2700      1     4.467
## 66813  33645068471 4.040 2006     2700      1     2.771
## 66834  33745029895 3.832 2006     2700      1     2.768
## 66845  33644979597 4.290 2006     2700      1     3.325
## 66847  33644984085 4.144 2006     2700      1     2.975
## 66856  33644967633 3.995 2006     2700      1     3.030
## 66872  33644957630 4.031 2006     2700      1     2.862
## 66880  33644833147 4.678 2006     2700      1     3.509
## 66883  33644807437 5.185 2006     2700      1     3.916
## 66925  33644584065 4.544 2006     2700      1     3.275
## 66926  33644605043 5.148 2006     2700      1     3.979
## 66930  33644638676 3.816 2006     2700      1     2.647
## 66932  33644642974 3.921 2006     2700      1     2.752
## 66942  33644654777 4.106 2006     2700      1     2.937
## 66944  33644660402 3.689 2006     2700      1     2.625
## 67001  33645355399 4.176 2006     2700      1     3.007
## 67354  33344467770 4.351 2006     2700      1     3.182
## 67363  33644536465 4.565 2006     2700      1     3.396
## 67394  32644457434 4.090 2006     2700      1     2.921
## 67395  32644467389 5.178 2006     2700      1     3.909
## 67432  32144436098 4.713 2006     2700      1     3.444
## 67435  32144455695 4.106 2006     2700      1     2.937
## 67438  32544459770 3.748 2006     1000      2     2.579
## 67445  32144432645 4.141 2006     2700      1     3.077
## 67446  32144441493 4.605 2006     2700      1     3.640
## 67447  32144443648 4.776 2006     2700      1     3.712
## 67455  32544444748 3.983 2006     2700      1     2.714
## 67461  32044452414 4.031 2006     2700      1     2.862
## 67474  31444436871 3.689 2006     2700      1     2.520
## 67506  32444436464 3.716 2006     2700      1     2.547
## 67508  32444438623 3.828 2006     2700      1     2.659
```

```
## 67510  32444443951 4.627 2006     2700      1     3.358
## 67517  32444450909 3.534 2006     2700      1     2.569
## 67538  33644842676 4.351 2006     2700      1     3.182
## 67781  31544462628 3.679 2006     1000      2     2.510
## 67786  31344467254 4.933 2006     2700      1     3.968
## 67787  31344470705 5.219 2006     2700      1     3.950
## 67790  31344479326 4.396 2006     2700      1     3.227
## 67795  31344454603 3.772 2006     2700      1     2.603
## 67804  30844432417 3.675 2006     2700      1     2.506
## 67813  30944452384 4.432 2006     2700      1     3.263
## 67815  30944457531 4.478 2006     2700      1     3.414
## 67876  32244433703 3.636 2006     2700      1     2.671
## 67920  30944433554 4.252 2006     2700      1     3.083
## 67933  31044431832 4.425 2006     2700      1     3.256
## 67938  31044436630 4.688 2006     2700      1     3.519
## 67941  31044442783 3.734 2006     2700      1     2.565
## 67942  31044446194 4.154 2006     2700      1     3.189
## 67944  31044453008 4.376 2006     2700      1     3.207
## 67945  31044453148 4.025 2006     2700      1     2.856
## 67950  31044456529 5.039 2006     2700      1     3.870
## 67972  32044434020 3.847 2006     2700      1     2.678
## 67974  32044435429 4.349 2006     2700      1     3.080
## 67991  32144461605 3.995 2006     2700      1     2.726
## 68363  33748146129 3.511 2006     2700      1     2.546
## 68847  37549036689 4.103 2007     2700      1     3.151
## 68848  37549039095 3.740 2007     2700      1     2.888
## 68868  37349092105 3.950 2007     2700      1     2.893
## 68877  37349011273 3.858 2007     2700      1     2.701
## 68878  37349028816 4.130 2007     2700      1     3.073
## 68923  37249041477 3.789 2007     2700      1     2.732
## 68924  37249061691 4.883 2007     2700      1     3.826
## 68931  37149028276 4.444 2007     2700      1     3.592
## 68952  38449085341 3.654 2007     2700      1     2.597
## 68959  36549078679 4.337 2007     2700      1     3.280
## 68962  37249065591 3.858 2007     2700      1     2.906
## 69125  36048934283 3.807 2007     2700      1     2.955
## 69128  36049048260 3.795 2007     2700      1     2.638
## 69155  36849011221 4.476 2007     2700      1     3.419
## 69168  36849035575 3.973 2007     2700      1     2.916
## 69181  36849068108 4.107 2007     2700      1     2.950
## 69184  36849079348 3.564 2007     2700      1     2.612
## 69237  38049023539 4.325 2007     2700      1     3.373
## 69238  38049038389 3.992 2007     2700      1     2.935
## 69239  38049053061 4.415 2007     2700      1     3.358
## 69241  38049087576 4.711 2007     2700      1     3.654
## 69968  56149101731 3.783 2007     2700      1     2.626
## 70668  36549006511 3.668 2007     2700      1     2.611
## 70677  36549033449 4.134 2007     2700      1     2.977
## 70678  36549066931 4.045 2007     2700      1     3.193
## 70725  36348961929 4.189 2007     2700      1     3.132
```

```
## 70727  36348982774 4.230 2007     2700      1     3.173
## 70728  36349035495 4.062 2007     2700      1     2.905
## 70766  36148981435 4.360 2007     2700      1     3.303
## 70798  36148951701 3.556 2007     2700      1     2.604
## 70799  36148966515 4.227 2007     2700      1     3.070
## 70800  36149000149 5.380 2007     2700      1     4.323
## 70832  35848963822 4.711 2007     2700      1     3.654
## 70833  35848964343 4.234 2007     2700      1     3.282
## 70835  35848969753 4.853 2007     2700      1     3.796
## 70842  35848935201 5.343 2007     2700      1     4.186
## 70844  35848964571 4.402 2007     2700      1     3.450
## 70845  35848968871 6.256 2007     2700      1     5.199
## 70888  35248896228 3.992 2007     2700      1     2.835
## 70910  35748974550 4.877 2007     2700      1     3.820
## 70923  36048931015 4.213 2007     2700      1     3.156
## 70927  36048938340 3.610 2007     2700      1     2.553
## 70929  36048953291 3.747 2007     2700      1     2.590
## 70932  36048958880 3.498 2007     2700      1     2.646
## 70934  36048963838 4.213 2007     2700      1     3.156
## 70935  36048968753 4.402 2007     2700      1     3.345
## 70950  36049037185 3.835 2007     2700      1     2.778
## 70951  36049042525 3.835 2007     2700      1     2.778
## 70959  36248973101 3.783 2007     2700      1     2.626
## 71265  41149161956 4.099 2007     2700      1     3.147
## 71308  48849113878 4.849 2007     2700      1     3.792
## 71611  35348968910 3.761 2007     1000      2     2.704
## 71623  35348981185 3.708 2007     2700      1     2.856
## 71624  35348997893 4.319 2007     2700      1     3.467
## 71653  35148831175 4.279 2007     2700      1     3.427
## 71667  34948830633 3.376 2007     2700      1     2.524
## 71669  35248829599 4.645 2007     2700      1     3.693
## 71673  35248895346 4.762 2007     2700      1     3.705
## 71681  35348866794 3.930 2007     2700      1     2.773
## 71714  34948849235 3.498 2007     2700      1     2.546
## 71717  34948891641 4.829 2007     2700      1     3.772
## 71720  34948845507 4.340 2007     2700      1     3.283
## 71722  34948882262 4.028 2007     2700      1     2.871
## 71753  34548827323 4.070 2007     2700      1     3.013
## 71755  34548835744 4.466 2007     2700      1     3.409
## 71788  34948860191 3.564 2007     2700      1     2.507
## 71791  34948872683 3.695 2007     2700      1     2.638
## 71809  35348828533 4.122 2007     2700      1     2.965
## 71831  35348985396 3.688 2007     2700      1     2.531
## 71842  35448937632 3.617 2007     2700      1     2.765
## 71847  35648946842 3.807 2007     2700      1     2.955
## 71864  35649003880 4.876 2007     2700      1     3.719
## 72101  40649100791 3.868 2007     2700      1     2.811
## 72327  34648832073 4.784 2007     2700      1     3.727
## 72334  34748875855 4.118 2007     2700      1     3.061
## 72398  34548833077 4.091 2007     2700      1     3.034
```

```
## 72400  34548849816 4.178 2007     2700      1     3.121
## 72454  34548776934 3.753 2007     2700      1     2.696
## 72465  34548570632 4.484 2007     2700      1     3.427
## 72494  34547426346 3.681 2007     2700      1     2.624
## 72517  34548419640 3.595 2007     2700      1     2.743
## 72546  34548418912 4.594 2007     2700      1     3.537
## 72548  34548433987 4.078 2007     2700      1     3.021
## 72555  34548426480 4.019 2007     2700      1     2.862
## 72556  34548430952 4.304 2007     2700      1     3.147
## 72568  34848816502 4.152 2007     2700      1     3.200
## 72569  34848818486 3.879 2007     2700      1     2.822
## 72572  34848831462 4.045 2007     2700      1     2.988
## 72575  34848838902 4.019 2007     2700      1     2.962
## 72578  34848858523 4.244 2007     2700      1     3.187
## 72581  34848881905 4.352 2007     2700      1     3.295
## 72582  34848881907 3.681 2007     2700      1     2.524
## 72583  34848888949 3.954 2007     2700      1     2.897
## 72587  34848901974 4.316 2007     2700      1     3.259
## 72588  34848902937 4.337 2007     2700      1     3.280
## 72591  34848920368 3.801 2007     2700      1     2.644
## 72606  34547502731 3.531 2007     2700      1     2.579
## 72608  34547541570 3.406 2007     2700      1     2.554
## 72609  34547545269 4.719 2007     2700      1     3.767
## 72613  34548040666 3.695 2007     2700      1     2.638
## 72619  34548172901 3.625 2007     2700      1     2.568
## 72623  34548262715 4.493 2007     2700      1     3.336
## 72673  34648836403 4.569 2007     2700      1     3.617
## 72681  34748862577 3.830 2007     2700      1     2.673
## 72807  36048990829 3.632 2007     2700      1     2.575
## 72822  36248973631 3.841 2007     2700      1     2.784
## 73355  34548221486 3.438 2007     1704      3     2.586
## 73492  34548066571 3.759 2007     2700      1     2.702
## 73493  34548075820 5.741 2007     2700      1     4.684
## 73494  34548083742 4.058 2007     2700      1     3.001
## 73500  34548092397 4.740 2007     2700      1     3.683
## 73551  34547950500 4.540 2007     2700      1     3.588
## 73558  34547876797 4.649 2007     2700      1     3.592
## 73559  34547906405 4.663 2007     2700      1     3.606
## 73563  34548361951 3.587 2007     2700      1     2.530
## 73594  34547757915 5.372 2007     2700      1     4.215
## 73599  34547730740 4.006 2007     2700      1     2.949
## 73600  34547755634 3.747 2007     2700      1     2.895
## 73649  34547647628 4.354 2007     2700      1     3.297
## 73681  34547614056 4.149 2007     2700      1     3.092
## 73683  34547630852 3.983 2007     2700      1     3.031
## 73684  34547634884 3.728 2007     2700      1     2.671
## 73685  34547643177 3.734 2007     2700      1     2.882
## 73686  34547652884 4.074 2007     2700      1     2.917
## 73697  34547681974 3.954 2007     2700      1     2.897
## 73699  34547686398 4.247 2007     2700      1     3.090
```

```
## 73747  34548397297 3.789 2007     2700      1     2.732
## 73760  34548456959 3.595 2007     2700      1     2.538
## 74193  34547148985 4.266 2007     2700      1     3.209
## 74194  34547170053 6.085 2007     2700      1     5.028
## 74199  34547230736 4.019 2007     2700      1     2.962
## 74203  34547159840 4.045 2007     2700      1     2.988
## 74204  34547167603 3.695 2007     2700      1     2.638
## 74226  34447521458 4.639 2007     2700      1     3.482
## 74241  34447544253 3.830 2007     2700      1     2.773
## 74246  34447300446 4.669 2007     2700      1     3.612
## 74247  34447512756 4.710 2007     2700      1     3.858
## 74249  34447520136 5.342 2007     2700      1     4.285
## 74250  34447558140 5.428 2007     2700      1     4.371
## 74312  34447340945 3.992 2007     2700      1     2.935
## 74314  34548165110 3.506 2007     2700      1     2.554
## 74322  34447257725 3.920 2007     2700      1     2.763
## 74323  34447291144 4.282 2007     2700      1     3.225
## 74369  34447116875 3.695 2007     2700      1     2.638
## 74375  34447121852 3.852 2007     2700      1     2.795
## 74385  34447137331 4.461 2007     2700      1     3.304
## 74398  34250840592 3.920 2007     2700      1     2.863
## 74704  35248881664 4.126 2007     2700      1     3.069
## 75040  34548567516 3.824 2007     2700      1     2.972
## 75044  34347210370 4.041 2007     2700      1     2.984
## 75045  34347237664 4.019 2007     2700      1     2.962
## 75047  34347255039 4.665 2007     2700      1     3.508
## 75057  34347272329 3.753 2007     2700      1     2.901
## 75085  34250696907 4.873 2007     2700      1     3.921
## 75086  34250722018 4.152 2007     2700      1     3.200
## 75089  34250779802 4.066 2007     2700      1     3.009
## 75092  34250835050 3.709 2007     1000      2     2.652
## 75142  34250212715 6.307 2007     2700      1     5.455
## 75144  34250309164 4.292 2007     2700      1     3.235
## 75168  34249899377 4.920 2007     2700      1     3.763
## 75183  34250024383 3.783 2007     2700      1     2.831
## 75202  34249982970 4.196 2007     2700      1     3.139
## 75203  34250027957 3.983 2007     2700      1     2.826
## 75236  34248149838 3.765 2007     2700      1     2.708
## 75239  34248169298 3.795 2007     2700      1     2.738
## 75242  34248201151 4.476 2007     2700      1     3.419
## 75243  34248206205 3.688 2007     2700      1     2.631
## 75256  34249037494 3.356 2007     2700      1     2.504
## 75260  34249662628 3.925 2007     2700      1     2.868
## 75262  34249863368 3.852 2007     2700      1     3.000
## 75274  34249913494 4.647 2007     2700      1     3.490
## 75275  34249915334 4.224 2007     2700      1     3.167
## 75382  34447648357 3.807 2007     2700      1     2.855
## 75793  34547309949 3.854 2007     2700      2     2.797
## 75956  34249302025 3.580 2007     2700      1     2.523
## 75959  34249803312 4.122 2007     2700      1     3.270
```

```
## 75961  34249830160 4.754 2007     2700      1     3.902
## 75973  34249027978 3.959 2007     2700      1     2.902
## 75990  34249657825 4.383 2007     2700      1     3.326
## 75993  34249673868 4.807 2007     2700      1     3.750
## 76036  34249047454 4.895 2007     2700      1     3.838
## 76042  34248551882 4.301 2007     2700      1     3.144
## 76043  34248559834 4.710 2007     2700      1     3.653
## 76044  34248564820 3.531 2007     2700      1     2.679
## 76093  34248399089 4.292 2007     2700      1     3.235
## 76121  34247615981 3.835 2007     2700      1     2.778
## 76133  34247862190 5.281 2007     2700      1     4.224
## 76164  34247268567 3.610 2007     2700      1     2.553
## 76199  34248184445 3.702 2007     2700      1     2.645
## 76200  34248185926 4.203 2007     2700      1     3.046
## 76214  34248214561 3.765 2007     2700      1     2.913
## 76218  34248227278 3.734 2007     2700      1     2.577
## 76835  34247863975 5.154 2007     2700      1     3.997
## 76849  34247476744 5.028 2007     2700      1     3.971
## 76864  34247470278 4.213 2007     2700      1     3.056
## 76865  34247498668 4.800 2007     2700      1     3.643
## 76883  34247167109 3.675 2007     2700      1     2.518
## 76903  34247241630 5.117 2007     2700      1     4.060
## 76945  34247118878 3.813 2007     2700      1     2.961
## 76946  34247144499 4.185 2007     2700      1     3.128
## 76947  34247149303 4.041 2007     2700      1     2.984
## 76950  34247523621 4.713 2007     2700      1     3.656
## 76959  34047096257 4.791 2007     2700      1     3.939
## 76970  34047243811 4.396 2007     2700      1     3.544
## 76972  34047268898 3.900 2007     2700      1     2.843
## 76981  34047237367 5.526 2007     2700      1     4.674
## 76988  34147096980 4.461 2007     2700      1     3.404
## 76993  34147124264 4.538 2007     2700      1     3.381
## 76995  34147139988 3.610 2007     2700      1     2.553
## 76998  34147153376 3.564 2007     2700      1     2.507
## 76999  34147153377 4.107 2007     2700      1     3.050
## 77015  33847667949 4.049 2007     2700      1     3.197
## 77028  33947671432 3.747 2007     2700      1     2.690
## 77077  34247476430 3.747 2007     2700      1     2.590
## 77081  34247498181 3.668 2007     2700      1     2.511
## 77084  34247545653 3.639 2007     2700      1     2.687
## 77570  34247282706 4.091 2007     2700      1     3.139
## 77574  34147113263 3.593 2007     1000      2     2.536
## 77582  34047185285 3.728 2007     2700      1     2.671
## 77604  33947210018 3.884 2007     2700      1     2.827
## 77607  33947595236 5.510 2007     2700      1     4.558
## 77643  33947519044 3.632 2007     2700      1     2.680
## 77661  33947196096 4.621 2007     2700      1     3.564
## 77664  33947279831 4.110 2007     2700      1     3.158
## 77685  33947271991 3.734 2007     2700      1     2.677
## 77712  33847705701 5.514 2007     2700      1     4.357
```

```
## 77720  33847723397 4.950 2007     2700      1     4.098
## 77730  33847393654 3.396 2007     2700      1     2.544
## 77752  33846926962 3.715 2007     2700      1     2.763
## 77761  33847253945 4.880 2007     2700      1     3.723
## 77762  33847267289 3.396 2007     2700      1     2.544
## 77772  33847368173 5.051 2007     2700      1     3.994
## 77773  33847369469 4.655 2007     2700      1     3.598
## 77776  33847381116 5.213 2007     2700      1     4.156
## 77777  33847391669 4.160 2007     2700      1     3.103
## 77779  33847404482 4.431 2007     2700      1     3.374
## 77780  33847406983 3.747 2007     2700      1     2.590
## 77781  33847407101 3.987 2007     2700      1     2.930
## 77786  33847413701 3.920 2007     2700      1     2.863
## 77819  33947356563 3.540 2007     2700      1     2.688
## 77839  33947710793 4.749 2007     2700      1     3.692
## 78473  34547138943 3.721 2007     2700      2     2.664
## 78607  33847353773 4.810 2007     2700      1     3.858
## 78631  33847692398 3.435 2007     2700      1     2.583
## 78651  33847165011 4.126 2007     2700      1     3.069
## 78654  33847103656 4.307 2007     2700      1     3.455
## 78655  33847174197 4.203 2007     2700      1     3.351
## 78704  33846964515 3.819 2007     2700      1     2.762
## 78707  33846996345 5.465 2007     2700      1     4.613
## 78728  33846857305 3.602 2007     2700      1     2.545
## 78737  33846866271 4.955 2007     2700      1     3.898
## 78738  33846875851 5.005 2007     2700      1     3.948
## 78746  33846863146 4.895 2007     2700      1     3.838
## 78749  33846878293 3.715 2007     2700      1     2.558
## 78780  33846571916 4.217 2007     2700      1     3.160
## 78792  33846075235 4.032 2007     2700      1     2.975
## 78810  33846661903 3.959 2007     2700      1     2.902
## 78812  33846673816 4.935 2007     2700      1     3.878
## 78813  33846688887 5.379 2007     2700      1     4.222
## 78816  33846783284 3.715 2007     2700      1     2.658
## 78821  33846787125 3.835 2007     2700      1     2.678
## 78822  33846794896 5.202 2007     2700      1     4.145
## 78825  33846810181 3.564 2007     2700      1     2.507
## 78832  33846817495 4.130 2007     2700      1     3.278
## 78836  33846829811 3.983 2007     2700      1     3.131
## 78839  33846839628 3.915 2007     2700      1     2.758
## 78843  33846850600 3.777 2007     2700      1     2.620
## 78847  33846913993 3.531 2007     2700      1     2.679
## 78899  33847610733 4.122 2007     2700      1     3.065
## 78900  33847617509 3.987 2007     2700      1     2.830
## 79061  34248581736 3.523 2007     2700      1     2.571
## 79359  33846524262 3.396 2007     2700      1     2.544
## 79377  33846294746 5.096 2007     2700      1     4.039
## 79386  33846413027 3.572 2007     2700      1     2.515
## 79416  33846126939 4.946 2007     2700      1     3.789
## 79420  33846175081 3.777 2007     2700      1     2.720
```

```
## 79427  33846125284 5.173 2007     2700      1     4.116
## 79428  33846129459 4.066 2007     2700      1     3.114
## 79489  33845970192 5.126 2007     2700      1     4.069
## 79491  33845972969 3.789 2007     2700      1     2.632
## 79492  33845979811 3.889 2007     2700      1     2.732
## 79495  33845994359 4.653 2007     2700      1     3.496
## 79496  33846006533 3.661 2007     2700      1     2.504
## 79498  33846012847 3.841 2007     2700      1     2.684
## 79499  33846017279 4.078 2007     2700      1     2.921
## 79510  33845984538 4.757 2007     2700      1     3.700
## 79511  33846021262 4.182 2007     2700      1     3.125
## 79512  33846026341 4.599 2007     2700      1     3.542
## 79513  33846027094 4.070 2007     2700      1     2.913
## 79540  33845870422 3.734 2007     2700      1     2.882
## 79560  33846037400 3.935 2007     2700      1     2.983
## 79602  33846668526 3.874 2007     2700      1     2.817
## 79607  33846671273 3.654 2007     2700      1     2.597
## 79609  33846672853 4.091 2007     2700      1     3.034
## 79671  33847184284 3.632 2007     2700      1     2.680
## 79978  34250801440 3.925 2007     2700      1     3.073
## 80011  34347336319 3.646 2007     2700      1     2.589
## 80013  34347343739 3.668 2007     2700      1     2.611
## 80106  35649016098 4.629 2007     2700      1     3.572
## 80111  35848953625 3.728 2007     2700      1     2.671
## 80112  35848954757 4.777 2007     2700      1     3.825
## 80217  42449096529 3.935 2007     2700      1     3.083
## 80636  58149083708 4.778 2008     2700      1     3.724
## 80643  58149218477 3.923 2008     2700      1     3.073
## 80683  58049213696 4.583 2008     2700      1     3.529
## 80684  58049216683 3.839 2008     2700      1     2.785
## 80728  57349192265 3.722 2008     2700      1     2.668
## 80744  57449112519 4.616 2008     2700      1     3.562
## 80797  57349093766 3.663 2008     2700      1     2.609
## 80814  57349133922 3.823 2008     2700      1     2.769
## 80816  57349184260 3.522 2008     2700      1     2.672
## 80943  56149095348 3.843 2008     2700      1     2.893
## 80980  57449084208 4.389 2008     2700      1     3.335
## 80983  57449095883 3.663 2008     2700      1     2.609
## 81062  58149387739 4.794 2008     2700      1     3.740
## 81349  64349102964 3.423 2008     2700      1     2.573
## 82714  57049128384 3.717 2008     2700      1     2.563
## 82785  56649112752 4.452 2008     2700      1     3.398
## 82845  56249128276 4.627 2008     2700      1     3.573
## 82846  56249141853 3.916 2008     2700      1     2.862
## 82880  55849115999 3.659 2008     1000      2     2.605
## 82899  55549095137 3.894 2008     2700      1     2.840
## 82911  55849084700 3.909 2008     2700      1     2.855
## 82920  55849139852 3.698 2008     2700      1     2.748
## 82950  54049103289 3.570 2008     2700      1     2.516
## 83109  58149467012 3.754 2008     2700      1     2.700
```

```
## 83609  55249119765 4.514 2008     2700      1     3.460
## 83619  54349091094 3.478 2008     2700      1     2.528
## 83648  55549139481 4.069 2008     2700      1     3.015
## 83661  54949144335 4.216 2008     2700      1     3.162
## 83714  54049121079 3.886 2008     2700      1     2.832
## 83728  54049104424 3.768 2008     2700      1     2.818
## 83729  54049106963 4.294 2008     2700      1     3.240
## 83794  53749095835 4.145 2008     2700      1     3.091
## 83849  54849146500 4.137 2008     2700      1     3.083
## 83859  54949141082 3.589 2008     1704      3     2.535
## 83861  50249096174 4.078 2008     2700      1     3.128
## 83872  51749084267 3.451 2008     2700      1     2.601
## 83891  52949103244 3.781 2008     2700      1     2.831
## 83892  52949106040 4.278 2008     2700      1     3.224
## 83897  52949144507 4.137 2008     2700      1     3.083
## 83898  52949144863 3.722 2008     2700      1     2.772
## 83901  53149119591 3.401 2008     2700      1     2.551
## 83969  55849112296 3.599 2008     2700      1     2.649
## 83975  55849150592 4.342 2008     2700      1     3.288
## 84213  65549087532 3.839 2008     2700      1     2.889
## 84545  52649151888 3.916 2008     2700      1     2.966
## 84550  52049088955 4.048 2008     2700      1     2.894
## 84553  52049108262 4.242 2008     2700      1     3.188
## 84555  52449113766 3.626 2008     2700      1     2.776
## 84585  52649120886 4.627 2008     2700      1     3.473
## 84598  52249089796 3.843 2008     2700      1     2.789
## 84607  51949095058 4.868 2008     2700      1     3.918
## 84682  51449097277 4.098 2008     2700      1     2.944
## 84728  50949089029 4.489 2008     2700      1     3.435
## 84734  51649088233 4.453 2008     2700      1     3.399
## 84761  51349169034 3.673 2008     2700      1     2.619
## 84831  52749095854 3.593 2008     2700      1     2.539
## 85169  59049103004 3.503 2008     2700      1     2.653
## 85277  65249106599 3.570 2008     2700      1     2.620
## 85688  50149098605 3.746 2008     1000      2     2.692
## 85697  50449091665 3.772 2008     2700      1     2.618
## 85698  50449093260 4.360 2008     2700      1     3.306
## 85709  50449095362 4.693 2008     2700      1     3.639
## 85728  49749143019 3.576 2008     2700      1     2.726
## 85743  50149121231 4.030 2008     1000      2     2.976
## 85752  49949087905 4.118 2008     2700      1     3.168
## 85753  49949091204 3.678 2008     2700      1     2.624
## 85765  49449098216 3.958 2008     2700      1     3.008
## 85828  49449100639 4.042 2008     2700      1     2.988
## 85831  49249087104 4.356 2008     2700      1     3.406
## 85886  49249104701 5.150 2008     2700      1     4.096
## 85962  48149090717 3.855 2008     2700      1     2.801
## 85972  48149110000 3.647 2008     2700      1     2.797
## 86469  47849130092 4.213 2008     2700      1     3.159
## 86470  47849130608 3.731 2008     2700      1     2.677
```

```
## 86483  47849100908 3.663 2008     2700      1     2.509
## 86505  47949092540 4.171 2008     2700      1     3.221
## 86507  47949114668 5.232 2008     2700      1     4.078
## 86564  47549096598 3.768 2008     2700      1     2.818
## 86565  47549105407 4.511 2008     2700      1     3.357
## 86620  46949093558 3.886 2008     2700      1     2.732
## 86634  47049126504 3.827 2008     2700      1     2.773
## 86680  46449113799 4.877 2008     2700      1     3.723
## 86684  46449128469 3.546 2008     1000      2     2.596
## 86729  45749146166 3.855 2008     2700      1     3.005
## 86742  46449098821 3.879 2008     2700      1     2.929
## 86743  46749084688 3.843 2008     2700      1     2.789
## 86746  46749099010 4.045 2008     2700      1     3.195
## 86749  46749117203 3.712 2008     2700      1     2.658
## 86755  46749156706 4.460 2008     2700      1     3.306
## 86780  47949131252 3.387 2008     2700      1     2.537
## 86795  48749129542 4.693 2008     2700      1     3.743
## 87045  55149104425 3.768 2008     2700      1     2.818
## 87540  45949092534 4.637 2008     2700      1     3.583
## 87543  45949107235 3.831 2008     2700      1     2.777
## 87544  46749131798 5.204 2008     2700      1     4.150
## 87592  45549084293 4.115 2008     2700      1     3.061
## 87602  45349084317 4.501 2008     2700      1     3.447
## 87603  45349103702 4.595 2008     2700      1     3.441
## 87663  45149127545 4.115 2008     2700      1     3.061
## 87691  44849103815 4.757 2008     2700      1     3.907
## 87696  44849115945 4.501 2008     2700      1     3.447
## 87735  45749100010 3.951 2008     2700      1     2.897
## 87739  45749114895 4.710 2008     2700      1     3.556
## 87740  45749124735 3.409 2008     2700      1     2.559
## 87745  45749149476 3.642 2008     2700      1     2.588
## 87792  44949119632 3.451 2008     2700      1     2.601
## 87828  46349104324 3.831 2008     2700      1     2.777
## 88491  44349145194 4.455 2008     2700      1     3.401
## 88496  44449106055 3.641 2008     1000      2     2.587
## 88499  44349153009 3.978 2008     2700      1     2.924
## 88500  44349155992 4.235 2008     2700      1     3.285
## 88501  44349165737 5.327 2008     2700      1     4.377
## 88546  44249122799 5.169 2008     2700      1     4.115
## 88634  43549085056 3.948 2008     2700      1     3.098
## 88637  43549110959 4.509 2008     2700      1     3.355
## 88680  43049170245 4.574 2008     2700      1     3.624
## 88682  43049174930 3.968 2008     2700      1     2.814
## 88683  43049176549 4.334 2008     2700      1     3.384
## 88750  42949149810 3.901 2008     2700      1     3.051
## 88757  43049090817 4.174 2008     2700      1     3.120
## 88759  43049094836 3.740 2008     2700      1     2.790
## 88760  43049095682 4.414 2008     2700      1     3.360
## 88784  43049136576 3.570 2008     2700      1     2.516
## 88857  44449096577 3.731 2008     2700      1     2.677
```

```
## 88859  44449122269 3.944 2008     2700      1     3.094
## 89498  42449106956 4.533 2008     2700      1     3.683
## 89500  42449141936 3.802 2008     2700      1     2.648
## 89501  44149123099 4.666 2008     2700      1     3.816
## 89566  42249112261 3.621 2008     1000      2     2.567
## 89567  42249098105 4.181 2008     2700      1     3.127
## 89601  42249083269 4.686 2008     2700      1     3.836
## 89602  42249092533 4.619 2008     2700      1     3.769
## 89605  48149094782 3.698 2008     2700      1     2.544
## 89617  42049123233 4.048 2008     2700      1     3.198
## 89645  40849108663 3.574 2008     1000      2     2.520
## 89663  41649112174 3.663 2008     2700      1     2.509
## 89739  41849105806 3.930 2008     2700      1     2.776
## 89751  41849148289 3.992 2008     2700      1     2.938
## 89783  43249086840 3.683 2008     2700      1     2.529
## 89784  43249087616 3.859 2008     2700      1     2.805
## 89878  46349092759 3.712 2008     2700      1     2.558
## 90277  43049134750 3.647 2008     2700      1     2.593
## 90284  41449104685 3.576 2008     2700      1     2.522
## 90287  41449112372 5.371 2008     2700      1     4.317
## 90289  41449113183 4.795 2008     2700      1     3.741
## 90338  40949100835 3.944 2008     2700      1     2.994
## 90347  40949102607 5.086 2008     2700      1     3.932
## 90361  40949127393 4.633 2008     2700      1     3.783
## 90407  40849124009 4.384 2008     2700      1     3.330
## 90409  40849142102 4.916 2008     2700      1     3.762
## 90412  40749132230 3.356 2008     1000      2     2.506
## 90415  40749141596 4.476 2008     2700      1     3.526
## 90432  41649090186 4.800 2008     2700      1     3.746
## 90458  40449141326 4.620 2008     2700      1     3.566
## 90464  40449104676 4.174 2008     2700      1     3.324
## 90509  39749103691 3.615 2008     2700      1     2.765
## 90521  40449086087 4.445 2008     2700      1     3.291
## 90523  40549084318 3.879 2008     2700      1     2.825
## 90545  40749105777 3.541 2008     2700      1     2.691
## 91382  39549099909 3.693 2008     2700      1     2.639
## 91383  39549100109 3.988 2008     2700      1     2.934
## 91384  39549102834 4.683 2008     2700      1     3.629
## 91386  39549111874 3.958 2008     2700      1     2.904
## 91388  39549096658 4.328 2008     2700      1     3.174
## 91415  39749102149 4.416 2008     2700      1     3.566
## 91426  39349097864 3.575 2008     1000      2     2.521
## 91488  38949105879 4.256 2008     2700      1     3.306
## 91489  38949106566 3.754 2008     2700      1     2.700
## 91490  38949188680 3.871 2008     2700      1     2.817
## 91499  38949152338 4.285 2008     2700      1     3.231
## 91520  38649125875 4.145 2008     2700      1     3.091
## 91534  38649100024 3.604 2008     2700      1     2.754
## 91552  38849130851 3.827 2008     2700      1     2.773
## 91565  38849174979 4.368 2008     2700      1     3.418
```

```
## 91567  38849182472 3.972 2008     2700      1     2.918
## 91602  39849097577 4.313 2008     2700      1     3.159
## 91603  39849099704 3.516 2008     2700      1     2.666
## 91605  39849102541 3.615 2008     2700      1     2.665
## 91725  43349085680 3.955 2008     2700      1     2.901
## 91726  43349087466 3.772 2008     2700      1     2.718
## 91730  43349096215 3.707 2008     2700      1     2.653
## 91738  43349107306 3.975 2008     2700      1     2.921
## 91845  38949168818 3.621 2008     1300      2     2.771
## 91993  38749111066 5.445 2008     2700      1     4.391
## 91995  38749127156 4.616 2008     2700      1     3.562
## 92001  38749133726 3.560 2008     1000      2     2.506
## 92033  38549088096 4.842 2008     2700      1     3.788
## 92037  38549134052 4.324 2008     2700      1     3.270
## 92040  38549177642 4.453 2008     2700      1     3.399
## 92046  38349164176 4.531 2008     2700      1     3.477
## 92068  38349049478 5.357 2008     2700      1     4.303
## 92111  38049146378 4.148 2008     2700      1     3.094
## 92113  38049169559 4.935 2008     2700      1     3.881
## 92134  38049062200 4.629 2008     2700      1     3.575
## 92136  38049077991 3.843 2008     2700      1     2.789
## 92140  38049082149 4.298 2008     2700      1     3.244
## 92148  38149064606 3.631 2008     2700      1     2.681
## 92161  38149141767 3.958 2008     2700      1     2.804
## 92193  37549009899 4.066 2008     2700      1     3.116
## 92197  37549020305 3.582 2008     2700      1     2.528
## 92199  37549027612 4.244 2008     2700      1     3.190
## 92203  37549064327 3.776 2008     2700      1     2.722
## 92219  38149074110 3.934 2008     2700      1     2.880
## 92231  38349085066 3.610 2008     2700      1     2.556
## 92494  41149157013 3.547 2008     2700      1     2.597
## 92576  42149177179 4.216 2008     2700      1     3.162
## 92681  44949253343 4.140 2008     2700      1     2.986
## 92735  46749123025 3.731 2008     2700      1     2.677
## 92922  57149102701 3.912 2008     2700      1     3.062
## 93538  74049108635 3.752 2009     2700      1     2.653
## 93541  73949086506 3.919 2009     2700      1     2.721
## 93592  73549083853 3.805 2009     2700      1     2.706
## 93602  77954601328 4.143 2009     2700      1     3.044
## 93632  70350733579 3.757 2009     2700      1     2.559
## 93633  70449636557 4.842 2009     2700      1     3.644
## 93634  71849097187 3.961 2009     2700      1     2.763
## 93635  71849098875 3.947 2009     2700      1     2.848
## 93652  70350733425 3.985 2009     2700      1     2.991
## 93654  70350738289 4.729 2009     2700      1     3.630
## 93659  72549091935 4.039 2009     2700      1     3.145
## 93707  71849118976 3.968 2009     2700      1     2.869
## 93720  71549150905 3.500 2009     2700      1     2.606
## 93760  73349094732 4.184 2009     2700      1     3.190
## 93764  73349134686 4.036 2009     2700      1     2.937
```

```
## 93775  84870055817 4.573 2009     2700      1     3.679
## 93843  65549167028 3.801 2009     2700      1     2.702
## 93889  67650760521 3.826 2009     2700      1     2.727
## 93962  70749114018 3.645 2009     2700      1     2.546
## 93964  70749145982 3.814 2009     2700      1     2.715
## 93975  72049129932 3.889 2009     2700      1     2.790
## 93995  72849125358 3.858 2009     2700      1     2.864
## 94224  77049105843 3.904 2009     2700      1     2.910
## 94281  77649213935 3.838 2009     2700      1     2.844
## 95604  72049129430 4.683 2009     2700      1     3.584
## 95638  70449652591 3.805 2009     2700      1     2.706
## 95723  70349873173 5.082 2009     2700      1     4.088
## 95726  70449636163 5.302 2009     2700      1     4.308
## 95743  77950669604 5.246 2009     2700      1     4.147
## 95828  70449355084 3.784 2009     2700      1     2.685
## 95831  70449377960 3.743 2009     2700      1     2.644
## 95834  70449394580 4.039 2009     2700      1     2.940
## 95842  70449448312 4.397 2009     2700      1     3.298
## 96014  73949153085 4.118 2009     2700      1     3.019
## 96406  73349109845 3.667 2009     2700      2     2.568
## 96572  84881264341 4.415 2009     2700      1     3.316
## 96673  70350064091 3.705 2009     2700      1     2.711
## 96696  70350458703 4.870 2009     2700      1     3.876
## 96716  70350212398 4.751 2009     2700      1     3.652
## 96717  70350220313 3.629 2009     2700      1     2.530
## 96718  70350227015 3.975 2009     2700      1     2.777
## 96719  70350236228 4.659 2009     2700      1     3.765
## 96753  70349506141 4.798 2009     2700      1     3.904
## 96773  70349943834 3.625 2009     1000      2     2.526
## 96854  65749111503 3.762 2009     2700      1     2.663
## 96867  69949132983 3.549 2009     2700      1     2.655
## 96870  69949144253 3.696 2009     2700      1     2.597
## 96895  70349481938 4.575 2009     2700      1     3.476
## 96897  70349610473 4.853 2009     2700      1     3.754
## 96899  70349625757 5.493 2009     2700      1     4.394
## 96900  70349629019 5.086 2009     2700      1     3.987
## 96919  70349684791 4.187 2009     2700      1     2.989
## 96920  70349687102 4.017 2009     2700      1     2.918
## 96953  70350455774 3.691 2009     2700      1     2.697
## 97145  77952988108 4.600 2009     2700      1     3.501
## 97523  70349437154 3.951 2009     2700      1     2.852
## 97545  70349308645 3.601 2009     2700      1     2.502
## 97546  70349308646 4.248 2009     2700      1     3.149
## 97548  70349319816 3.705 2009     2700      1     2.711
## 97550  70349335796 4.765 2009     2700      1     3.666
## 97633  70349238733 4.938 2009     2700      1     4.044
## 97634  70349244812 4.857 2009     2700      1     3.963
## 97636  70349254065 3.850 2009     2700      1     2.751
## 97666  66849121699 3.710 2009     2700      1     2.611
## 97668  66849128097 3.947 2009     2700      1     2.848
```

```
## 97740  70149093912 5.343 2009     2700      1     4.145
## 97741  70149098511 3.748 2009     2700      1     2.649
## 97748  70249111091 3.839 2009     1000      2     2.641
## 97755  70149112350 4.282 2009     2700      1     3.084
## 97852  68949150765 3.725 2009     2700      1     2.731
## 97861  69849097352 3.958 2009     2700      1     2.859
## 97897  70349205537 3.506 2009     2700      1     2.512
## 97906  70349213381 3.691 2009     2700      1     2.797
## 97912  70349220917 3.757 2009     2700      1     2.658
## 97930  70349636711 4.014 2009     2700      1     2.915
## 98067  73449108112 3.846 2009     2700      1     2.852
## 98073  73449123799 4.061 2009     2700      1     2.962
## 98210  75349114036 3.766 2009     2700      1     2.667
## 98756  69249219296 5.111 2009     2700      1     4.012
## 98759  69249229866 3.850 2009     2700      1     2.751
## 98778  69449098588 4.214 2009     2700      1     3.115
## 98813  68949090752 4.850 2009     2700      1     3.652
## 98814  68949092632 4.197 2009     2700      1     3.303
## 98815  68949102467 3.830 2009     2700      1     2.731
## 98816  68949133346 4.250 2009     2700      1     3.052
## 98817  68949171916 3.775 2009     2700      1     2.881
## 98921  68849084949 4.231 2009     2700      1     3.033
## 98930  68849119553 4.573 2009     2700      1     3.474
## 98931  68849123400 4.335 2009     2700      1     3.236
## 98992  68249154875 5.117 2009     2700      1     4.018
## 99008  68349117223 3.919 2009     2700      1     3.025
## 99032  68849104003 3.613 2009     2700      1     2.514
## 99033  68849112640 3.607 2009     2700      1     2.508
## 99081  67650635227 3.991 2009     2700      1     2.793
## 99104  67651160419 4.452 2009     2700      1     3.558
## 99146  69349094143 4.819 2009     2700      1     3.720
## 99564  68049123594 4.967 2009     2700      1     3.868
## 99616  68149144251 3.826 2009     2700      1     2.727
## 99644  67651218924 4.011 2009     2700      1     2.912
## 99664  67650391945 3.818 2009     2700      1     2.620
## 99679  67650318719 3.926 2009     2700      1     2.728
## 99686  67650899270 4.135 2009     2700      1     3.036
## 99688  67650909125 3.968 2009     2700      1     2.869
## 99748  68149112408 3.440 2009     2700      1     2.546
## 99827  67649664439 3.975 2009     2700      1     3.081
## 99858  67649814915 3.743 2009     2700      1     2.545
## 99874  67649967683 4.202 2009     2700      1     3.308
## 99913  68049122014 3.725 2009     2700      1     2.831
## 99961  68849122050 3.988 2009     2700      1     2.889
## 99962  68849124811 4.132 2009     2700      1     3.238
## 100692 67649390890 3.995 2009     2700      1     2.896
## 100705 67649218781 4.174 2009     2700      1     2.976
## 100727 67149116154 4.250 2009     2700      1     3.052
## 100761 67449110743 5.734 2009     2700      1     4.536
## 100762 67449152290 4.549 2009     2700      1     3.655
```

```
## 100764 67449164373 3.944 2009     2700      1     2.950
## 100831 67149142042 5.168 2009     2700      1     4.069
## 100844 67049116139 3.882 2009     2700      1     2.783
## 100938 66649113371 4.055 2009     2700      1     2.857
## 100947 66449100700 4.172 2009     2700      1     2.974
## 101089 67651005390 3.748 2009     2700      1     2.754
## 101093 67651006589 3.433 2009     2700      1     2.539
## 101786 67650169237 4.101 2009     2700      1     3.002
## 101788 65849187703 5.312 2009     2700      1     4.213
## 101794 66249084112 4.991 2009     2700      1     3.892
## 101808 66149152644 4.179 2009     2700      1     2.981
## 101854 66149118225 3.814 2009     2700      1     2.715
## 101861 66249137759 3.789 2009     1000      2     2.690
## 101904 66149123748 3.968 2009     1000      2     2.869
## 101942 65549163328 4.148 2009     2700      1     3.154
## 101994 65449136284 3.759 2009     1303      7     2.660
## 102096 66449111715 3.766 2009     2700      1     2.667
## 102501 65549092261 3.656 2009     1000      2     2.557
## 102867 65449152185 5.067 2009     2700      1     3.968
## 102988 64749093541 4.983 2009     2700      1     3.884
## 103006 64749086505 4.036 2009     2700      1     2.937
## 103008 64749114159 4.724 2009     2700      1     3.526
## 103061 62549134121 3.909 2009     1000      2     2.810
## 103846 63249124498 3.838 2009     2700      1     2.739
## 103851 63349094964 3.846 2009     2700      1     2.952
## 103903 62749097289 4.481 2009     2700      1     3.587
## 103904 62749131642 4.113 2009     2700      1     2.915
## 103921 62649136134 4.286 2009     2700      1     3.292
## 103922 62649155822 4.030 2009     2700      1     3.036
## 103984 61849137328 3.830 2009     2700      1     2.731
## 103998 62349094336 3.870 2009     2700      1     2.771
## 104003 62249133709 4.143 2009     1000      2     3.044
## 104012 62149120632 4.393 2009     2700      1     3.399
## 104068 61849118968 4.245 2009     2700      1     3.251
## 104126 60049098815 4.857 2009     2700      1     3.758
## 104132 61549094077 5.418 2009     2700      1     4.524
## 104148 62449262646 3.640 2009     2700      1     2.646
## 105045 60649118584 4.654 2009     2700      1     3.660
## 105055 60649090430 4.316 2009     2700      1     3.217
## 105061 61449183682 5.109 2009     2700      1     4.010
## 105074 61449162171 3.752 2009     2700      1     2.653
## 105118 60849115270 5.804 2009     2700      1     4.705
## 105126 60749089639 3.858 2009     2700      1     2.759
## 105140 59849099670 3.904 2009     2700      1     3.010
## 105210 60249096283 3.797 2009     2700      1     2.803
## 105211 60249099307 4.478 2009     2700      1     3.280
## 105240 59249084303 4.172 2009     2700      1     3.073
## 105242 59449101337 4.234 2009     2700      1     3.036
## 105251 59749096464 4.506 2009     2700      1     3.407
## 105271 58949085856 3.525 2009     2700      1     2.631
```

```
## 105282 61749094289 4.261 2009     2700      1     3.063
## 105283 61749094852 3.775 2009     2700      1     2.577
## 105852 59449094890 4.255 2009     2700      1     3.261
## 105884 58749094444 5.536 2009     2700      1     4.338
## 105903 58749109459 3.607 2009     2700      1     2.613
## 105947 58349091054 3.908 2009     2700      1     2.710
## 105956 58249110796 3.725 2009     1000      2     2.527
## 105970 58249088833 3.991 2009     2700      1     2.793
## 105990 58149385407 4.564 2009     2700      1     3.465
## 105992 58149389215 5.900 2009     2700      1     4.801
## 106003 58149385649 4.940 2009     2700      1     3.742
## 106031 61749090233 4.005 2009     2700      1     2.906
## 106037 61749102618 4.145 2009     2700      1     3.046
## 106063 57749092004 3.743 2009     2700      1     2.644
## 106066 57749093663 3.623 2009     2700      1     2.524
## 106107 58249087710 3.985 2009     2700      1     2.886
## 106187 61449122079 3.671 2009     2700      1     2.777
## 106195 61449149483 4.101 2009     2700      1     3.002
## 107838 78650858488 4.531 2010     2700      1     3.388
## 107864 78650664067 3.685 2010     2700      1     2.542
## 107918 78649906060 4.743 2010     2700      1     3.600
## 108011 78649880804 4.124 2010     2700      1     2.981
## 108234 78649822277 3.520 2010     2700      1     2.581
## 108250 78649888666 3.441 2010     2700      1     2.502
## 108252 78649895980 3.904 2010     2700      1     2.761
## 109519 78649429727 4.440 2010     2700      1     3.297
## 109522 78649439268 4.337 2010     2700      1     3.194
## 109593 78549288927 3.760 2010     2700      1     2.821
## 109846 78149330579 3.601 2010     2700      1     2.563
## 109891 78049510428 4.327 2010     2700      1     3.084
## 109892 78049515807 4.522 2010     2700      1     3.379
## 109919 78049407029 4.053 2010     2700      1     2.910
## 110787 77957768543 4.019 2010     2700      1     2.876
## 110790 78049428879 4.101 2010     2700      1     2.958
## 110803 78049337151 3.858 2010     2700      1     2.715
## 110805 78049351929 4.511 2010     2700      1     3.368
## 110864 77958149840 4.229 2010     2700      1     2.986
## 111024 77957934893 3.955 2010     2700      1     2.712
## 111045 77957928301 3.775 2010     2700      1     2.632
## 111115 77957725001 3.841 2010     2700      1     2.698
## 111130 77957671700 3.646 2010     2700      1     2.503
## 111485 80051785529 3.594 2010     2700      1     2.556
## 112042 77957333252 3.808 2010     2700      1     2.665
## 112043 77957341503 4.430 2010     2700      1     3.187
## 112124 77957139539 3.680 2010     1000      2     2.537
## 112173 77956929653 3.485 2010     2700      1     2.546
## 112174 77956942602 3.749 2010     2700      1     2.506
## 112259 77956634601 3.549 2010     2700      1     2.610
## 113861 77955877759 5.447 2010     2700      1     4.204
## 113873 77955760684 3.730 2010     2700      1     2.587
```

```
## 114067 77955357246 3.858 2010     2700      1     2.715
## 114203 77955296243 3.591 2010     2700      1     2.553
## 114219 77955286305 3.791 2010     2700      1     2.648
## 114324 77955573674 4.157 2010     2700      1     3.218
## 114349 77956057548 3.989 2010     2700      1     2.846
## 115221 77955066199 5.339 2010     2700      1     4.196
## 115230 77954988618 4.006 2010     2700      1     2.863
## 115361 77954635020 4.231 2010     2700      1     3.088
## 115455 77954421915 4.566 2010     2700      1     3.423
## 115465 77954374543 4.143 2010     2700      1     2.900
## 115552 77954183132 3.685 2010     2700      1     2.542
## 115604 77954974268 3.717 2010     2700      1     2.574
## 116866 77953956205 4.405 2010     2700      1     3.367
## 117020 77953577951 4.518 2010     2700      1     3.375
## 117076 77952962769 3.929 2010     2700      1     2.686
## 117084 77952977625 3.827 2010     2700      1     2.684
## 117093 77952998648 3.957 2010     2700      1     3.018
## 117101 77953013128 3.801 2010     2700      1     2.658
## 117127 77953173228 4.019 2010     2700      1     2.981
## 117148 77952995935 3.649 2010     2700      1     2.506
## 117177 77952766987 3.851 2010     2700      1     2.708
## 117245 77953360245 3.933 2010     2700      1     2.994
## 118168 77952757705 4.297 2010     2700      1     3.358
## 118169 77952760196 3.757 2010     2700      1     2.514
## 118172 77952784840 4.247 2010     2700      1     3.308
## 118242 77952470787 3.736 2010     2700      1     2.797
## 118247 77952356171 3.714 2010     2700      1     2.571
## 118249 77952389110 3.851 2010     2700      1     2.708
## 118307 77952301687 4.418 2010     2700      1     3.275
## 118351 77951868945 4.370 2010     2700      1     3.227
## 118352 77951874018 4.849 2010     2700      1     3.910
## 118367 77951894040 4.247 2010     2700      1     3.308
## 118392 77951589703 4.220 2010     2700      1     3.077
## 118410 77951861324 3.897 2010     2700      1     2.754
## 119263 77951639385 4.329 2010     2700      1     3.086
## 119325 77951179995 4.040 2010     2700      1     3.101
## 119326 77951195763 4.197 2010     2700      1     3.159
## 119394 77950894766 3.527 2010     2700      1     2.588
## 119487 77950847452 3.722 2010     2700      1     2.579
## 119497 77950477601 4.327 2010     2700      1     3.184
## 119499 77950486825 3.875 2010     2700      1     2.732
## 119582 77950475726 4.241 2010     2700      1     3.098
## 119583 77950478147 4.036 2010     2700      1     3.097
## 119585 77950494853 4.319 2010     2700      1     3.076
## 120218 77950246404 4.319 2010     2700      1     3.176
## 120282 77949511388 3.778 2010     2700      1     2.635
## 120453 77649329576 4.566 2010     2700      1     3.323
## 120619 77950355514 3.621 2010     2700      1     2.682
## 121458 77249162677 4.349 2010     2700      1     3.206
## 121459 77249165024 3.722 2010     2700      1     2.579
```

```
## 121517 77149173502 3.864 2010     2700      1     2.826
## 121526 77149120471 3.773 2010     2700      1     2.530
## 121527 77149136099 3.803 2010     2700      1     2.560
## 121596 76649086071 4.375 2010     2700      1     3.436
## 122262 74549178560 4.208 2010     1315      2     3.065
## 122299 74849098405 4.260 2010     2700      1     3.321
## 122328 74549140993 5.722 2010     2700      1     4.479
## 122329 74549159752 5.207 2010     2700      1     4.169
## 122330 74549184608 3.581 2010     2700      1     2.642
## 122358 74949117631 4.260 2010     2700      1     3.222
## 122403 74149095074 3.682 2010     2700      1     2.743
## 122451 74049098673 4.147 2010     2700      1     3.109
## 122453 74049114688 4.434 2010     2700      1     3.291
## 122489 74949089659 3.781 2010     2700      1     2.638
## 122867 77950651061 4.285 2010     2700      1     3.346
## 122994 77952974496 3.813 2010     2700      1     2.670
## 123000 77952987107 3.798 2010     2700      1     2.655
## 123134 77955050878 3.858 2010     2700      1     2.615
## 125455 84855370035 4.953 2011     2700      1     3.917
## 125459 84855425106 5.298 2011     2700      1     4.361
## 125468 84555187324 3.856 2011     2700      1     2.615
## 125541 84255194001 4.947 2011     2700      1     3.806
## 125573 83755178774 3.993 2011     2700      1     2.957
## 125605 83355163991 3.976 2011     2700      1     2.940
## 125606 83455210809 3.567 2011     2700      1     2.630
## 125608 83455235074 3.980 2011     2700      1     3.043
## 125659 83255182742 3.895 2011     2700      1     2.958
## 125678 83155182848 4.324 2011     2700      1     3.183
## 125681 83055177179 3.952 2011     1000      2     2.811
## 125689 83055168450 3.885 2011     2700      1     2.948
## 125690 83055173186 4.057 2011     2700      1     2.916
## 125692 83155168389 3.830 2011     2700      1     2.893
## 125693 83155177076 3.989 2011     2700      1     2.953
## 125777 79960457303 3.482 2011     2700      1     2.545
## 125944 82555168440 4.771 2011     2700      1     3.630
## 125993 84055200171 3.731 2011     2700      1     2.590
## 125994 84055217017 3.890 2011     2700      1     2.749
## 125995 84055217292 3.856 2011     2700      1     2.715
## 127205 82255175382 5.249 2011     2700      1     4.108
## 127216 81855218115 4.488 2011     2700      1     3.347
## 127225 81855161048 3.950 2011     2700      1     3.013
## 127277 81555213207 3.799 2011     2700      1     2.763
## 127287 81455154874 4.549 2011     2700      1     3.408
## 127289 81455159320 4.127 2011     2700      1     2.986
## 127302 81255208675 4.283 2011     2700      1     3.142
## 127362 80755130317 4.428 2011     2700      1     3.491
## 127364 80755188166 4.426 2011     2700      1     3.285
## 127437 80455162464 4.949 2011     2700      1     3.808
## 127457 80355135138 4.850 2011     2700      1     3.913
## 127458 80355148407 4.102 2011     2700      1     2.861
```

```
## 127461 80455179713 3.713 2011     2700      1     2.572
## 127572 80555136815 4.124 2011     2700      1     3.088
## 128312 80054769456 4.528 2011     2700      1     3.387
## 128352 80054714783 3.880 2011     2700      1     2.739
## 128374 80054003232 4.343 2011     2700      1     3.202
## 128462 80053469106 4.076 2011     2700      1     3.139
## 128473 80053510713 3.861 2011     2700      1     2.620
## 128565 80053402552 3.851 2011     2700      1     2.610
## 128577 80053426842 4.065 2011     2700      1     2.824
## 129148 80053359692 4.030 2011     2700      1     2.889
## 129173 84555215993 3.637 2011     2700      1     2.601
## 129220 80053069652 4.106 2011     2700      1     2.965
## 129236 80053065502 3.945 2011     2700      1     2.804
## 129326 80052834654 5.140 2011     2700      1     3.999
## 129352 80052792378 4.382 2011     2700      1     3.241
## 129501 80052462089 3.945 2011     2700      1     2.804
## 129504 80052515621 3.713 2011     2700      1     2.572
## 129538 80052482300 3.725 2011     2700      1     2.584
## 129610 80052371597 3.972 2011     2700      1     2.831
## 129611 80052376347 3.895 2011     2700      1     2.754
## 129619 80052436335 4.001 2011     2700      1     2.860
## 130467 80052358982 3.650 2011     2700      1     2.713
## 130507 80051720194 5.682 2011     2700      1     4.541
## 130522 80052047215 4.316 2011     2700      1     3.175
## 130523 80052063328 3.904 2011     2700      1     2.763
## 130603 84860389821 4.428 2011     2700      1     3.287
## 130622 80051723862 4.452 2011     2700      1     3.311
## 130627 80051831092 4.268 2011     2700      1     3.127
## 130629 84867338003 4.503 2011     2700      1     3.262
## 130677 80051640525 4.042 2011     2700      1     2.901
## 130688 80051571901 4.590 2011     2700      1     3.554
## 130690 80051775476 5.397 2011     2700      1     4.256
## 130735 80051468836 4.057 2011     2700      1     2.916
## 130778 79961102782 3.743 2011     2700      1     2.707
## 130779 79961108629 6.254 2011     2700      1     5.218
## 130780 79961116424 4.087 2011     2700      1     2.946
## 130798 79961040968 3.804 2011     2700      1     2.663
## 130883 79961006281 3.941 2011     2700      1     2.700
## 131346 79960855656 4.776 2011     2700      1     3.635
## 131363 79960704300 3.825 2011     2700      1     2.584
## 131364 79960709532 3.890 2011     2700      1     2.749
## 131435 79960575079 4.596 2011     2700      1     3.455
## 131438 79960611010 3.670 2011     2700      1     2.529
## 131494 79960230129 3.950 2011     2700      1     2.809
## 131599 79960056173 4.241 2011     2700      1     3.100
## 131600 79960066694 3.552 2011     2700      1     2.516
## 131724 79959997891 3.895 2011     2700      1     2.654
## 131727 79960015327 4.590 2011     2700      1     3.554
## 131728 79960015997 5.714 2011     2700      1     4.777
## 131731 79960027556 4.472 2011     2700      1     3.436
```

```
## 132437 79959794787 5.141 2011     2700      1     4.000
## 132474 79959772357 4.719 2011     2700      1     3.578
## 132530 79959478455 5.314 2011     2700      1     4.173
## 132531 79959478969 4.354 2011     2700      1     3.213
## 132542 79959392796 3.748 2011     2700      1     2.507
## 132545 79959426992 4.106 2011     2700      1     2.965
## 132546 79959487586 4.565 2011     2700      1     3.424
## 132606 79959363092 3.574 2011     1000      2     2.637
## 132615 79959297709 3.871 2011     2700      1     2.835
## 132662 79958724181 4.832 2011     2700      1     3.591
## 132668 79959316645 3.976 2011     2700      1     2.835
## 132736 79958195203 4.483 2011     2700      1     3.342
## 132738 79958226299 4.639 2011     2700      1     3.498
## 132745 79958066836 4.674 2011     2700      1     3.533
## 132790 79957795474 4.944 2011     2700      1     3.803
## 132794 79957877087 4.014 2011     2700      1     3.077
## 132801 79952593642 3.682 2011     2700      1     2.541
## 132832 79957520670 3.959 2011     2700      1     2.718
## 132838 79957805100 4.354 2011     2700      1     3.113
## 132839 79957808383 4.120 2011     2700      1     2.979
## 132850 79957889288 3.777 2011     2700      1     2.636
## 132905 79958782230 3.713 2011     2700      1     2.572
## 133601 79957444504 4.382 2011     2700      1     3.346
## 133602 79957445348 4.662 2011     2700      1     3.725
## 133644 79956204057 3.846 2011     2700      1     2.705
## 133684 79956311926 3.968 2011     1000      2     2.727
## 133688 79956189911 4.080 2011     2700      1     2.939
## 133690 79956218863 5.583 2011     2700      1     4.442
## 133696 79956258982 3.993 2011     2700      1     2.752
## 133759 79955909191 4.421 2011     2700      1     3.280
## 133811 79955823001 3.760 2011     2700      1     2.619
## 133813 79955877895 3.880 2011     2700      1     2.739
## 133897 79955624770 4.409 2011     2700      1     3.268
## 133939 79955517235 4.745 2011     2700      1     3.504
## 133996 79955115145 3.695 2011     2700      1     2.554
## 134076 79958006776 3.885 2011     2700      1     2.948
## 134662 79955376345 3.993 2011     2700      1     2.752
## 134719 79955009821 4.244 2011     2700      1     3.003
## 134722 79955016374 3.663 2011     2700      1     2.726
## 134723 79955035027 3.993 2011     2700      1     2.852
## 134724 79955038968 3.967 2011     2700      1     2.826
## 134759 79954561234 4.639 2011     2700      1     3.498
## 134775 79954487919 5.185 2011     2700      1     4.248
## 134778 79954553689 4.428 2011     2700      1     3.287
## 134783 79953848750 3.799 2011     2700      1     2.763
## 134788 79954542852 3.731 2011     2700      1     2.590
## 134883 79953748230 4.910 2011     2700      1     3.669
## 134959 79953297488 4.005 2011     2700      1     2.864
## 134968 79953307878 4.700 2011     2700      1     3.559
## 134978 79953329970 4.175 2011     2700      1     3.034
```

```
## 134979 79953331206 4.492 2011     2700      1     3.351
## 135559 79953225169 4.658 2011     2700      1     3.721
## 135593 79952941296 3.793 2011     2700      1     2.856
## 135629 79953232077 4.567 2011     2700      1     3.426
## 135637 79953118839 5.685 2011     2700      1     4.444
## 135705 79952795093 4.650 2011     2700      1     3.713
## 135762 79952376771 4.188 2011     2700      1     3.047
## 135766 79952424506 4.158 2011     2700      1     3.017
## 135793 79952476762 4.836 2011     2700      1     3.800
## 135911 79952363727 5.013 2011     2700      1     3.872
## 135923 79952267051 4.395 2011     2700      1     3.154
## 135924 79952271716 3.900 2011     2700      1     2.759
## 136005 79952219919 4.001 2011     2700      1     2.860
## 136036 79952570027 3.688 2011     2700      1     2.547
## 136042 79952749048 3.603 2011     2700      1     2.567
## 136941 79951961755 3.835 2011     2700      1     2.694
## 136954 79951978171 3.825 2011     2700      1     2.584
## 136956 79952204045 3.688 2011     2700      1     2.547
## 136959 79951778987 3.914 2011     2700      1     2.773
## 136973 79951850287 4.863 2011     2700      1     3.722
## 136988 79951821800 4.881 2011     2700      1     3.845
## 136997 79951699237 4.608 2011     2700      1     3.367
## 137041 79951514365 3.458 2011     1000      2     2.521
## 137056 79851493086 4.470 2011     2700      1     3.229
## 137064 79751522951 5.633 2011     2700      1     4.696
## 137065 79851475199 3.707 2011     2700      1     2.770
## 137169 79551566712 4.247 2011     2700      1     3.106
## 137251 79251491762 3.799 2011     2700      1     2.658
## 137265 79551523377 4.106 2011     2700      1     3.169
## 137278 79551666741 3.927 2011     2700      1     2.686
## 137775 79251501962 3.820 2011     2700      1     2.679
## 137776 79251525181 3.985 2011     2700      1     2.844
## 137777 79251527865 4.001 2011     2700      1     2.760
## 137786 78751642476 3.804 2011     2700      1     2.663
## 137800 78751610848 4.907 2011     2700      1     3.666
## 137805 79952148734 4.771 2011     2700      1     3.630
## 137813 78751519245 3.731 2011     2700      1     2.695
## 137814 78751532610 3.719 2011     2700      1     2.683
## 137817 78751667739 3.989 2011     2700      1     2.848
## 137857 78651351928 3.959 2011     2700      1     3.022
## 137860 78651378996 3.950 2011     2700      1     2.809
## 137861 78651379500 5.349 2011     2700      1     4.208
## 137865 78651393550 4.594 2011     2700      1     3.453
## 137867 78651505523 3.777 2011     2700      1     2.636
## 137896 78650893810 3.993 2011     2700      1     2.752
## 137905 78650940402 3.574 2011     2700      1     2.538
## 137909 78650949272 4.632 2011     2700      1     3.391
## 137988 78650987230 3.560 2011     2700      1     2.623
## 137995 78651245300 4.046 2011     2700      1     3.010
## 138009 79151481493 3.670 2011     2700      1     2.529
```

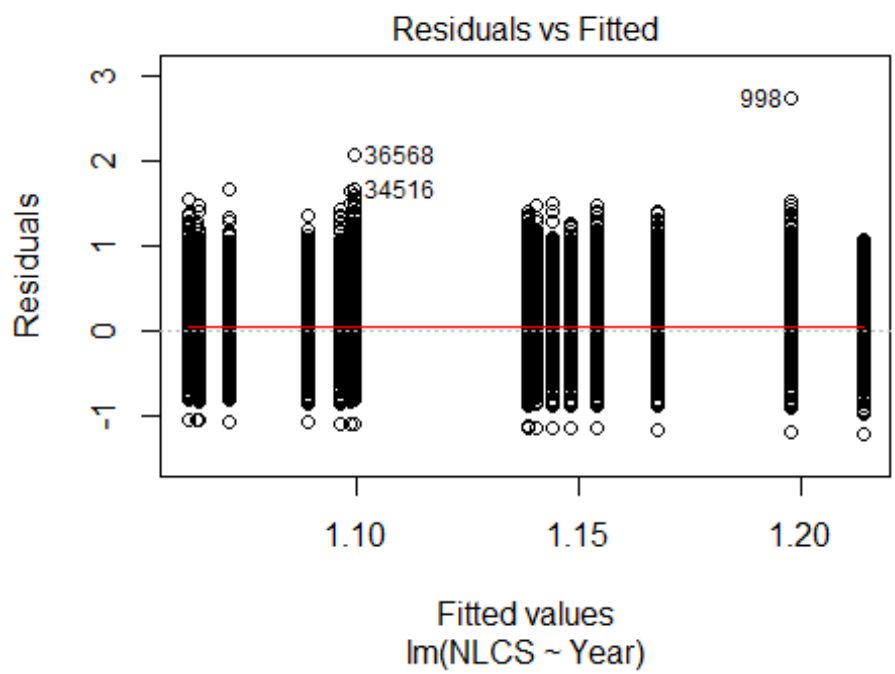
```
## 138381 79955523146 3.731 2011     2700      1     2.590
## 138610 80051964418 3.980 2011     2700      1     2.839
## 139175 79952183161 3.858 2011     2700      2     2.717
## 139231 79958822104 4.047 2011     2700      2     3.110
## 139857 84871293779 4.110 2012     2700      1     2.991
## 139901 84871314601 3.955 2012     2700      1     2.836
## 140049 84870877380 3.955 2012     2700      1     2.836
## 140050 84870887067 4.610 2012     2700      1     3.491
## 140140 84870549609 4.439 2012     2700      1     3.320
## 141706 84870012939 4.380 2012     2700      1     3.261
## 141721 84869792699 3.633 2012     2700      1     2.619
## 141725 84870312106 4.184 2012     2700      1     2.965
## 141912 84868691029 4.249 2012     2700      1     3.130
## 142043 84868305308 4.241 2012     2700      1     3.022
## 142202 84868609657 3.465 2012     2700      1     2.551
## 142208 84868613782 3.623 2012     2700      1     2.504
## 142976 84867828250 4.526 2012     2700      1     3.512
## 143180 84867221377 4.243 2012     2700      1     3.024
## 143205 84867437406 3.667 2012     2700      1     2.548
## 143276 84867273101 3.658 2012     2700      1     2.644
## 144201 84866555961 4.254 2012     2700      1     3.135
## 144204 84866628742 3.794 2012     2700      1     2.780
## 144331 84866098277 4.037 2012     2700      1     2.918
## 144357 84866271313 3.691 2012     2700      1     2.677
## 144361 84866395310 4.216 2012     2700      1     3.097
## 144362 84866395319 3.995 2012     2700      1     2.876
## 144505 84865395104 3.895 2012     2700      1     2.776
## 144517 84865987121 3.945 2012     2700      1     2.826
## 145514 84865454335 3.901 2012     2700      1     2.987
## 145517 84865487218 4.045 2012     2700      1     2.826
## 145540 84865701980 4.297 2012     2700      1     3.178
## 145542 84865820981 3.939 2012     2700      1     2.720
## 145641 84865320778 4.511 2012     2700      1     3.392
## 145642 84865323065 4.426 2012     2700      1     3.512
## 145875 84864842736 4.100 2012     2700      1     2.881
## 146034 84864286532 3.526 2012     2700      1     2.512
## 146664 84864242034 4.379 2012     2700      1     3.260
## 146677 84864138716 3.932 2012     2700      1     2.813
## 146785 84864008031 5.104 2012     2700      1     3.985
## 146790 84930197839 3.672 2012     2700      1     2.553
## 146935 84863923855 3.855 2012     2700      1     2.736
## 147014 84863524725 3.731 2012     2700      1     2.817
## 147015 84863524855 5.149 2012     2700      1     3.930
## 147018 84863533887 4.258 2012     2700      1     3.039
## 147075 84863550146 3.806 2012     2700      1     2.587
## 147080 84863554398 3.686 2012     2700      1     2.772
## 147811 84862859820 6.239 2012     2700      1     5.020
## 147814 84862903106 5.924 2012     2700      1     4.705
## 147913 84862511733 3.543 2012     2700      1     2.629
## 147933 84862859054 3.790 2012     2700      1     2.571
```

```
## 148038 84862226987 4.247 2012     2700      1     3.333
## 148062 84862489245 3.731 2012     2700      1     2.717
## 148141 84862118837 3.961 2012     2700      1     2.842
## 148146 84862260833 4.332 2012     2700      1     3.113
## 148915 84861689703 3.695 2012     2700      1     2.576
## 149011 84861394764 4.694 2012     2700      1     3.575
## 149012 84861409254 3.735 2012     2700      1     2.616
## 149013 84861409551 4.560 2012     2700      1     3.546
## 149126 84861018546 4.042 2012     2700      1     3.128
## 149127 84861022041 4.793 2012     2700      1     3.879
## 149136 84862777742 3.759 2012     1000      2     2.540
## 149215 84860914047 3.560 2012     2700      1     2.546
## 149238 84860741191 4.704 2012     2700      1     3.585
## 149259 84860776122 4.355 2012     2700      1     3.236
## 149264 84860910413 3.658 2012     2700      1     2.539
## 149362 84860718270 4.087 2012     2700      1     2.968
## 149363 84860720391 4.234 2012     2700      1     3.115
## 150121 84860344255 3.718 2012     2700      1     2.599
## 150154 84860116634 3.813 2012     2700      1     2.694
## 150155 84860136615 5.159 2012     2700      1     4.040
## 150175 84860200127 4.100 2012     2700      1     2.981
## 150177 84860358233 3.618 2012     2700      1     2.604
## 150256 84859819731 4.273 2012     2700      1     3.154
## 150277 84859834632 4.025 2012     2700      1     2.906
## 150279 84859846433 4.100 2012     2700      1     2.981
## 150384 84859555074 4.313 2012     2700      1     3.194
## 150445 84860383862 4.156 2012     2700      1     3.037
## 150503 84858665432 4.608 2012     2700      1     3.489
## 150508 84859544287 4.701 2012     2700      1     3.582
## 150550 84859092587 3.942 2012     2700      1     2.823
## 150563 84859718265 4.683 2012     2700      1     3.564
## 151282 84859406106 3.740 2012     2700      1     2.621
## 151344 84858638369 4.465 2012     2700      1     3.346
## 151347 84863337617 4.454 2012     2700      1     3.335
## 151349 84858796262 3.993 2012     1000      2     2.874
## 151408 84858331964 4.546 2012     1000      2     3.427
## 151434 84858265970 4.576 2012     2700      1     3.457
## 151459 84858329412 3.813 2012     2700      1     2.694
## 151554 84863229939 4.007 2012     2700      1     3.093
## 151641 84858019974 4.876 2012     2700      1     3.757
## 151895 84863267488 3.633 2012     2700      1     2.514
## 152472 84857848576 3.905 2012     2700      1     2.686
## 152564 84857427752 4.330 2012     2700      1     3.211
## 152567 84863115197 3.667 2012     2700      1     2.548
## 152654 84857132739 4.189 2012     2700      1     3.070
## 152684 84857065859 4.573 2012     2700      1     3.354
## 152856 84856433589 5.714 2012     2700      1     4.700
## 152866 84856487711 5.900 2012     2700      1     4.886
## 152870 84856552278 3.704 2012     2700      1     2.585
## 152908 84857643783 4.842 2012     2700      1     3.723
```

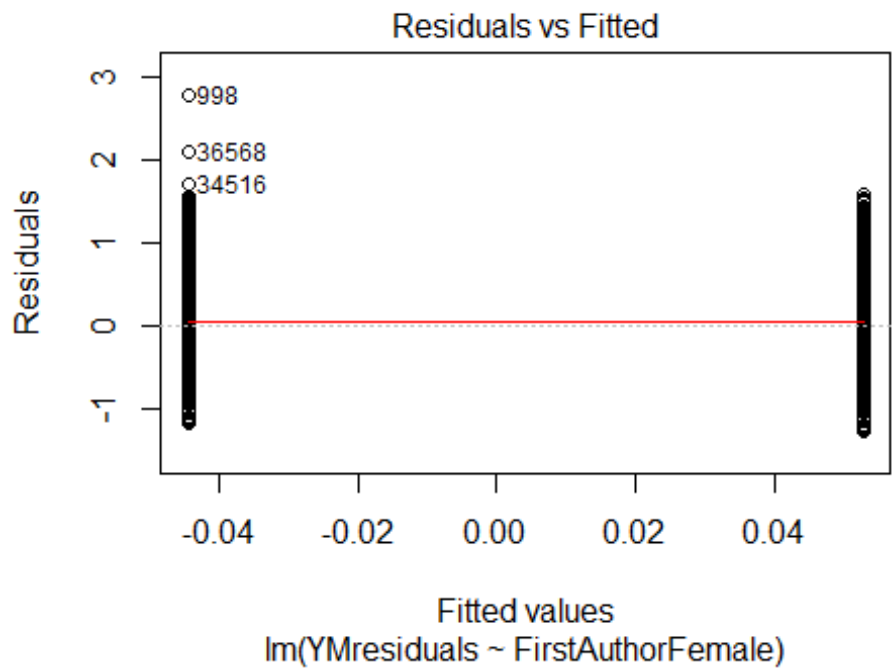
```
## 153052 84863012865 4.087 2012     2700      1     3.073
## 153390 84856245226 4.422 2012     2700      1     3.303
## 153392 84856249964 3.802 2012     2700      1     2.683
## 153393 84856253589 4.184 2012     2700      1     3.065
## 153404 84856133161 3.633 2012     2700      1     2.619
## 153473 84856159009 4.542 2012     2700      1     3.323
## 153475 84862908900 3.581 2012     2700      1     2.567
## 153490 84855993116 3.748 2012     2700      1     2.734
## 153532 84855843640 3.790 2012     2700      1     2.876
## 153533 84855854046 4.736 2012     2700      1     3.617
## 153548 84856176782 4.001 2012     2700      1     3.087
## 153549 84862909100 4.062 2012     2700      1     2.943
## 153663 84855426656 3.662 2012     2700      1     2.543
## 153672 84855459760 4.667 2012     2700      1     3.548
## 153732 84055199809 4.449 2012     2700      1     3.330
## 153835 84856628070 3.840 2012     2700      1     2.721
## 153838 84856657579 4.089 2012     2700      1     2.870
## 154252 84863718620 3.870 2012     2700      1     2.956
## 154329 84864861863 4.382 2012     2700      1     3.263
## 154561 84870478439 4.172 2012     2700      1     3.053
## 154562 84870494510 4.054 2012     2700      1     2.935
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##       Min        1Q    Median        3Q       Max
## -2.681999 -0.923123 -0.000956  0.711044  5.408060
##
## Coefficients:
##                   Estimate Std. Error t value Pr(>|t|)
## (Intercept)        1.93664    0.03108   62.30  < 2e-16 ***
## LastAuthorFemale1 -0.15526    0.00778  -19.95  < 2e-16 ***
## Year1997          -0.00605    0.04247   -0.14     0.89
## Year1998           0.03947    0.04239    0.93     0.35
## Year1999          -0.01676    0.04215   -0.40     0.69
## Year2000           0.74535    0.09340    7.98  1.5e-15 ***
## Year2001          -0.06021    0.04274   -1.41     0.16
## Year2002          -0.53890    0.03668  -14.69  < 2e-16 ***
## Year2003          -0.60135    0.03770  -15.95  < 2e-16 ***
## Year2004          -0.68650    0.03573  -19.21  < 2e-16 ***
## Year2005          -0.71705    0.03545  -20.23  < 2e-16 ***
## Year2006          -0.74544    0.03403  -21.90  < 2e-16 ***
## Year2007          -0.85769    0.03336  -25.71  < 2e-16 ***
## Year2008          -0.85826    0.03285  -26.13  < 2e-16 ***
## Year2009          -0.81321    0.03271  -24.86  < 2e-16 ***
## Year2010          -0.76838    0.03218  -23.88  < 2e-16 ***
## Year2011          -0.76960    0.03297  -23.34  < 2e-16 ***
```

```
## Year2012          -0.79179    0.03239  -24.45  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 1.07
## Multiple R-squared:  0.0737, Adjusted R-squared:  0.0735
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
##  8 observations c(14272,15323,22486,33665,36647,39540,73247,82361)
##   are outliers with |weight| = 0 ( < 1.2e-06);
##  7033 weights are ~= 1. The remaining 79617 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.000   0.894   0.935   0.906   0.983   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.15e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Sample size for the above analysis:  86658"
## [1] ""
## [1] ""
## [1] "####################################"
## [1] "Analysis of AJSC 2701"
## [1] "####################################"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 1601 1737 1427 1407 1680 1584 1476 1575 1688 1688 1998 2302 2239 2335 2451
## 2011 2012
## 2563 2747
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 1145 1197  988 1046  991  766 1171 1272 1310 1345 1602 1838 1757 1788 1910
## 2011 2012
## 1996 2178
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
```

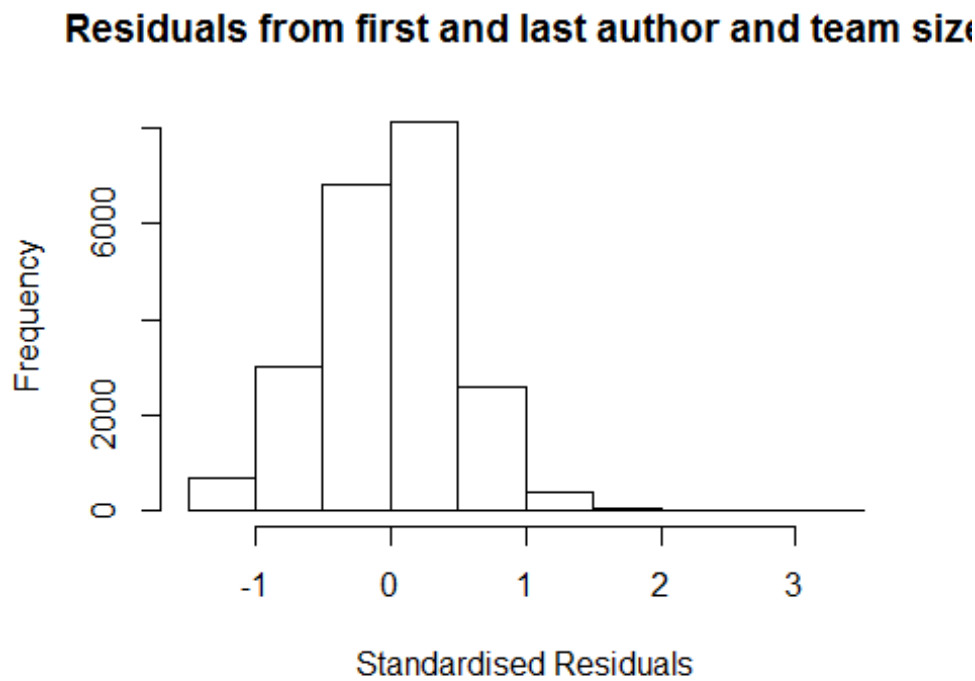
```
## 1030 1083  889  929  893  690 1049 1152 1166 1173 1449 1628 1549 1605 1687
## 2011 2012
## 1768 1924
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
##  Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 190, df = 16, p-value <2e-16
```




```
##
##  Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 120, df = 1, p-value <2e-16
```

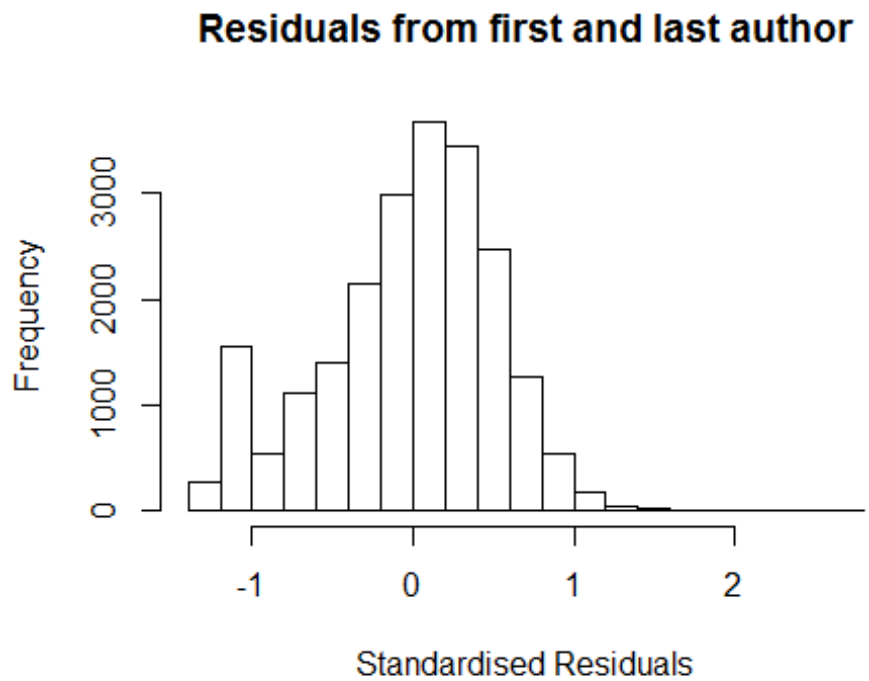



```
## [1] "Female first author team size 2018 geometric mean: 4.56673594068569"
## [1] "Male first author team size 2018 geometric mean: 4.12174019377263"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 510000, p-value = 0.005
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 4.33717253822796"
## [1] "Male last author team size 2018 geometric mean: 4.39348144080591"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 470000, p-value = 0.5
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.050  1           1.025
## LastAuthorFemale  1.028  1           1.014
## UniqueAuthors     1.079  4           1.010
## Year              1.082 16           1.002
```



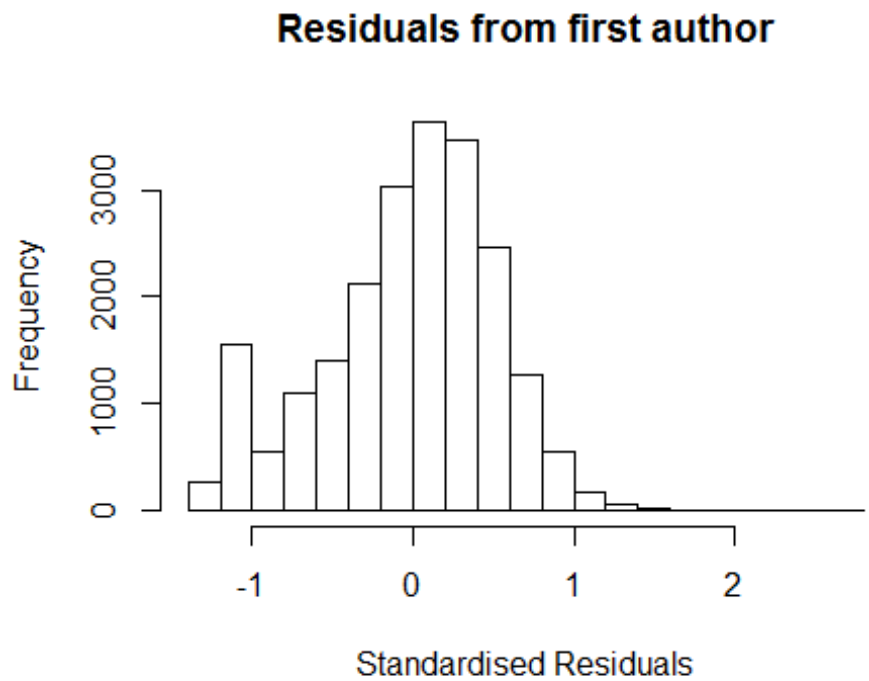

```
## [1] "List of  1 outliers with residuals above  2.5"
##       ScopusId  NLCS Year OneField Fields residuals
## 998 0030175198 3.938 1996     2701      1     3.173
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
UniqueAuthors +
##     Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##     k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.4853 -0.3227  0.0179  0.3185  3.1726
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)          0.76538    0.02356   32.48  < 2e-16 ***
## FirstAuthorFemale1   0.04986    0.00693    7.20  6.3e-13 ***
## LastAuthorFemale1    0.01087    0.00706    1.54   0.1234
## UniqueAuthors2       0.42856    0.01628   26.32  < 2e-16 ***
## UniqueAuthors3       0.52509    0.01516   34.63  < 2e-16 ***
## UniqueAuthors4       0.59425    0.01489   39.91  < 2e-16 ***
## UniqueAuthors5       0.67007    0.01365   49.10  < 2e-16 ***
## Year1997            -0.03252    0.02696   -1.21   0.2278
## Year1998            -0.05840    0.02831   -2.06   0.0391 *
## Year1999            -0.07308    0.02759   -2.65   0.0081 **
```

```
## Year2000           -0.03842    0.02544   -1.51   0.1309
## Year2001           -0.08814    0.02889   -3.05   0.0023 **
## Year2002           -0.11737    0.02533   -4.63  3.6e-06 ***
## Year2003           -0.18039    0.02439   -7.40  1.5e-13 ***
## Year2004           -0.21315    0.02467   -8.64  < 2e-16 ***
## Year2005           -0.19077    0.02413   -7.90  2.8e-15 ***
## Year2006           -0.21964    0.02349   -9.35  < 2e-16 ***
## Year2007           -0.20118    0.02366   -8.50  < 2e-16 ***
## Year2008           -0.21271    0.02382   -8.93  < 2e-16 ***
## Year2009           -0.17081    0.02349   -7.27  3.7e-13 ***
## Year2010           -0.20738    0.02320   -8.94  < 2e-16 ***
## Year2011           -0.18688    0.02316   -8.07  7.5e-16 ***
## Year2012           -0.20092    0.02323   -8.65  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.453
## Multiple R-squared:  0.198,  Adjusted R-squared:  0.197
## Convergence in 15 IRWLS iterations
##
## Robustness weights:
##  observation 555 is an outlier with |weight| = 0 ( < 4.6e-06);
##  1808 weights are ~= 1. The remaining 19855 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0348  0.8550  0.9450  0.8930  0.9840  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          4.62e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.022  1           1.011
## LastAuthorFemale  1.014  1           1.007
## Year              1.028 16           1.001
```



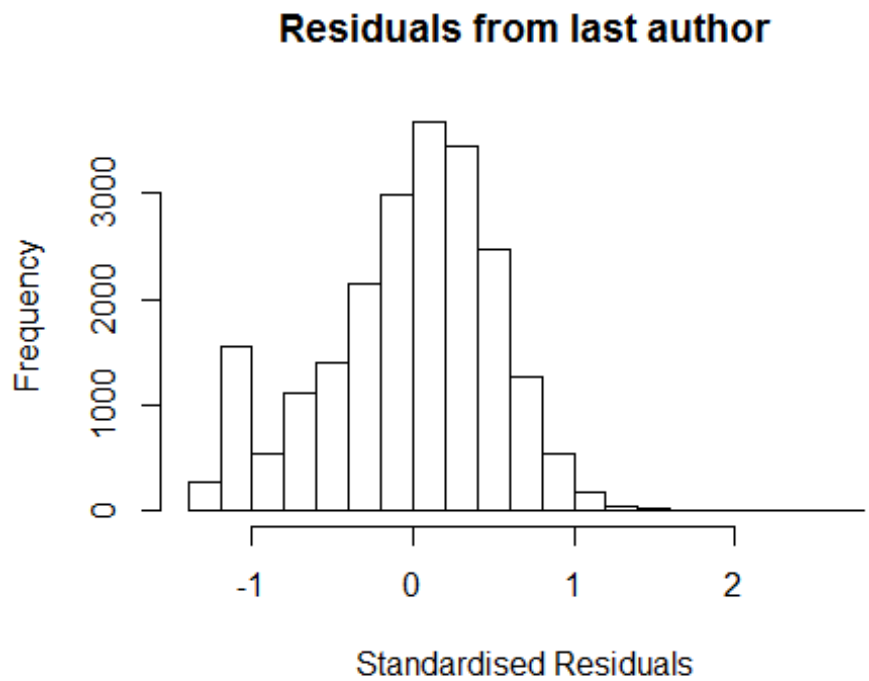

```
## [1] "List of  1 outliers with residuals above  2.5"
##       ScopusId  NLCS Year OneField Fields residuals
## 998 0030175198 3.938 1996     2701      1     2.754
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.3029 -0.3447  0.0474  0.3430  2.7537
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)          1.18429    0.02097   56.48  < 2e-16 ***
## FirstAuthorFemale1   0.10434    0.00746   13.98  < 2e-16 ***
## LastAuthorFemale1    0.01430    0.00758    1.89  0.05934 .
## Year1997            -0.03833    0.02732   -1.40  0.16060
## Year1998            -0.05029    0.03027   -1.66  0.09664 .
## Year1999            -0.06826    0.02855   -2.39  0.01681 *
## Year2000            -0.00438    0.02576   -0.17  0.86496
## Year2001            -0.05960    0.03076   -1.94  0.05273 .
## Year2002            -0.06582    0.02633   -2.50  0.01243 *
## Year2003            -0.12098    0.02568   -4.71  2.5e-06 ***
## Year2004            -0.16054    0.02649   -6.06  1.4e-09 ***
## Year2005            -0.13262    0.02538   -5.22  1.8e-07 ***
```

```
## Year2006           -0.15766    0.02480   -6.36  2.1e-10 ***
## Year2007           -0.15469    0.02493   -6.20  5.6e-10 ***
## Year2008           -0.14456    0.02473   -5.84  5.2e-09 ***
## Year2009           -0.08571    0.02413   -3.55  0.00038 ***
## Year2010           -0.11585    0.02398   -4.83  1.4e-06 ***
## Year2011           -0.09509    0.02387   -3.98  6.8e-05 ***
## Year2012           -0.12892    0.02410   -5.35  8.9e-08 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.488
## Multiple R-squared:  0.0176, Adjusted R-squared:  0.0168
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
##  observation 555 is an outlier with |weight| = 0 ( < 4.6e-06);
##  1728 weights are ~= 1. The remaining 19935 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.191   0.852   0.947   0.891   0.985   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          4.62e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.021  1           1.010
## Year              1.021 16           1.001
```




```
## [1] "List of  1 outliers with residuals above  2.5"
##       ScopusId  NLCS Year OneField Fields residuals
## 998 0030175198 3.938 1996     2701      1     2.754
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.2953 -0.3443  0.0464  0.3431  2.7501
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)          1.18789    0.02077   57.18  < 2e-16 ***
## FirstAuthorFemale1  0.10737    0.00754   14.23  < 2e-16 ***
## Year1997           -0.03818    0.02732   -1.40  0.16221
## Year1998           -0.05096    0.03025   -1.68  0.09205 .
## Year1999           -0.06822    0.02854   -2.39  0.01683 *
## Year2000           -0.00380    0.02575   -0.15  0.88275
## Year2001           -0.05887    0.03076   -1.91  0.05571 .
## Year2002           -0.06592    0.02632   -2.51  0.01225 *
## Year2003           -0.12037    0.02567   -4.69  2.8e-06 ***
## Year2004           -0.16011    0.02648   -6.05  1.5e-09 ***
## Year2005           -0.13250    0.02539   -5.22  1.8e-07 ***
## Year2006           -0.15718    0.02480   -6.34  2.4e-10 ***
```

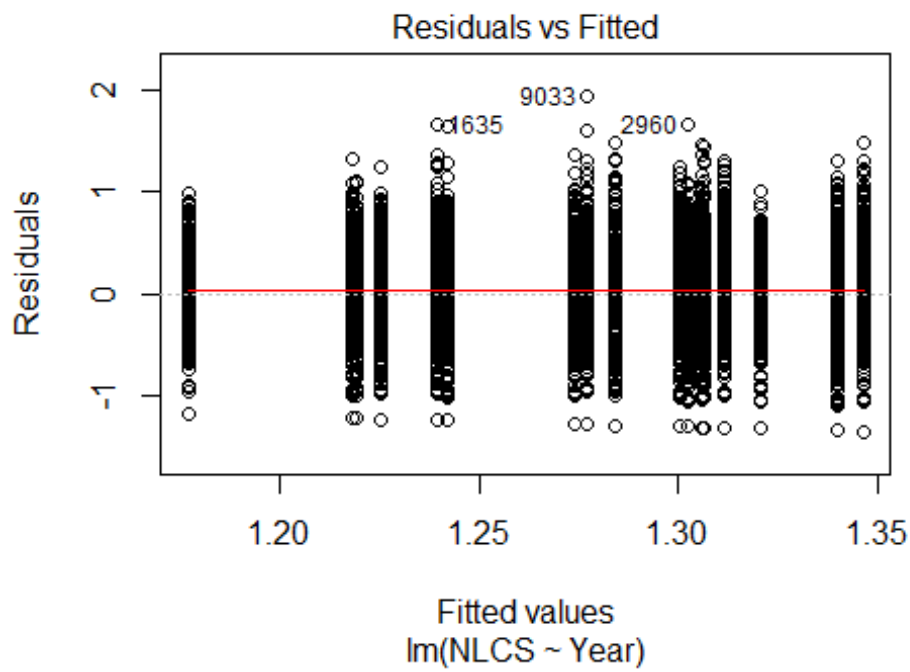
```
## Year2007           -0.15434    0.02493   -6.19  6.1e-10 ***
## Year2008           -0.14429    0.02473   -5.84  5.4e-09 ***
## Year2009           -0.08569    0.02412   -3.55  0.00038 ***
## Year2010           -0.11561    0.02398   -4.82  1.4e-06 ***
## Year2011           -0.09465    0.02386   -3.97  7.3e-05 ***
## Year2012           -0.12811    0.02409   -5.32  1.1e-07 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.488
## Multiple R-squared:  0.0174, Adjusted R-squared:  0.0166
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
##  observation 555 is an outlier with |weight| = 0 ( < 4.6e-06);
##  1753 weights are ~= 1. The remaining 19910 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.194   0.851   0.946   0.891   0.984   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          4.62e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##                   GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.012  1           1.006
## Year             1.012 16           1.000
```



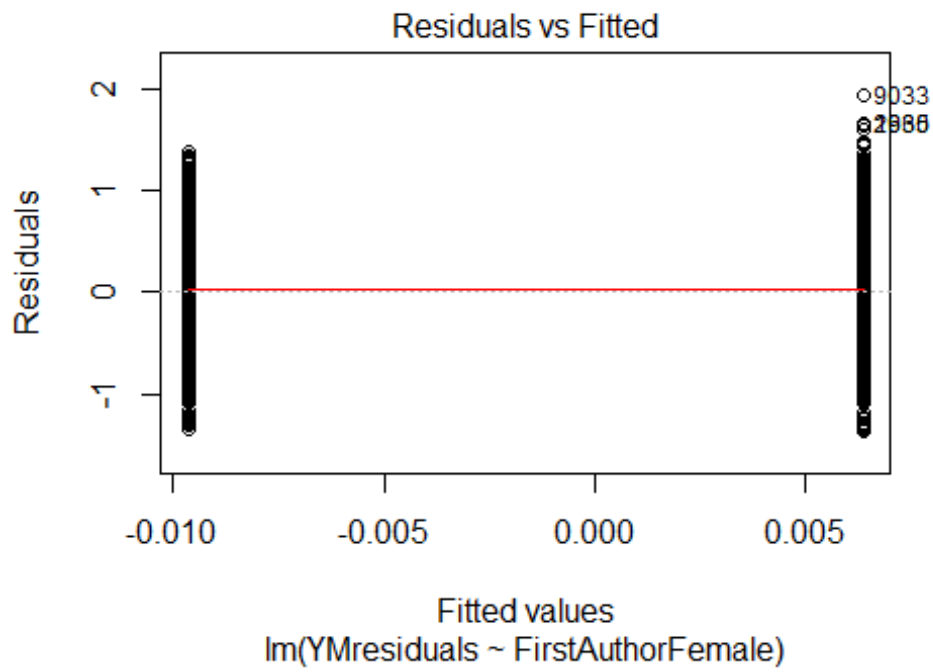

```
## [1] "List of  1 outliers with residuals above  2.5"
##       ScopusId  NLCS Year OneField Fields residuals
## 998 0030175198 3.938 1996     2701      1     2.754
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.2549 -0.3448  0.0493  0.3434  2.7207
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.217262   0.020677   58.87  < 2e-16 ***
## LastAuthorFemale1   0.037672   0.007665    4.91  9.0e-07 ***
## Year1997           -0.037732   0.027246   -1.38  0.16612
## Year1998           -0.051903   0.030164   -1.72  0.08532 .
## Year1999           -0.066088   0.028492   -2.32  0.02038 *
## Year2000           -0.000762   0.025651   -0.03  0.97631
## Year2001           -0.056479   0.030608   -1.85  0.06501 .
## Year2002           -0.059546   0.026299   -2.26  0.02357 *
## Year2003           -0.118129   0.025635   -4.61  4.1e-06 ***
## Year2004           -0.150976   0.026535   -5.69  1.3e-08 ***
## Year2005           -0.123567   0.025397   -4.87  1.1e-06 ***
## Year2006           -0.151197   0.024747   -6.11  1.0e-09 ***
```

```
## Year2007          -0.146286   0.024932   -5.87  4.5e-09 ***
## Year2008          -0.133907   0.024699   -5.42  6.0e-08 ***
## Year2009          -0.072438   0.024043   -3.01  0.00259 **
## Year2010          -0.104446   0.023926   -4.37  1.3e-05 ***
## Year2011          -0.081505   0.023791   -3.43  0.00061 ***
## Year2012          -0.115584   0.024046   -4.81  1.5e-06 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.489
## Multiple R-squared:  0.00826,    Adjusted R-squared:  0.00749
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
##  observation 555 is an outlier with |weight| = 0 ( < 4.6e-06);
##  1731 weights are ~= 1. The remaining 19932 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.220   0.850   0.947   0.891   0.985   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          4.62e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Sample size for the above analysis:  21664"
## [1] ""
## [1] ""
## [1] "####################################"
## [1] "Analysis of AJSC 2702"
## [1] "####################################"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  560  608  601  582  566  564  519  419  478  494  513  509  502  498  502
## 2011 2012
##  529  438
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  385  439  399  415  354  277  379  318  358  362  387  384  366  395  392
```

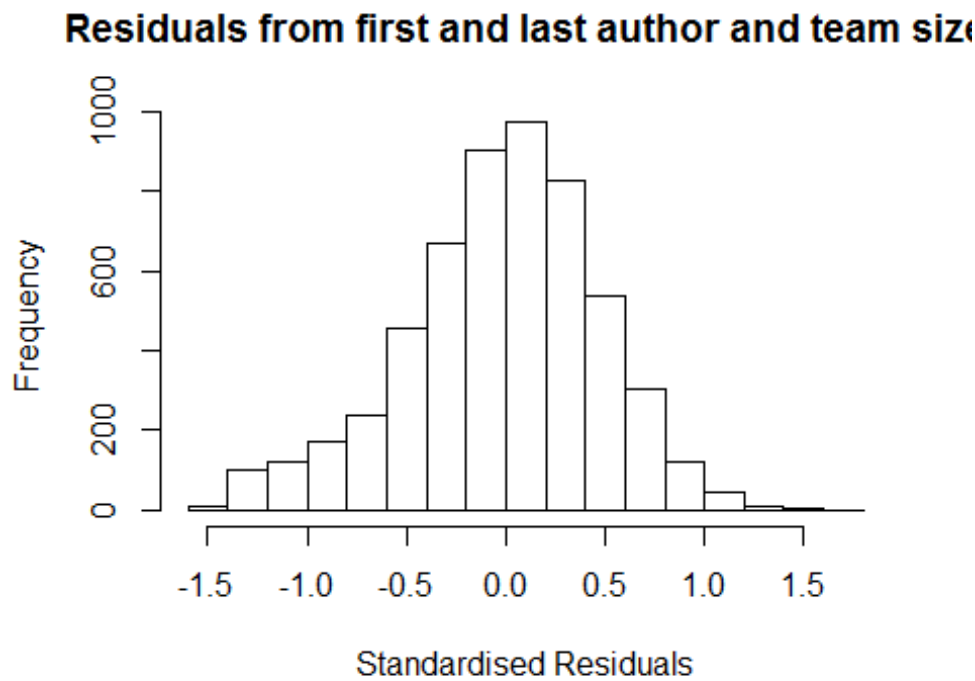
```
## 2011 2012
##  410  331
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  342  386  350  357  310  240  334  269  305  311  334  317  303  344  341
## 2011 2012
##  367  281
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
##  Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 110, df = 16, p-value = 4e-16
```




```
##
##  Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 14, df = 1, p-value = 2e-04
```

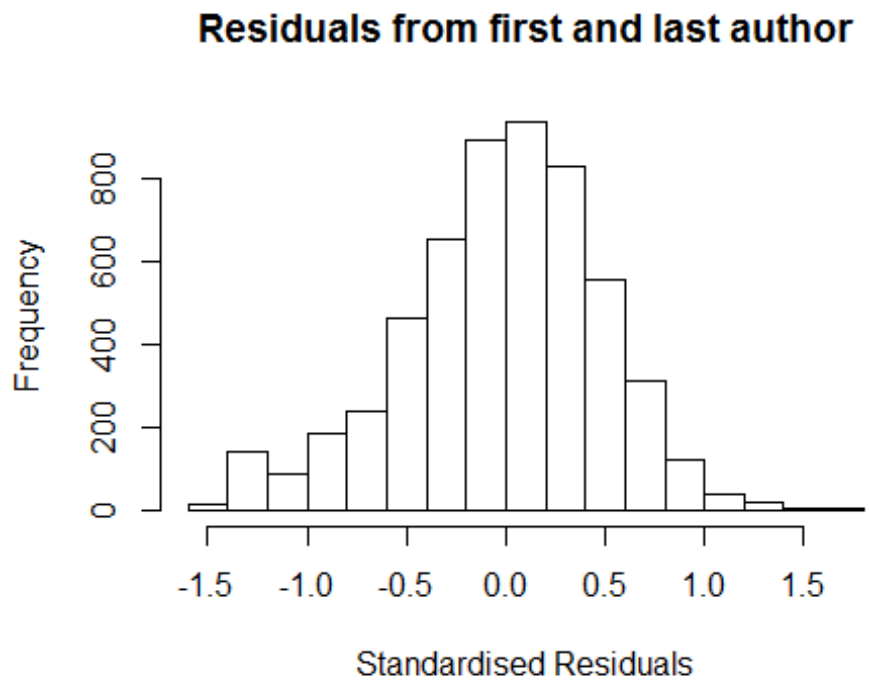



```
## [1] "Female first author team size 2018 geometric mean: 3.74155295145347"
## [1] "Male first author team size 2018 geometric mean: 3.38969759620642"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 19000, p-value = 0.2
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 3.40490706773862"
## [1] "Male last author team size 2018 geometric mean: 3.64118106720565"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 16000, p-value = 0.4
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.070  1           1.034
## LastAuthorFemale  1.069  1           1.034
## UniqueAuthors     1.128  4           1.015
## Year              1.142 16           1.004
```



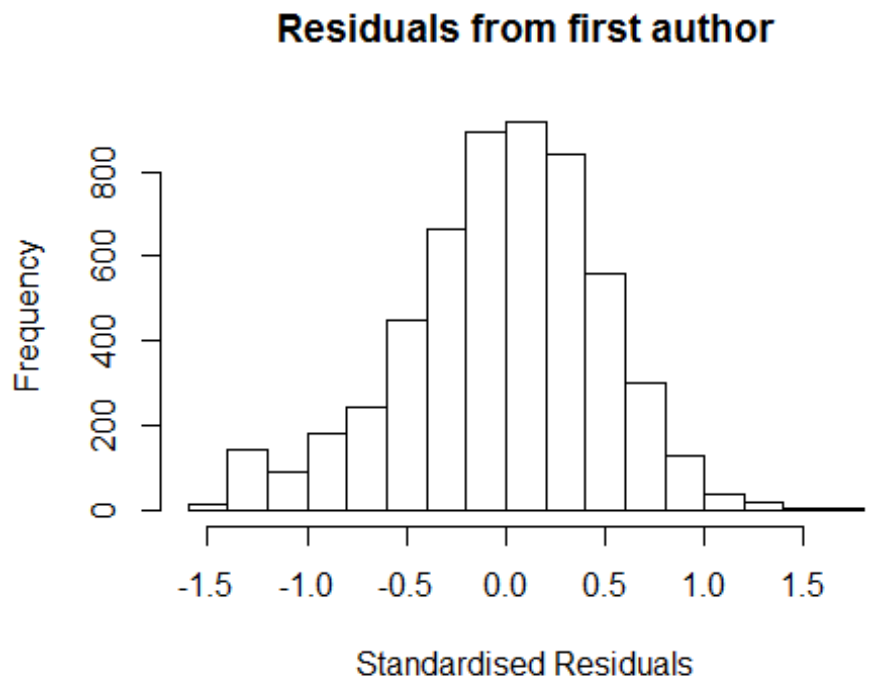

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
UniqueAuthors +
##     Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##     k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.5101 -0.3115  0.0184  0.3122  1.6408
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)          1.2820     0.0378   33.92  < 2e-16 ***
## FirstAuthorFemale1   0.0049     0.0137    0.36  0.71976
## LastAuthorFemale1   -0.0245     0.0151   -1.63  0.10372
## UniqueAuthors2       0.0969     0.0249    3.88  0.00010 ***
## UniqueAuthors3       0.1243     0.0251    4.96  7.3e-07 ***
## UniqueAuthors4       0.1090     0.0261    4.18  2.9e-05 ***
## UniqueAuthors5       0.2805     0.0237   11.85  < 2e-16 ***
## Year1997            -0.0524     0.0429   -1.22  0.22199
## Year1998            -0.1366     0.0436   -3.14  0.00172 **
## Year1999            -0.1428     0.0393   -3.63  0.00028 ***
```

```
## Year2000            -0.0842     0.0415   -2.03  0.04227 *
## Year2001            -0.0962     0.0417   -2.31  0.02099 *
## Year2002            -0.1710     0.0419   -4.08  4.6e-05 ***
## Year2003            -0.1733     0.0432   -4.02  6.0e-05 ***
## Year2004            -0.1876     0.0411   -4.57  5.0e-06 ***
## Year2005            -0.2481     0.0400   -6.20  6.1e-10 ***
## Year2006            -0.1248     0.0412   -3.03  0.00248 **
## Year2007            -0.1043     0.0405   -2.58  0.01004 *
## Year2008            -0.0849     0.0445   -1.91  0.05636 .
## Year2009            -0.0305     0.0436   -0.70  0.48447
## Year2010            -0.1114     0.0414   -2.69  0.00720 **
## Year2011            -0.1044     0.0405   -2.58  0.00991 **
## Year2012            -0.1286     0.0454   -2.83  0.00461 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.458
## Multiple R-squared:  0.0515, Adjusted R-squared:  0.0477
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
##  467 weights are ~= 1. The remaining 5024 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.172   0.867   0.949   0.896   0.985   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.82e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.073  1           1.036
## LastAuthorFemale  1.057  1           1.028
## Year              1.036 16           1.001
```



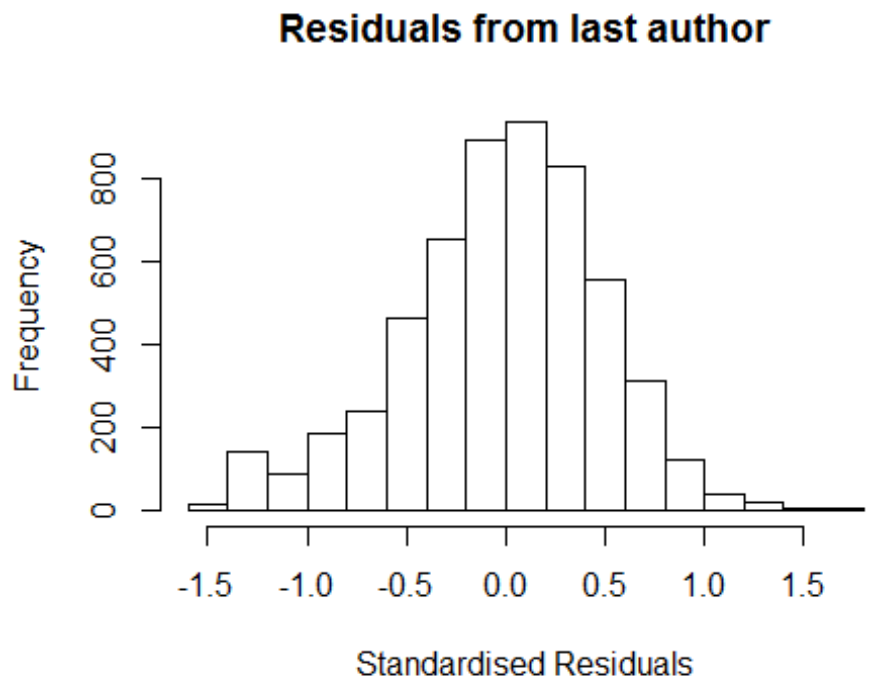

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.4229 -0.3175  0.0159  0.3106  1.6420
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)          1.40568    0.03330   42.21  < 2e-16 ***
## FirstAuthorFemale1  0.00662    0.01389    0.48  0.63362
## LastAuthorFemale1  -0.04083    0.01518   -2.69  0.00717 **
## Year1997           -0.04669    0.04348   -1.07  0.28296
## Year1998           -0.12867    0.04438   -2.90  0.00375 **
## Year1999           -0.13866    0.04018   -3.45  0.00056 ***
## Year2000           -0.07467    0.04228   -1.77  0.07740 .
## Year2001           -0.07528    0.04250   -1.77  0.07654 .
## Year2002           -0.15453    0.04277   -3.61  0.00031 ***
## Year2003           -0.15493    0.04383   -3.53  0.00041 ***
## Year2004           -0.16015    0.04184   -3.83  0.00013 ***
## Year2005           -0.21085    0.04065   -5.19  2.2e-07 ***
```

```
## Year2006           -0.09810    0.04177   -2.35  0.01889 *
## Year2007           -0.07515    0.04080   -1.84  0.06556 .
## Year2008           -0.05413    0.04504   -1.20  0.22950
## Year2009            0.01065    0.04403    0.24  0.80892
## Year2010           -0.07491    0.04207   -1.78  0.07501 .
## Year2011           -0.06596    0.04111   -1.60  0.10867
## Year2012           -0.09442    0.04572   -2.07  0.03894 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.466
## Multiple R-squared:  0.0148, Adjusted R-squared:  0.0115
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
##  458 weights are ~= 1. The remaining 5033 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.189   0.867   0.950   0.895   0.985   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.82e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.026  1           1.013
## Year              1.026 16           1.001
```




```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.4079 -0.3153  0.0158  0.3103  1.6496
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)           1.3973     0.0332   42.11  < 2e-16 ***
## FirstAuthorFemale1   -0.0008     0.0136   -0.06  0.95309
## Year1997             -0.0464     0.0437   -1.06  0.28825
## Year1998             -0.1283     0.0444   -2.89  0.00390 **
## Year1999             -0.1380     0.0402   -3.43  0.00060 ***
## Year2000             -0.0739     0.0423   -1.75  0.08067 .
## Year2001             -0.0742     0.0425   -1.74  0.08141 .
## Year2002             -0.1522     0.0428   -3.55  0.00038 ***
## Year2003             -0.1548     0.0440   -3.52  0.00043 ***
## Year2004             -0.1586     0.0418   -3.79  0.00015 ***
## Year2005             -0.2098     0.0406   -5.16  2.5e-07 ***
## Year2006             -0.0985     0.0419   -2.35  0.01870 *
```

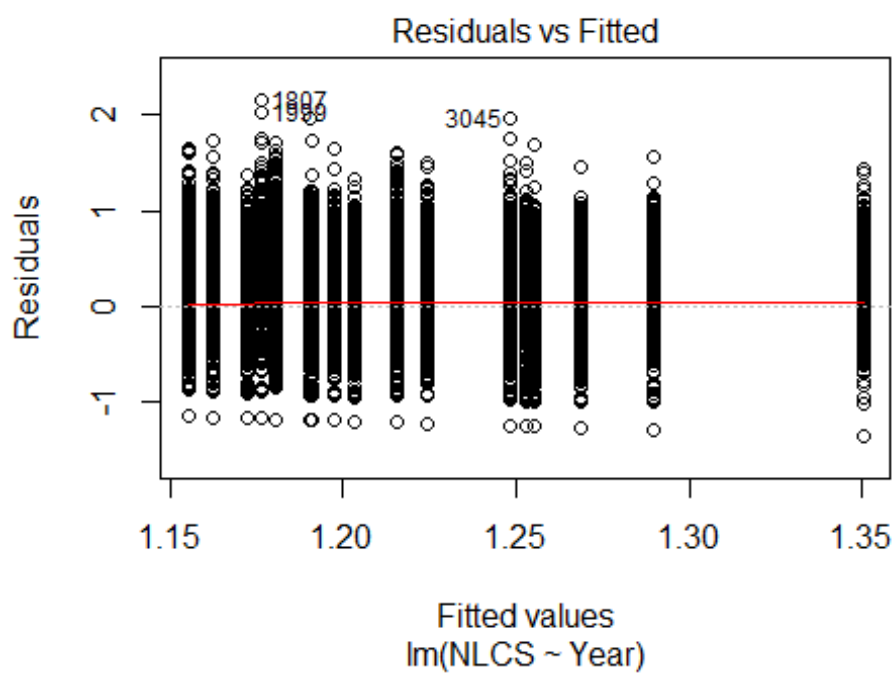
```
## Year2007            -0.0764     0.0408   -1.87  0.06137 .
## Year2008            -0.0552     0.0451   -1.22  0.22107
## Year2009             0.0106     0.0441    0.24  0.80978
## Year2010            -0.0766     0.0421   -1.82  0.06869 .
## Year2011            -0.0665     0.0412   -1.61  0.10680
## Year2012            -0.0963     0.0457   -2.11  0.03497 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.466
## Multiple R-squared:  0.0134, Adjusted R-squared:  0.0103
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
##  451 weights are ~= 1. The remaining 5040 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.185   0.868   0.950   0.896   0.985   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.82e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##                  GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.01  1           1.005
## Year             1.01 16           1.000
```



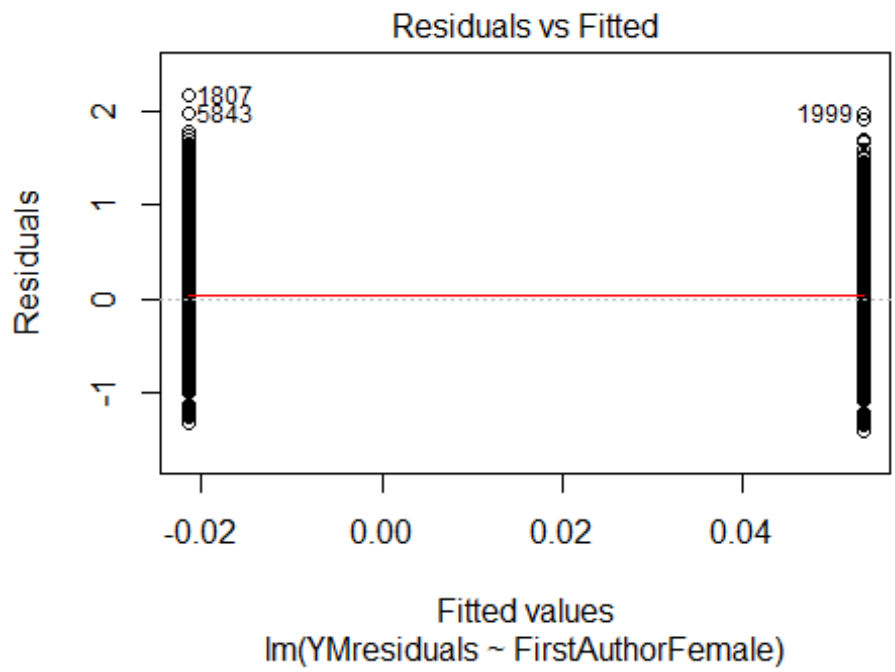

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.4190 -0.3169  0.0157  0.3117  1.6400
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.4079     0.0329   42.73  < 2e-16 ***
## LastAuthorFemale1  -0.0394     0.0148   -2.65  0.00798 **
## Year1997           -0.0466     0.0435   -1.07  0.28352
## Year1998           -0.1287     0.0444   -2.90  0.00376 **
## Year1999           -0.1388     0.0402   -3.46  0.00055 ***
## Year2000           -0.0749     0.0423   -1.77  0.07670 .
## Year2001           -0.0754     0.0425   -1.77  0.07612 .
## Year2002           -0.1541     0.0428   -3.61  0.00031 ***
## Year2003           -0.1548     0.0438   -3.53  0.00042 ***
## Year2004           -0.1602     0.0418   -3.83  0.00013 ***
## Year2005           -0.2106     0.0406   -5.18  2.3e-07 ***
## Year2006           -0.0985     0.0418   -2.36  0.01839 *
```

```
## Year2007           -0.0747     0.0408   -1.83  0.06701 .
## Year2008           -0.0543     0.0450   -1.21  0.22808
## Year2009            0.0111     0.0440    0.25  0.80011
## Year2010           -0.0750     0.0421   -1.78  0.07480 .
## Year2011           -0.0656     0.0411   -1.60  0.11065
## Year2012           -0.0940     0.0457   -2.06  0.03969 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.466
## Multiple R-squared:  0.0147, Adjusted R-squared:  0.0117
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
##  452 weights are ~= 1. The remaining 5039 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.190   0.867   0.949   0.896   0.985   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.82e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Sample size for the above analysis:  5491"
## [1] ""
## [1] ""
## [1] "####################################"
## [1] "Analysis of AJSC 2703"
## [1] "####################################"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 1075 1009  988  913  973  893  961  799  816  847 1002  967 1016 1143 1293
## 2011 2012
## 1462 1347
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  585  621  545  537  675  577  728  593  619  621  744  780  799  907 1012
## 2011 2012
```

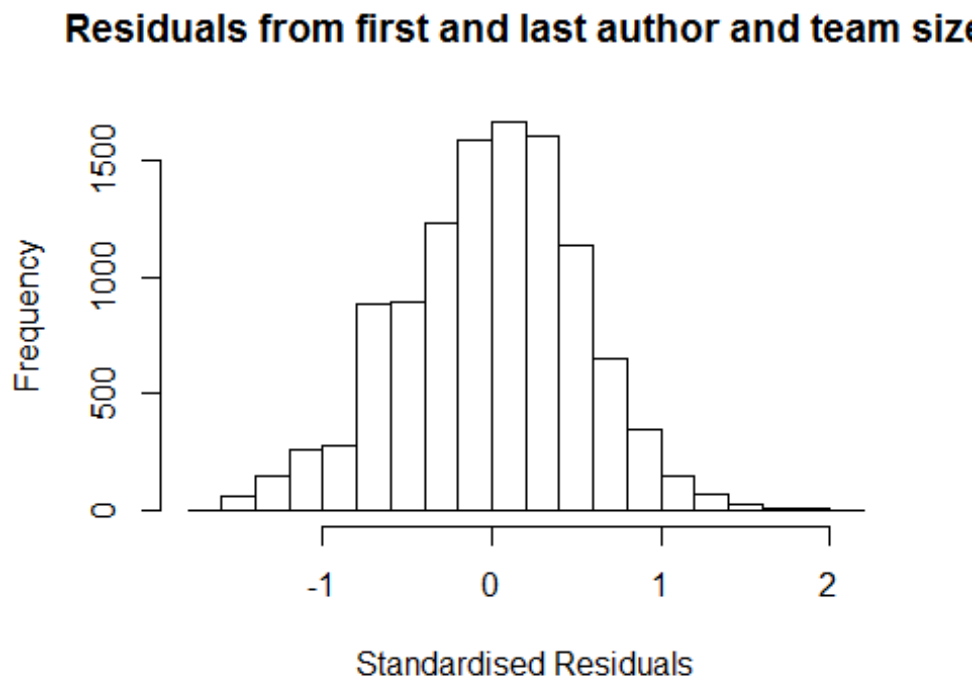
```
## 1127 1034
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  522  571  464  480  587  520  614  521  538  552  636  696  714  788  886
## 2011 2012
## 1000  922
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
##  Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 150, df = 16, p-value <2e-16
```




```
##
##  Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 7.7, df = 1, p-value = 0.006
```

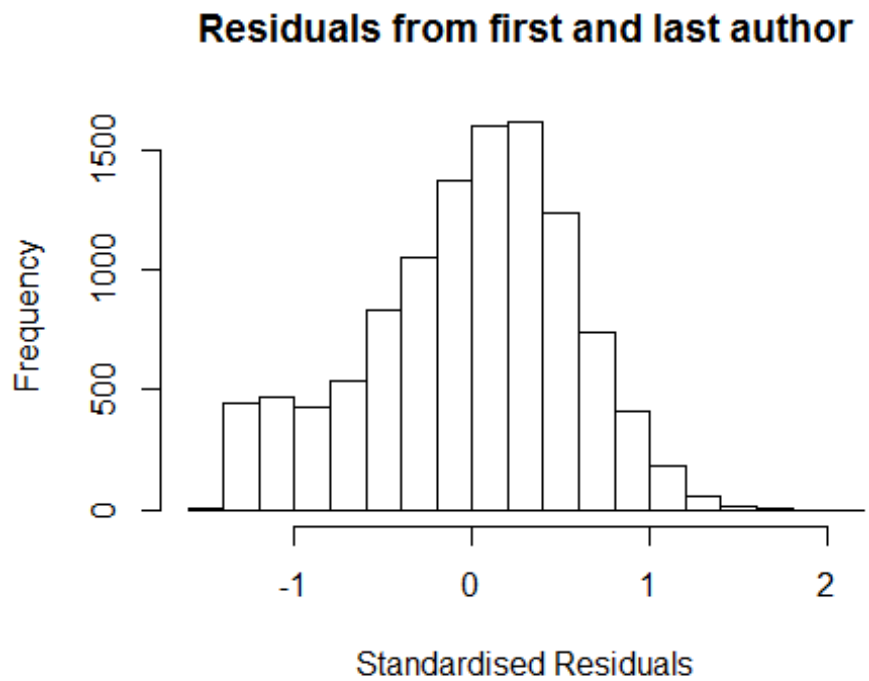



```
## [1] "Female first author team size 2018 geometric mean: 4.76318687514613"
## [1] "Male first author team size 2018 geometric mean: 4.17442521423667"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 130000, p-value = 0.02
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 4.75926904398964"
## [1] "Male last author team size 2018 geometric mean: 4.26529070198421"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 120000, p-value = 0.04
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.064  1           1.031
## LastAuthorFemale  1.047  1           1.023
## UniqueAuthors     1.060  4           1.007
## Year              1.078 16           1.002
```



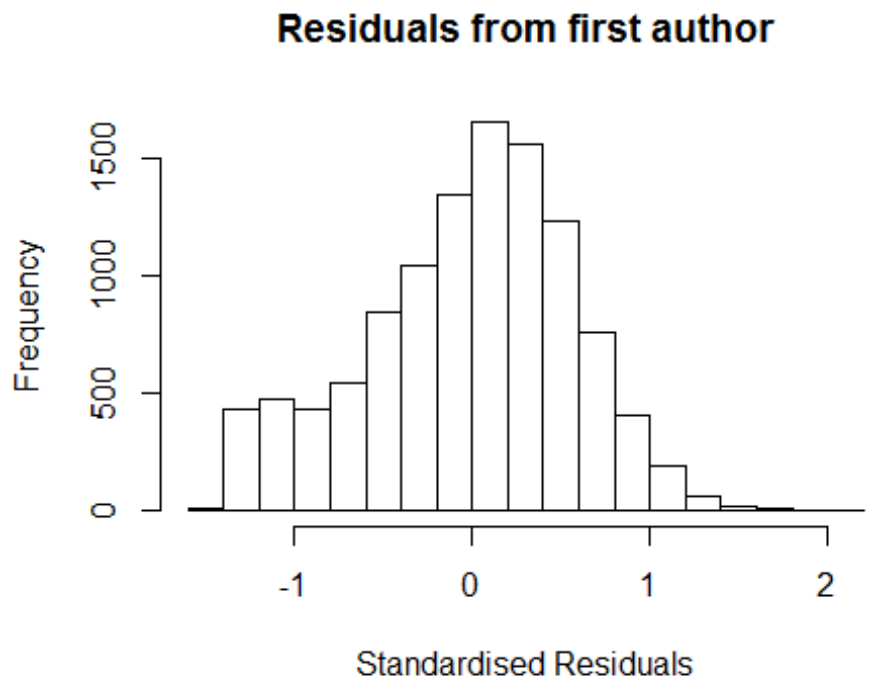

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
UniqueAuthors +
##     Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##     k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.7128 -0.3562  0.0174  0.3477  2.0708
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)           0.8498     0.0296   28.75  < 2e-16 ***
## FirstAuthorFemale1    0.0331     0.0115    2.89  0.00387 **
## LastAuthorFemale1     0.0282     0.0125    2.26  0.02375 *
## UniqueAuthors2        0.3947     0.0229   17.26  < 2e-16 ***
## UniqueAuthors3        0.5328     0.0218   24.46  < 2e-16 ***
## UniqueAuthors4        0.6207     0.0218   28.53  < 2e-16 ***
## UniqueAuthors5        0.8017     0.0193   41.50  < 2e-16 ***
## Year1997             -0.1487     0.0383   -3.88  0.00011 ***
## Year1998             -0.1230     0.0382   -3.22  0.00130 **
## Year1999             -0.0835     0.0355   -2.35  0.01870 *
```

```
## Year2000            -0.1436     0.0338   -4.25  2.2e-05 ***
## Year2001            -0.2043     0.0331   -6.18  6.7e-10 ***
## Year2002            -0.1986     0.0324   -6.14  8.7e-10 ***
## Year2003            -0.1614     0.0321   -5.03  4.9e-07 ***
## Year2004            -0.1646     0.0318   -5.18  2.2e-07 ***
## Year2005            -0.1631     0.0321   -5.08  3.8e-07 ***
## Year2006            -0.1915     0.0327   -5.86  4.8e-09 ***
## Year2007            -0.2462     0.0309   -7.97  1.7e-15 ***
## Year2008            -0.2362     0.0312   -7.56  4.3e-14 ***
## Year2009            -0.2491     0.0314   -7.94  2.1e-15 ***
## Year2010            -0.2322     0.0300   -7.73  1.2e-14 ***
## Year2011            -0.2895     0.0301   -9.61  < 2e-16 ***
## Year2012            -0.2573     0.0311   -8.28  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.508
## Multiple R-squared:  0.215,  Adjusted R-squared:  0.214
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
##  917 weights are ~= 1. The remaining 10094 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0597  0.8640  0.9480  0.9000  0.9840  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          9.08e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.032  1           1.016
## LastAuthorFemale  1.025  1           1.012
## Year              1.041 16           1.001
```



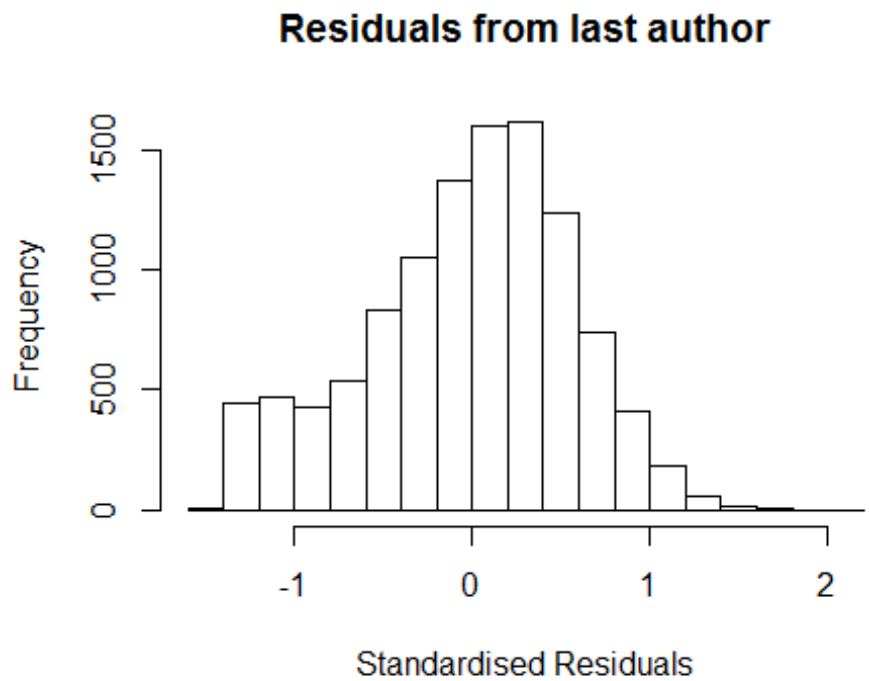

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.4782 -0.3928  0.0523  0.3866  2.1645
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)           1.3642     0.0263   51.88  < 2e-16 ***
## FirstAuthorFemale1    0.0815     0.0126    6.44  1.2e-10 ***
## LastAuthorFemale1     0.0326     0.0140    2.32  0.02041 *
## Year1997             -0.1987     0.0425   -4.67  3.0e-06 ***
## Year1998             -0.1224     0.0419   -2.92  0.00351 **
## Year1999             -0.0746     0.0381   -1.96  0.05010 .
## Year2000             -0.1825     0.0371   -4.92  9.0e-07 ***
## Year2001             -0.1886     0.0361   -5.23  1.8e-07 ***
## Year2002             -0.1723     0.0347   -4.97  6.7e-07 ***
## Year2003             -0.1227     0.0342   -3.59  0.00033 ***
## Year2004             -0.1318     0.0348   -3.79  0.00015 ***
## Year2005             -0.1056     0.0339   -3.12  0.00183 **
```

```
## Year2006            -0.1594     0.0355   -4.49  7.2e-06 ***
## Year2007            -0.2090     0.0336   -6.23  5.0e-10 ***
## Year2008            -0.2036     0.0344   -5.91  3.4e-09 ***
## Year2009            -0.2264     0.0348   -6.50  8.4e-11 ***
## Year2010            -0.1586     0.0336   -4.72  2.3e-06 ***
## Year2011            -0.2544     0.0328   -7.76  9.3e-15 ***
## Year2012            -0.2078     0.0344   -6.04  1.6e-09 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.559
## Multiple R-squared:  0.0143, Adjusted R-squared:  0.0127
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
##  845 weights are ~= 1. The remaining 10166 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0997  0.8590  0.9480  0.8980  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          9.08e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.025  1           1.013
## Year              1.025 16           1.001
```




```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.4568 -0.3969  0.0508  0.3862  2.1597
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)           1.3691     0.0262   52.21  < 2e-16 ***
## FirstAuthorFemale1    0.0877     0.0127    6.90  5.4e-12 ***
## Year1997             -0.1988     0.0426   -4.67  3.1e-06 ***
## Year1998             -0.1218     0.0420   -2.90  0.00373 **
## Year1999             -0.0741     0.0382   -1.94  0.05232 .
## Year2000             -0.1829     0.0371   -4.93  8.6e-07 ***
## Year2001             -0.1883     0.0361   -5.22  1.9e-07 ***
## Year2002             -0.1725     0.0347   -4.98  6.6e-07 ***
## Year2003             -0.1213     0.0342   -3.54  0.00040 ***
## Year2004             -0.1326     0.0348   -3.81  0.00014 ***
## Year2005             -0.1043     0.0339   -3.08  0.00210 **
## Year2006             -0.1592     0.0355   -4.48  7.5e-06 ***
```

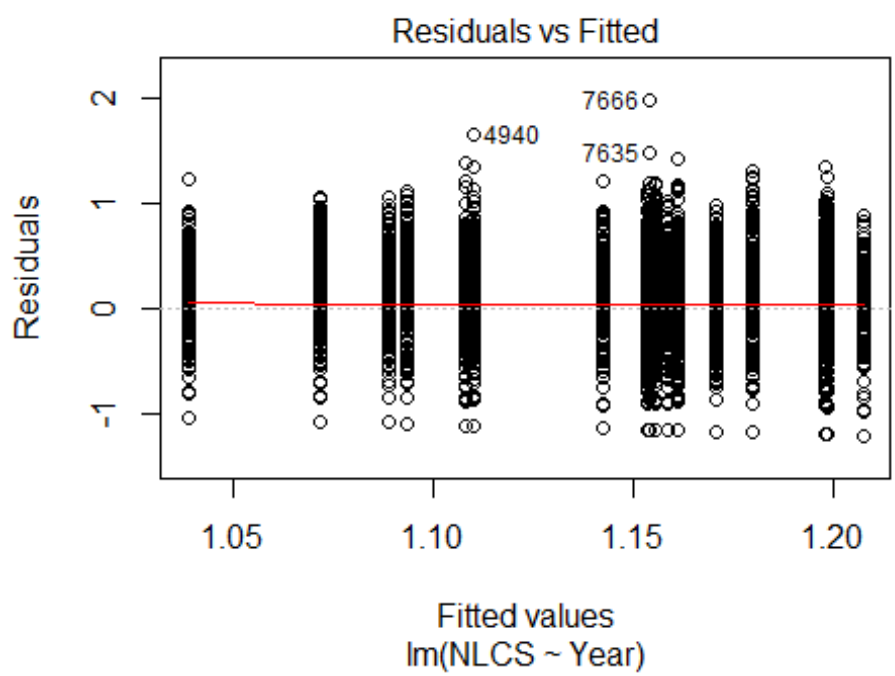
```
## Year2007            -0.2083     0.0336   -6.20  5.8e-10 ***
## Year2008            -0.2029     0.0344   -5.90  3.8e-09 ***
## Year2009            -0.2253     0.0349   -6.46  1.1e-10 ***
## Year2010            -0.1573     0.0336   -4.68  2.9e-06 ***
## Year2011            -0.2532     0.0328   -7.72  1.2e-14 ***
## Year2012            -0.2058     0.0344   -5.98  2.3e-09 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.56
## Multiple R-squared:  0.0137, Adjusted R-squared:  0.0121
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
##  846 weights are ~= 1. The remaining 10165 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.103   0.859   0.948   0.899   0.985   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          9.08e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##                   GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.017  1           1.009
## Year             1.017 16           1.001
```



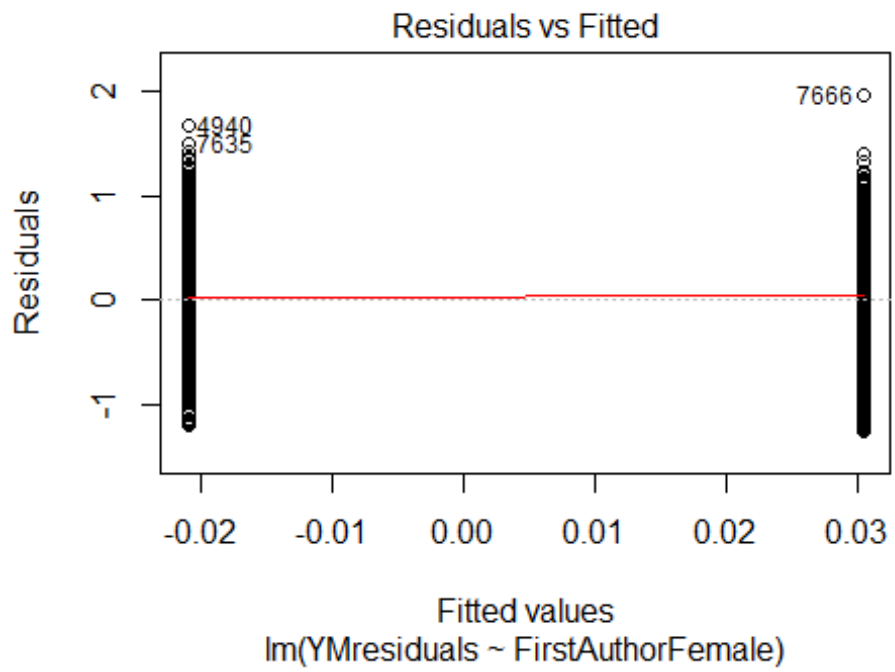

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.4274 -0.3926  0.0523  0.3897  2.1524
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)          1.3768     0.0262   52.54  < 2e-16 ***
## LastAuthorFemale1    0.0506     0.0140    3.60  0.00032 ***
## Year1997            -0.1992     0.0426   -4.68  2.9e-06 ***
## Year1998            -0.1196     0.0419   -2.86  0.00429 **
## Year1999            -0.0744     0.0380   -1.96  0.05060 .
## Year2000            -0.1786     0.0371   -4.81  1.5e-06 ***
## Year2001            -0.1859     0.0361   -5.14  2.7e-07 ***
## Year2002            -0.1663     0.0347   -4.80  1.6e-06 ***
## Year2003            -0.1170     0.0342   -3.42  0.00063 ***
## Year2004            -0.1217     0.0348   -3.50  0.00047 ***
## Year2005            -0.0965     0.0339   -2.85  0.00439 **
## Year2006            -0.1527     0.0354   -4.31  1.6e-05 ***
```

```
## Year2007           -0.2021     0.0336   -6.02  1.8e-09 ***
## Year2008           -0.1947     0.0345   -5.65  1.7e-08 ***
## Year2009           -0.2156     0.0349   -6.19  6.4e-10 ***
## Year2010           -0.1494     0.0335   -4.46  8.1e-06 ***
## Year2011           -0.2441     0.0328   -7.45  1.0e-13 ***
## Year2012           -0.1957     0.0344   -5.70  1.3e-08 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.56
## Multiple R-squared:  0.0105, Adjusted R-squared:  0.00897
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
##  844 weights are ~= 1. The remaining 10167 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.107   0.860   0.948   0.899   0.985   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          9.08e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Sample size for the above analysis:  11011"
## [1] ""
## [1] ""
## [1] "####################################"
## [1] "Analysis of AJSC 2704"
## [1] "####################################"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  407  391  379  378  402  391  400  276  337  339  315  401  469  475  467
## 2011 2012
##  538  472
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  261  239  249  239  222  148  268  186  240  237  232  295  346  336  327
## 2011 2012
```

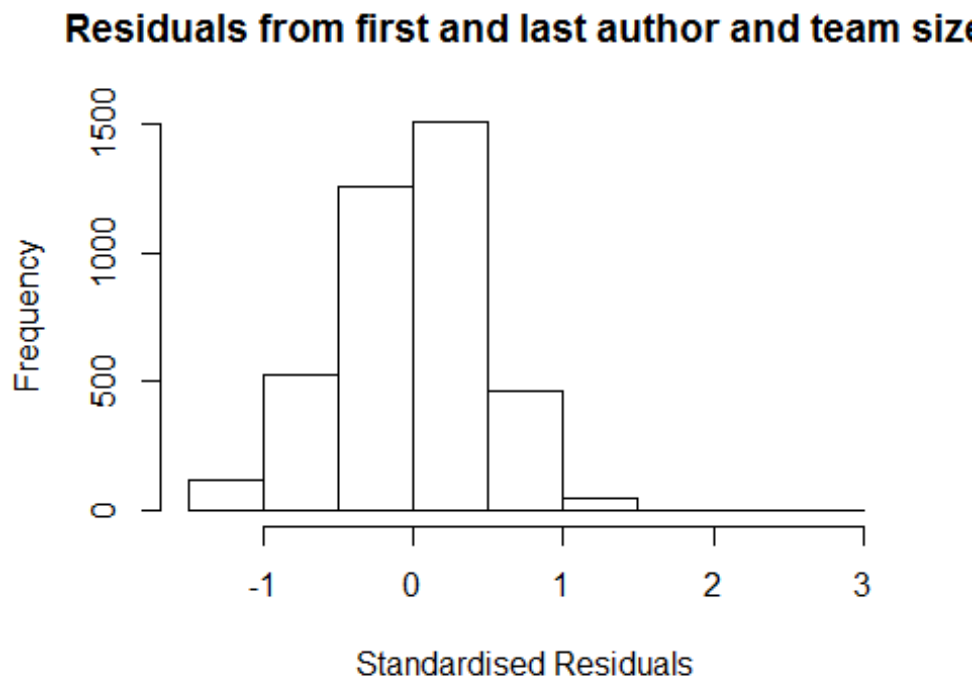
```
##  385  338
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  228  216  226  210  203  130  236  154  218  209  201  248  302  288  256
## 2011 2012
##  322  282
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
##  Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 46, df = 16, p-value = 8e-05
```




```
##
##  Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 0.029, df = 1, p-value = 0.9
```

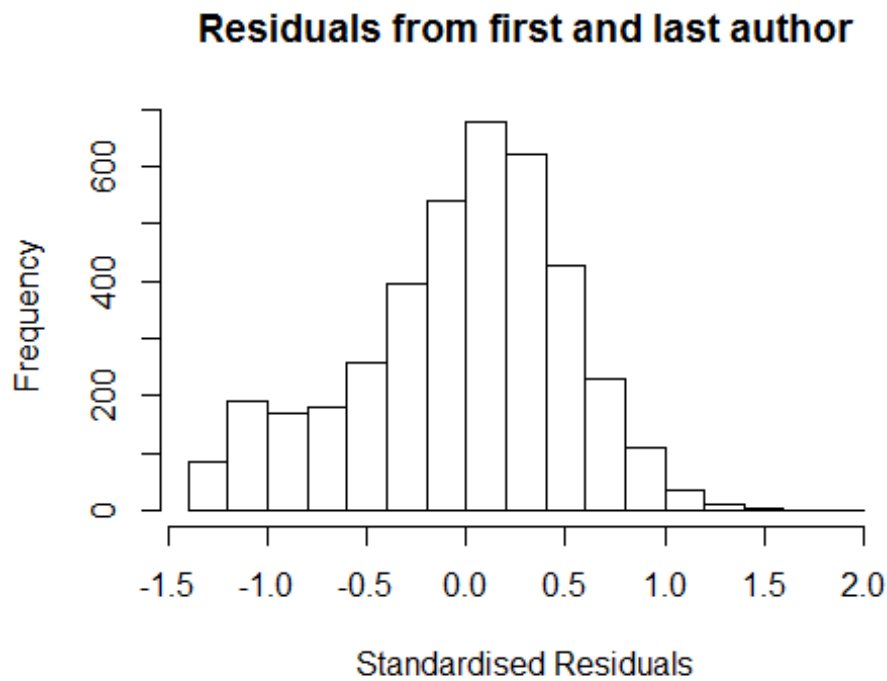



```
## [1] "Female first author team size 2018 geometric mean: 5.79951277723436"
## [1] "Male first author team size 2018 geometric mean: 4.73803810854271"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 6000, p-value = 0.07
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 5.08128910629005"
## [1] "Male last author team size 2018 geometric mean: 5.36766216968965"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 4400, p-value = 0.3
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.064  1           1.032
## LastAuthorFemale  1.037  1           1.018
## UniqueAuthors     1.112  4           1.013
## Year              1.119 16           1.004
```



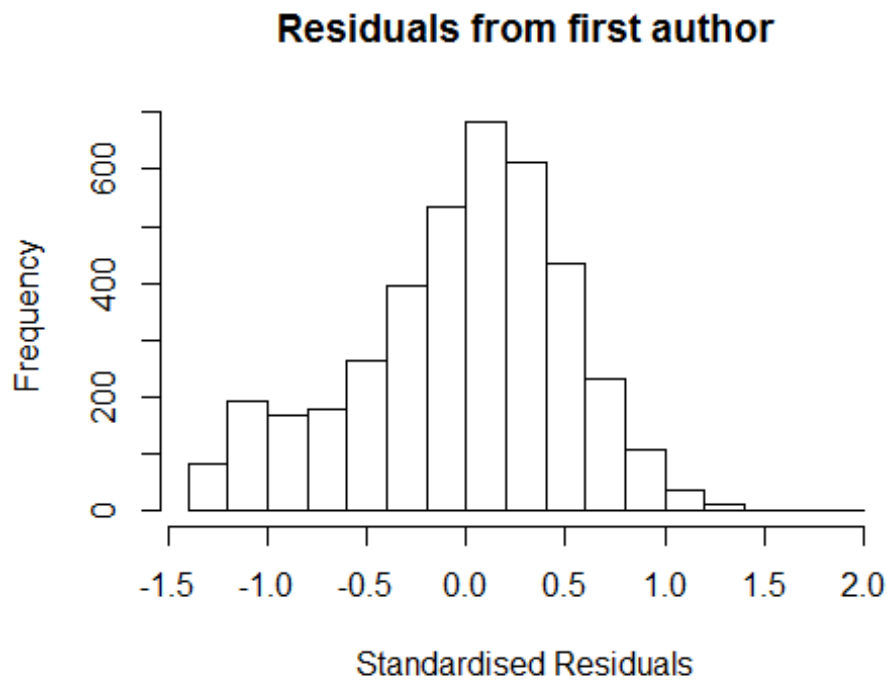

```
## [1] "List of  1 outliers with residuals above  2.5"
##         ScopusId  NLCS Year OneField Fields residuals
## 7666 84867276477 3.137 2012     1308      2     2.577
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
UniqueAuthors +
##     Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##     k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.4663 -0.3128  0.0154  0.3122  2.5774
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         0.65069    0.05139   12.66  < 2e-16 ***
## FirstAuthorFemale1  0.00705    0.01627    0.43  0.66488
## LastAuthorFemale1   0.00169    0.01766    0.10  0.92365
## UniqueAuthors2      0.25215    0.04665    5.40  6.9e-08 ***
## UniqueAuthors3      0.44299    0.04279   10.35  < 2e-16 ***
## UniqueAuthors4      0.57200    0.04100   13.95  < 2e-16 ***
## UniqueAuthors5      0.77896    0.03597   21.65  < 2e-16 ***
## Year1997            0.03661    0.05702    0.64  0.52083
## Year1998           -0.07049    0.05055   -1.39  0.16328
## Year1999           -0.08578    0.05442   -1.58  0.11508
```

```
## Year2000           -0.01426    0.05206   -0.27  0.78415
## Year2001           -0.02788    0.05191   -0.54  0.59129
## Year2002           -0.11057    0.05167   -2.14  0.03243 *
## Year2003           -0.10460    0.05577   -1.88  0.06081 .
## Year2004           -0.15907    0.05193   -3.06  0.00220 **
## Year2005           -0.19556    0.05192   -3.77  0.00017 ***
## Year2006           -0.15781    0.05363   -2.94  0.00328 **
## Year2007           -0.14364    0.05363   -2.68  0.00743 **
## Year2008           -0.07293    0.04977   -1.47  0.14287
## Year2009           -0.03787    0.04912   -0.77  0.44079
## Year2010           -0.03402    0.04976   -0.68  0.49425
## Year2011           -0.07062    0.04889   -1.44  0.14868
## Year2012           -0.09979    0.05042   -1.98  0.04785 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.453
## Multiple R-squared:  0.227,  Adjusted R-squared:  0.222
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
##  observation 3699 is an outlier with |weight| = 0 ( < 2.5e-05);
##  341 weights are ~= 1. The remaining 3587 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.129   0.858   0.948   0.896   0.984   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          2.55e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.033  1           1.016
## LastAuthorFemale  1.023  1           1.011
## Year              1.055 16           1.002
```



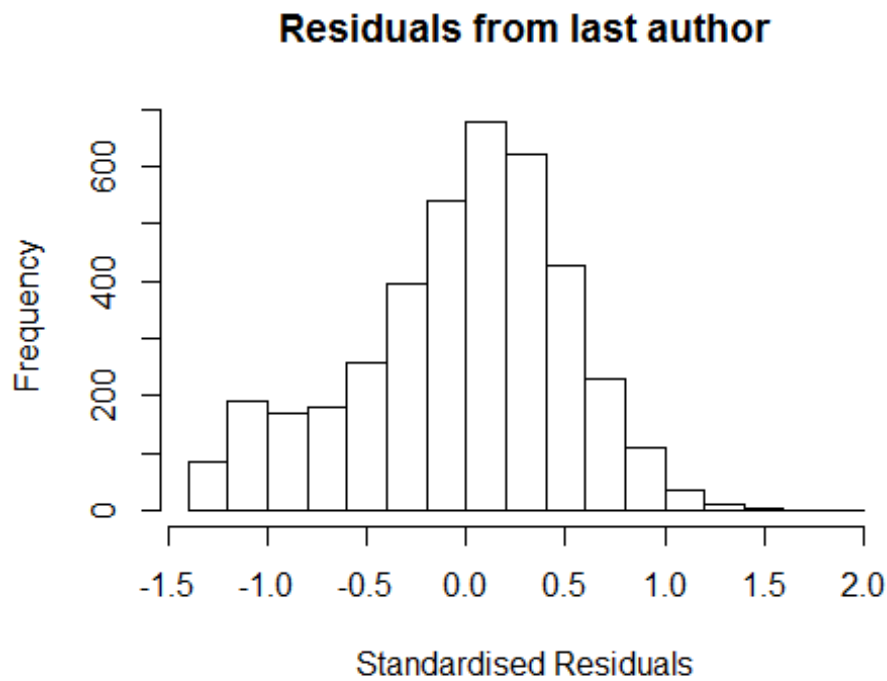

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.3108 -0.3489  0.0479  0.3395  1.9091
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)          1.18349    0.04842   24.44  < 2e-16 ***
## FirstAuthorFemale1   0.06981    0.01803    3.87  0.00011 ***
## LastAuthorFemale1    0.01416    0.02051    0.69  0.49011
## Year1997             0.04337    0.06591    0.66  0.51058
## Year1998            -0.01402    0.06046   -0.23  0.81665
## Year1999            -0.10118    0.06573   -1.54  0.12381
## Year2000            -0.00477    0.06154   -0.08  0.93826
## Year2001             0.03437    0.05995    0.57  0.56641
## Year2002            -0.08428    0.05852   -1.44  0.14989
## Year2003            -0.05139    0.06358   -0.81  0.41895
## Year2004            -0.09938    0.06074   -1.64  0.10190
## Year2005            -0.12294    0.06149   -2.00  0.04564 *
```

```
## Year2006           -0.10301    0.06303   -1.63  0.10227
## Year2007           -0.08377    0.06130   -1.37  0.17187
## Year2008           -0.04041    0.05918   -0.68  0.49472
## Year2009            0.01768    0.05727    0.31  0.75751
## Year2010           -0.00391    0.05802   -0.07  0.94622
## Year2011           -0.03439    0.05595   -0.61  0.53890
## Year2012           -0.03957    0.05894   -0.67  0.50204
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.491
## Multiple R-squared:  0.012,  Adjusted R-squared:  0.00747
## Convergence in 15 IRWLS iterations
##
## Robustness weights:
##  315 weights are ~= 1. The remaining 3614 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0963  0.8460  0.9480  0.8910  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          2.55e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.034  1           1.017
## Year              1.034 16           1.001
```




```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##    Min     1Q Median     3Q    Max
## -1.300 -0.348  0.048  0.340  1.918
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)          1.18528    0.04845   24.47  < 2e-16 ***
## FirstAuthorFemale1   0.07145    0.01814    3.94  8.3e-05 ***
## Year1997             0.04339    0.06583    0.66    0.510
## Year1998            -0.01357    0.06037   -0.22    0.822
## Year1999            -0.09969    0.06559   -1.52    0.129
## Year2000            -0.00384    0.06144   -0.06    0.950
## Year2001             0.03530    0.05983    0.59    0.555
## Year2002            -0.08299    0.05836   -1.42    0.155
## Year2003            -0.05043    0.06340   -0.80    0.426
## Year2004            -0.09873    0.06065   -1.63    0.104
## Year2005            -0.12150    0.06135   -1.98    0.048 *
## Year2006            -0.10177    0.06291   -1.62    0.106
```

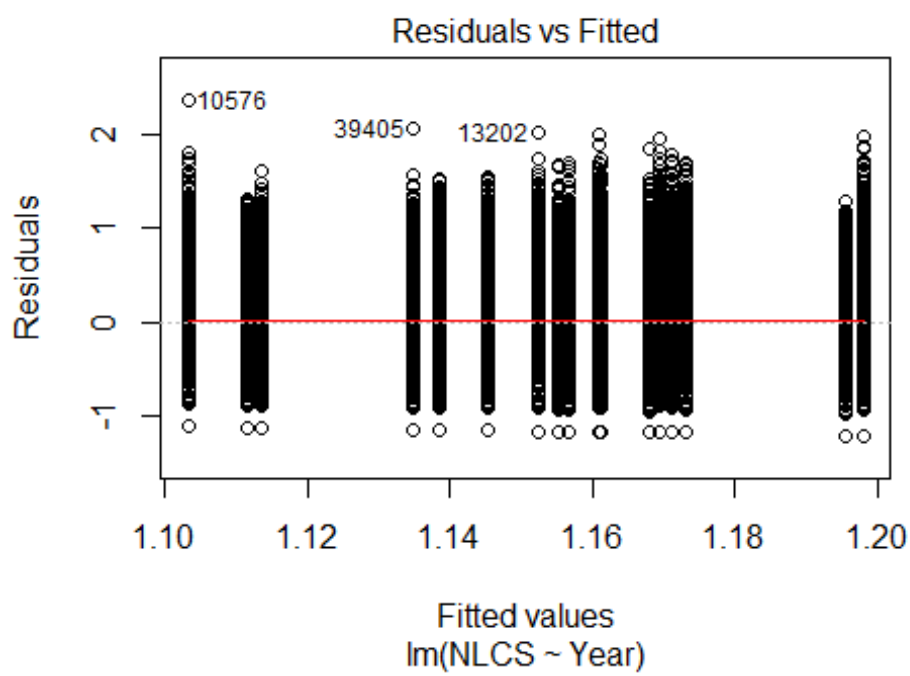
```
## Year2007           -0.08299    0.06116   -1.36    0.175
## Year2008           -0.03878    0.05896   -0.66    0.511
## Year2009            0.01831    0.05719    0.32    0.749
## Year2010           -0.00265    0.05788   -0.05    0.963
## Year2011           -0.03308    0.05579   -0.59    0.553
## Year2012           -0.03762    0.05865   -0.64    0.521
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.491
## Multiple R-squared:  0.0118, Adjusted R-squared:  0.00753
## Convergence in 15 IRWLS iterations
##
## Robustness weights:
##  309 weights are ~= 1. The remaining 3620 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0936  0.8470  0.9490  0.8910  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          2.55e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##                   GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.022  1           1.011
## Year             1.022 16           1.001
```



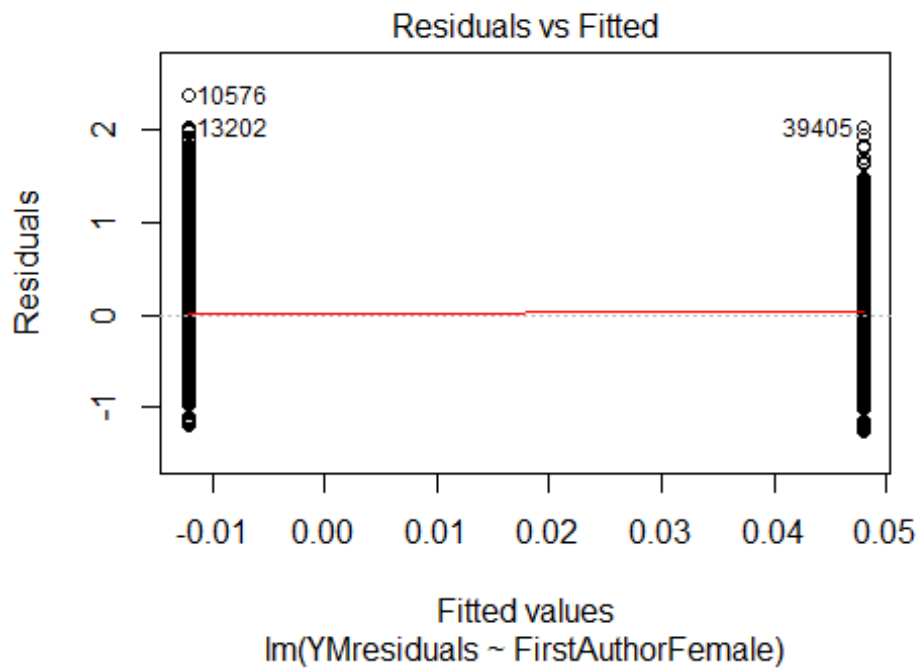

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.2722 -0.3476  0.0489  0.3418  1.9403
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.20523    0.04794   25.14   <2e-16 ***
## LastAuthorFemale1   0.02448    0.02058    1.19    0.234
## Year1997            0.04251    0.06554    0.65    0.517
## Year1998           -0.01546    0.06007   -0.26    0.797
## Year1999           -0.10123    0.06547   -1.55    0.122
## Year2000           -0.00749    0.06133   -0.12    0.903
## Year2001            0.03486    0.05976    0.58    0.560
## Year2002           -0.07981    0.05813   -1.37    0.170
## Year2003           -0.04714    0.06340   -0.74    0.457
## Year2004           -0.09741    0.06054   -1.61    0.108
## Year2005           -0.11743    0.06127   -1.92    0.055 .
## Year2006           -0.09352    0.06247   -1.50    0.134
```

```
## Year2007          -0.07757    0.06097   -1.27    0.203
## Year2008          -0.03564    0.05876   -0.61    0.544
## Year2009           0.02336    0.05698    0.41    0.682
## Year2010          -0.00214    0.05768   -0.04    0.970
## Year2011          -0.03064    0.05560   -0.55    0.582
## Year2012          -0.03302    0.05851   -0.56    0.572
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.496
## Multiple R-squared:  0.00787,    Adjusted R-squared:  0.00356
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
##  316 weights are ~= 1. The remaining 3613 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0921  0.8520  0.9480  0.8930  0.9840  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          2.55e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Sample size for the above analysis:  3929"
## [1] ""
## [1] ""
## [1] "####################################"
## [1] "Analysis of AJSC 2705"
## [1] "####################################"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 4915 5244 4958 4595 5200 5279 4713 3775 4165 4694 5196 5479 5876 5920 5847
## 2011 2012
## 6145 5810
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 2943 3251 2897 2764 3518 3337 3637 2916 3160 3577 3940 4180 4525 4580 4532
## 2011 2012
```

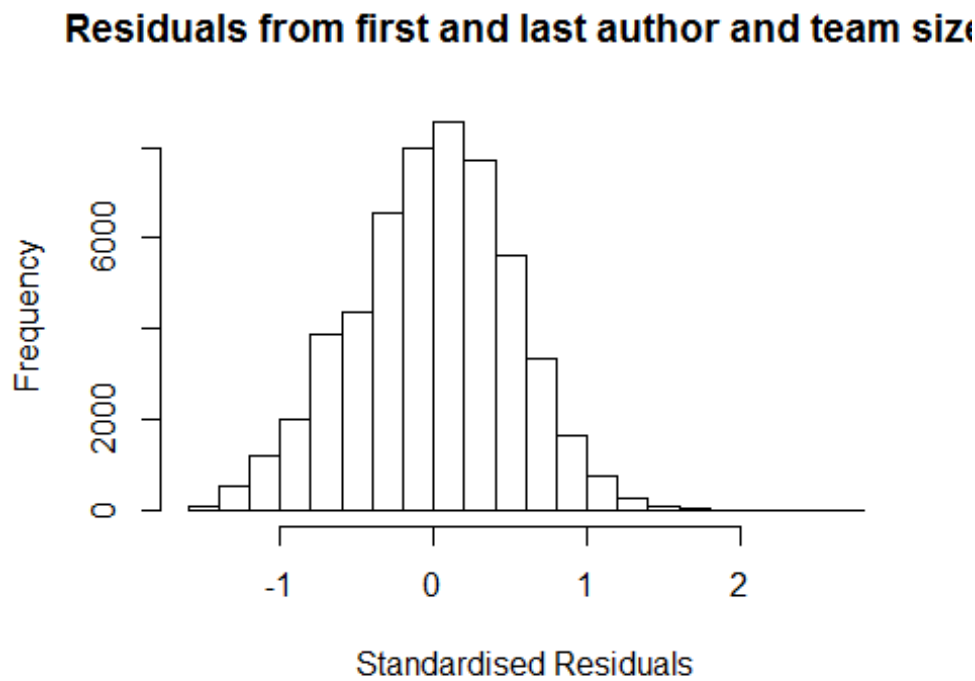
```
## 4756 4513
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 2459 2863 2489 2432 3071 2926 3162 2549 2762 3094 3469 3647 3923 3969 3878
## 2011 2012
## 4085 3835
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
##  Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 510, df = 16, p-value <2e-16
```




```
##
##  Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 51, df = 1, p-value = 8e-13
```

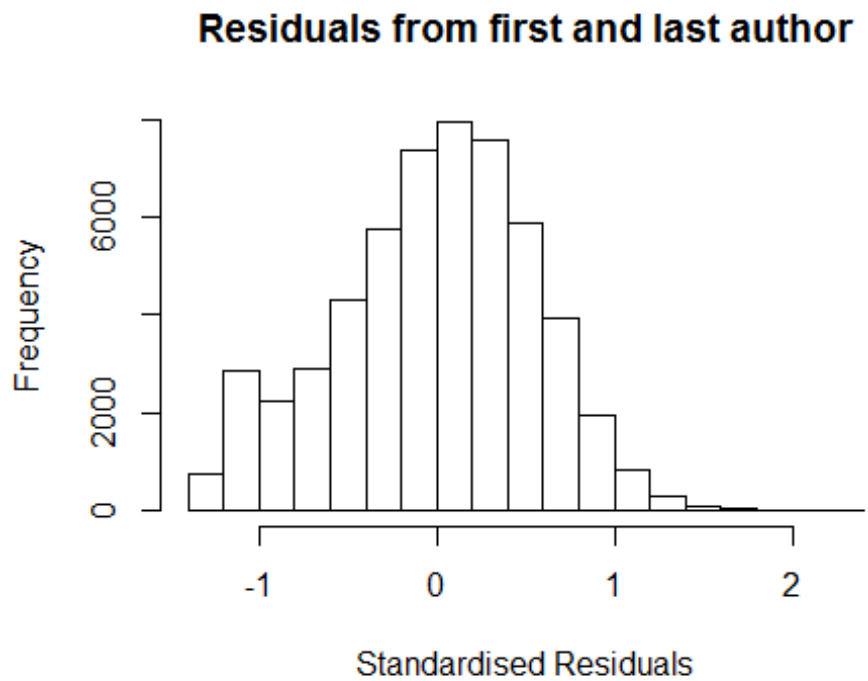



```
## [1] "Female first author team size 2018 geometric mean: 5.56921523333413"
## [1] "Male first author team size 2018 geometric mean: 5.57527229786139"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 1500000, p-value = 0.5
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 5.61327695732691"
## [1] "Male last author team size 2018 geometric mean: 5.56348693940119"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 1200000, p-value = 0.5
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##                     GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.036  1           1.018
## LastAuthorFemale  1.027  1           1.013
## UniqueAuthors     1.020  4           1.002
## Year              1.026 16           1.001
```



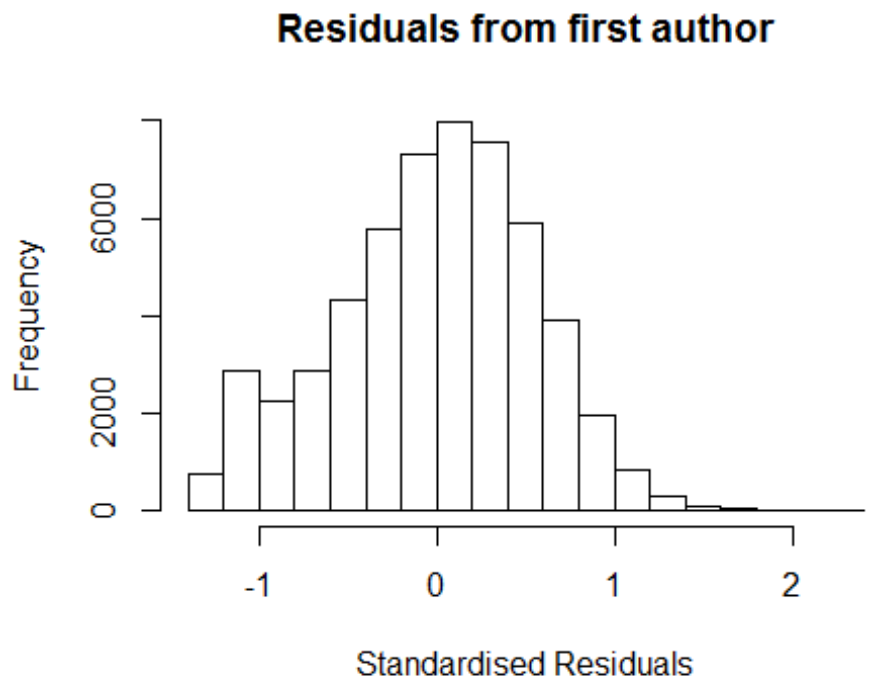

```
## [1] "List of  1 outliers with residuals above  2.5"
##         ScopusId  NLCS Year OneField Fields residuals
## 10576 0030912190 3.457 1997     2705      3     2.798
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
UniqueAuthors +
##     Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##     k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.5079 -0.3467  0.0175  0.3475  2.7977
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)          0.78968    0.01476   53.51  < 2e-16 ***
## FirstAuthorFemale1   0.03630    0.00551    6.59  4.5e-11 ***
## LastAuthorFemale1    0.03972    0.00669    5.94  3.0e-09 ***
## UniqueAuthors2       0.22871    0.01150   19.88  < 2e-16 ***
## UniqueAuthors3       0.32766    0.01089   30.08  < 2e-16 ***
## UniqueAuthors4       0.42871    0.01066   40.23  < 2e-16 ***
## UniqueAuthors5       0.67849    0.00912   74.40  < 2e-16 ***
## Year1997            -0.13033    0.01684   -7.74  1.0e-14 ***
## Year1998            -0.04112    0.01720   -2.39    0.017 *
## Year1999            -0.08358    0.01644   -5.09  3.7e-07 ***
```

```
## Year2000           -0.08478    0.01524   -5.56  2.7e-08 ***
## Year2001           -0.07628    0.01512   -5.05  4.5e-07 ***
## Year2002           -0.11875    0.01513   -7.85  4.2e-15 ***
## Year2003           -0.15099    0.01556   -9.70  < 2e-16 ***
## Year2004           -0.17393    0.01539  -11.30  < 2e-16 ***
## Year2005           -0.17786    0.01524  -11.67  < 2e-16 ***
## Year2006           -0.13784    0.01515   -9.10  < 2e-16 ***
## Year2007           -0.13619    0.01500   -9.08  < 2e-16 ***
## Year2008           -0.11980    0.01485   -8.07  7.4e-16 ***
## Year2009           -0.12689    0.01504   -8.44  < 2e-16 ***
## Year2010           -0.14875    0.01501   -9.91  < 2e-16 ***
## Year2011           -0.11368    0.01498   -7.59  3.3e-14 ***
## Year2012           -0.13830    0.01530   -9.04  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.511
## Multiple R-squared:  0.171,  Adjusted R-squared:  0.171
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  2 observations c(5052,51842) are outliers with |weight| = 0 ( < 1.8e-06);
##  4380 weights are ~= 1. The remaining 50231 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.017   0.866   0.950   0.906   0.985   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.83e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.023  1           1.011
## LastAuthorFemale  1.016  1           1.008
## Year              1.014 16           1.000
```



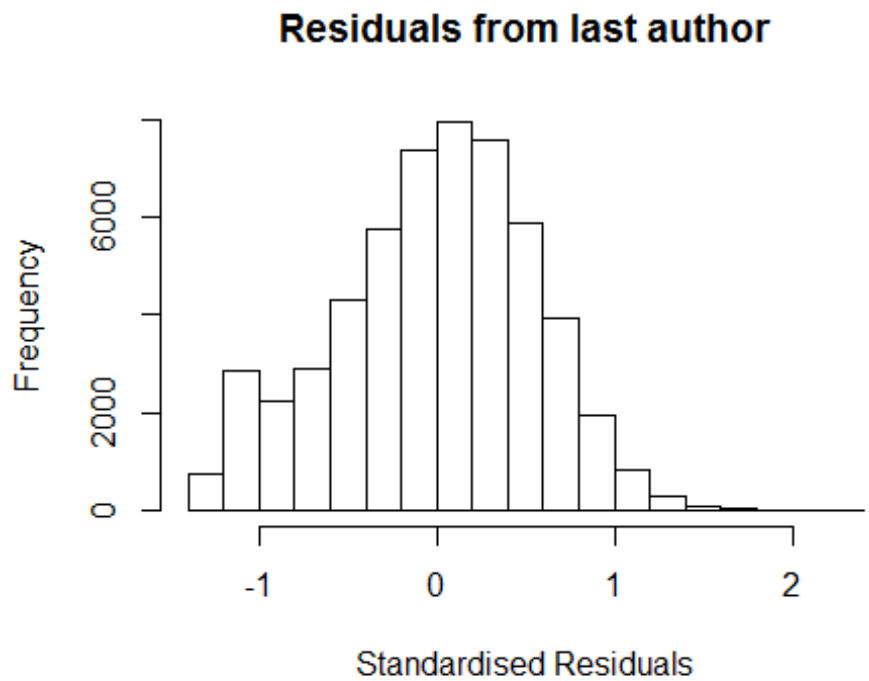

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##    Min     1Q Median     3Q    Max
## -1.322 -0.377  0.031  0.381  2.361
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)          1.24745    0.01298   96.08  < 2e-16 ***
## FirstAuthorFemale1   0.06263    0.00597   10.49  < 2e-16 ***
## LastAuthorFemale1    0.01209    0.00726    1.66  0.09618 .
## Year1997            -0.15172    0.01817   -8.35  < 2e-16 ***
## Year1998            -0.05420    0.01858   -2.92  0.00354 **
## Year1999            -0.08825    0.01781   -4.96  7.2e-07 ***
## Year2000            -0.08199    0.01637   -5.01  5.5e-07 ***
## Year2001            -0.04662    0.01614   -2.89  0.00386 **
## Year2002            -0.09387    0.01606   -5.85  5.1e-09 ***
## Year2003            -0.11290    0.01658   -6.81  9.9e-12 ***
## Year2004            -0.13991    0.01680   -8.33  < 2e-16 ***
## Year2005            -0.14058    0.01634   -8.60  < 2e-16 ***
```

```
## Year2006           -0.10949    0.01632   -6.71  2.0e-11 ***
## Year2007           -0.10678    0.01597   -6.69  2.3e-11 ***
## Year2008           -0.08703    0.01595   -5.46  4.9e-08 ***
## Year2009           -0.08052    0.01611   -5.00  5.8e-07 ***
## Year2010           -0.09285    0.01605   -5.78  7.3e-09 ***
## Year2011           -0.06018    0.01597   -3.77  0.00016 ***
## Year2012           -0.08126    0.01623   -5.01  5.5e-07 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.553
## Multiple R-squared:  0.00565,    Adjusted R-squared:  0.00533
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  4532 weights are ~= 1. The remaining 50081 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0282  0.8660  0.9490  0.9040  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.83e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.011  1           1.005
## Year              1.011 16           1.000
```




```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.3129 -0.3773  0.0313  0.3808  2.3602
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)          1.24844    0.01296   96.31  < 2e-16 ***
## FirstAuthorFemale1   0.06442    0.00595   10.83  < 2e-16 ***
## Year1997            -0.15162    0.01817   -8.35  < 2e-16 ***
## Year1998            -0.05409    0.01858   -2.91  0.00361 **
## Year1999            -0.08810    0.01780   -4.95  7.5e-07 ***
## Year2000            -0.08173    0.01637   -4.99  6.0e-07 ***
## Year2001            -0.04629    0.01613   -2.87  0.00411 **
## Year2002            -0.09369    0.01606   -5.84  5.4e-09 ***
## Year2003            -0.11264    0.01657   -6.80  1.1e-11 ***
## Year2004            -0.13957    0.01680   -8.31  < 2e-16 ***
## Year2005            -0.14032    0.01634   -8.59  < 2e-16 ***
## Year2006            -0.10928    0.01631   -6.70  2.1e-11 ***
```

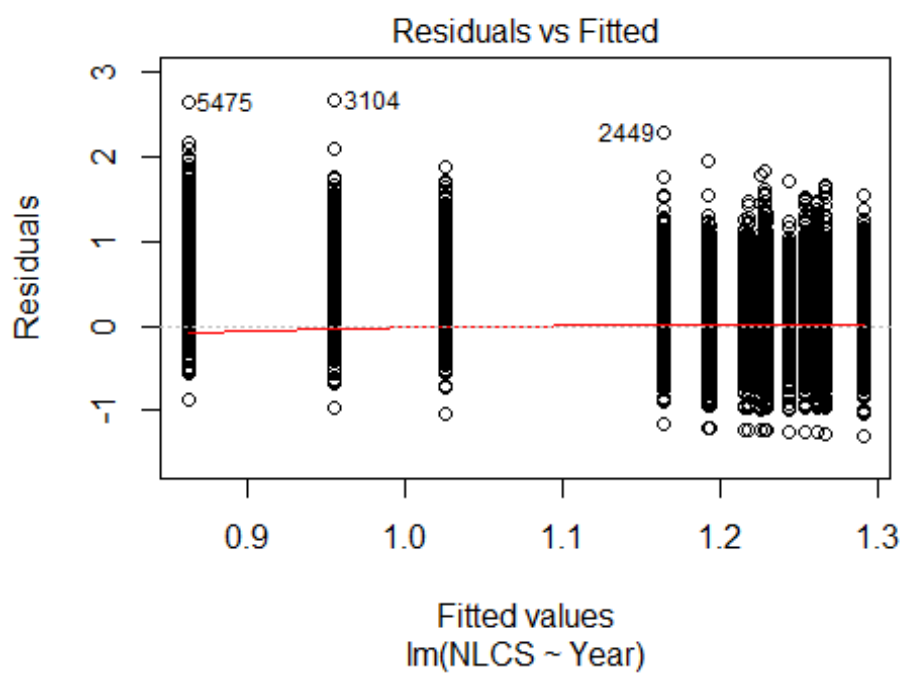
```
## Year2007           -0.10663    0.01597   -6.68  2.4e-11 ***
## Year2008           -0.08666    0.01595   -5.43  5.5e-08 ***
## Year2009           -0.08013    0.01610   -4.98  6.5e-07 ***
## Year2010           -0.09243    0.01604   -5.76  8.4e-09 ***
## Year2011           -0.05966    0.01596   -3.74  0.00019 ***
## Year2012           -0.08090    0.01622   -4.99  6.1e-07 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.553
## Multiple R-squared:  0.00559,    Adjusted R-squared:  0.00528
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  4514 weights are ~= 1. The remaining 50099 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0285  0.8650  0.9490  0.9040  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.83e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##                   GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.004  1           1.002
## Year             1.004 16           1.000
```



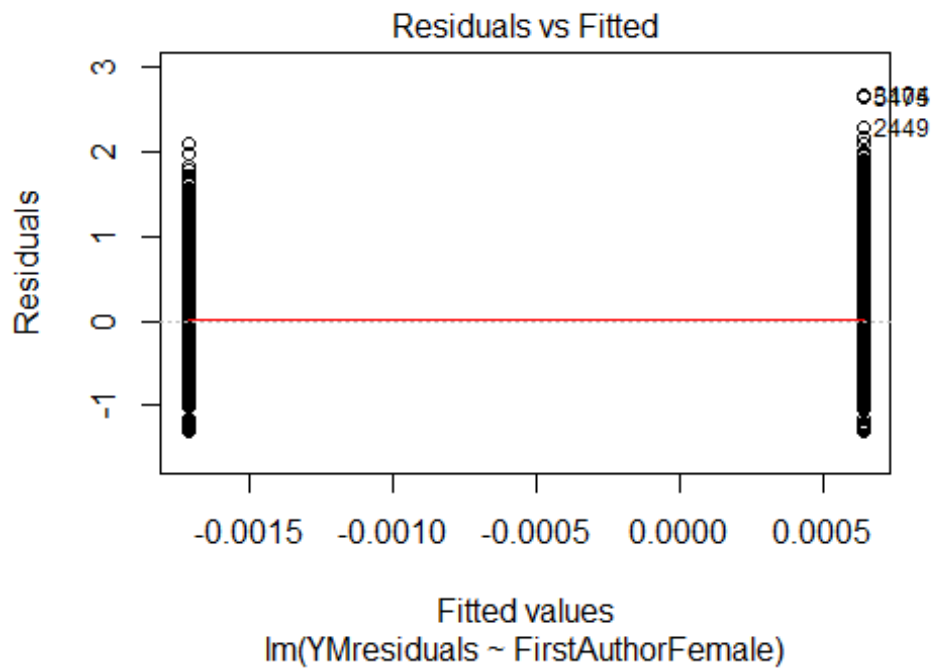

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.2799 -0.3774  0.0319  0.3822  2.3530
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)          1.2548     0.0130   96.70  < 2e-16 ***
## LastAuthorFemale1    0.0251     0.0072    3.49  0.00049 ***
## Year1997            -0.1509     0.0182   -8.29  < 2e-16 ***
## Year1998            -0.0531     0.0186   -2.86  0.00428 **
## Year1999            -0.0857     0.0178   -4.80  1.6e-06 ***
## Year2000            -0.0795     0.0164   -4.85  1.2e-06 ***
## Year2001            -0.0451     0.0162   -2.79  0.00523 **
## Year2002            -0.0912     0.0161   -5.67  1.4e-08 ***
## Year2003            -0.1095     0.0166   -6.60  4.2e-11 ***
## Year2004            -0.1358     0.0168   -8.07  7.0e-16 ***
## Year2005            -0.1365     0.0164   -8.34  < 2e-16 ***
## Year2006            -0.1052     0.0163   -6.44  1.2e-10 ***
```

```
## Year2007           -0.1033     0.0160   -6.46  1.1e-10 ***
## Year2008           -0.0812     0.0159   -5.09  3.5e-07 ***
## Year2009           -0.0748     0.0161   -4.64  3.5e-06 ***
## Year2010           -0.0868     0.0161   -5.40  6.7e-08 ***
## Year2011           -0.0534     0.0160   -3.35  0.00081 ***
## Year2012           -0.0750     0.0162   -4.62  3.8e-06 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.553
## Multiple R-squared:  0.00374,    Adjusted R-squared:  0.00342
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  4492 weights are ~= 1. The remaining 50121 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0306  0.8640  0.9490  0.9040  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.83e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Sample size for the above analysis:  54613"
## [1] ""
## [1] ""
## [1] "####################################"
## [1] "Analysis of AJSC 2706"
## [1] "####################################"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 1388 1022 1347 1760 1105 1050 1104  891  754  811  797 1031 1153 1209 1268
## 2011 2012
## 1418 1333
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  977  738  952 1346  742  635  909  733  619  676  645  866  944 1030 1068
## 2011 2012
```

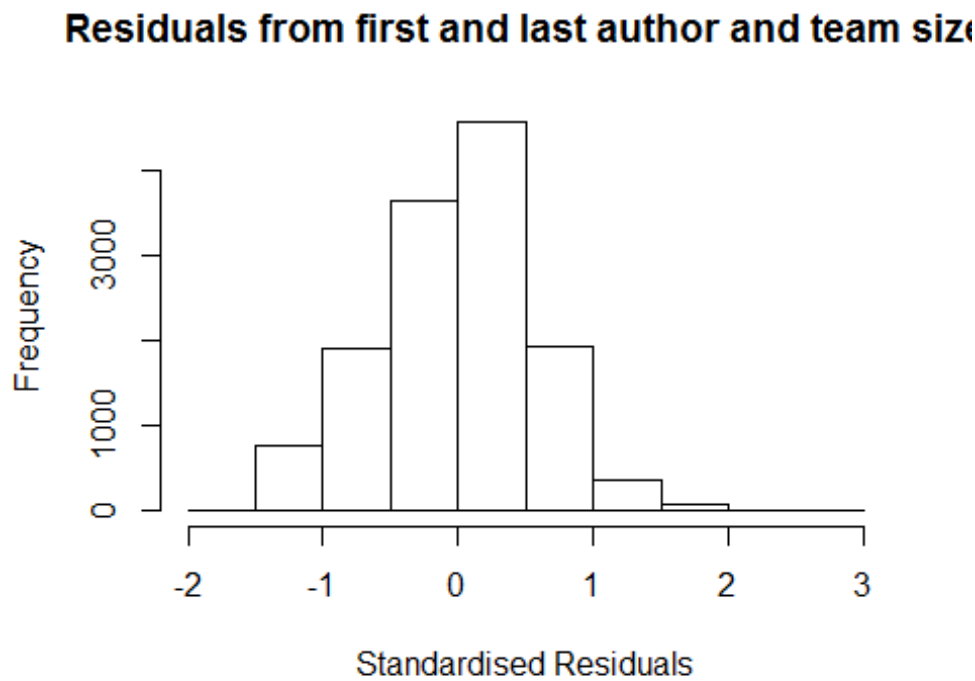
```
## 1175 1088
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  730  670  740 1190  659  550  800  657  550  589  576  776  832  922  947
## 2011 2012
## 1022  961
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
##  Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 860, df = 16, p-value <2e-16
```




```
##
##  Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 12, df = 1, p-value = 5e-04
```

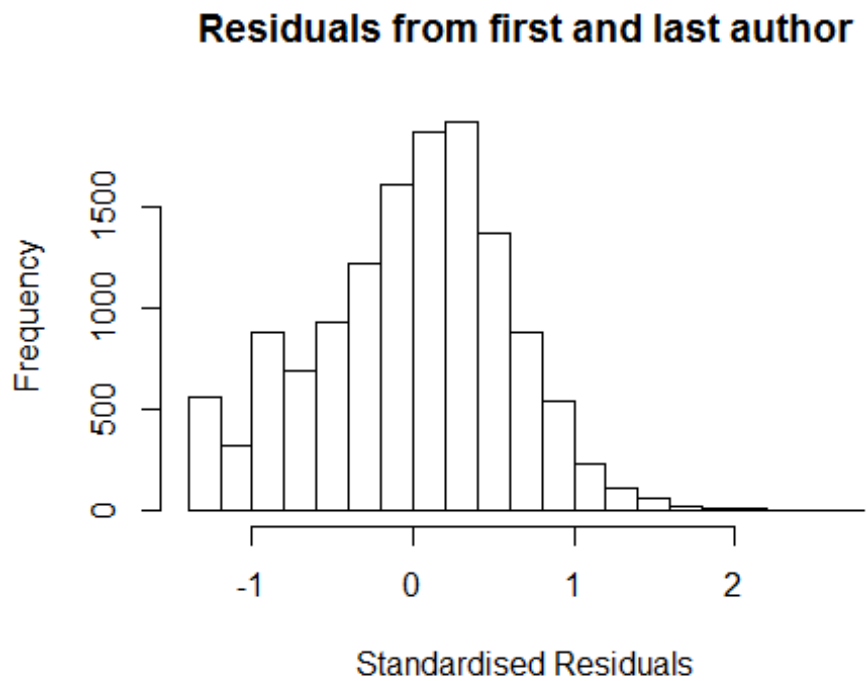



```
## [1] "Female first author team size 2018 geometric mean: 4.55771283092137"
## [1] "Male first author team size 2018 geometric mean: 4.42186517037694"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 94000, p-value = 0.7
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 4.40981010033401"
## [1] "Male last author team size 2018 geometric mean: 4.50718954994344"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 82000, p-value = 0.6
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##                     GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.078  1           1.038
## LastAuthorFemale  1.087  1           1.042
## UniqueAuthors     1.107  4           1.013
## Year              1.106 16           1.003
```



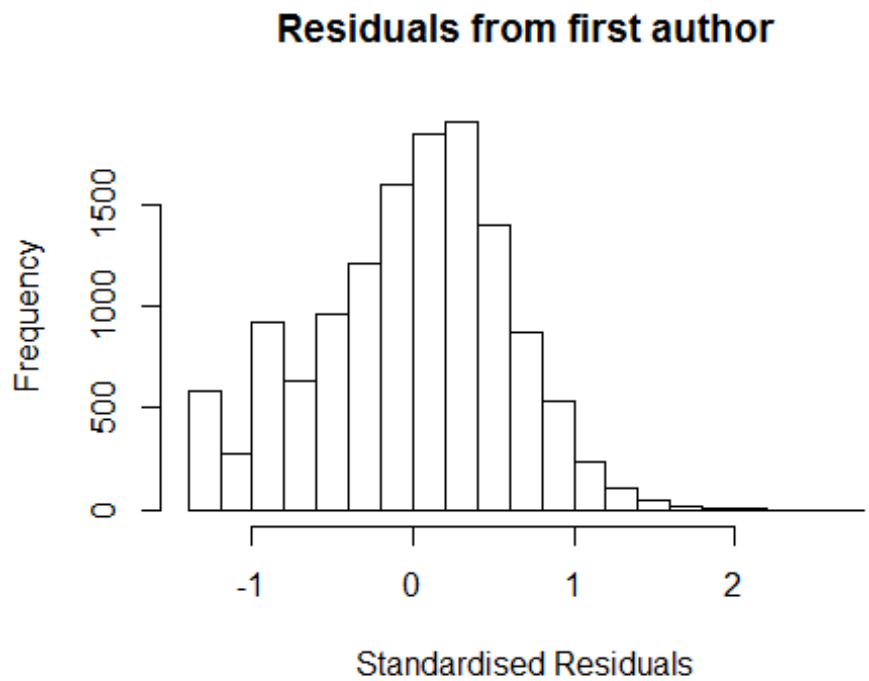

```
## [1] "List of  3 outliers with residuals above  2.5"
##        ScopusId  NLCS Year OneField Fields residuals
## 2449 0030912190 3.457 1997     2705      3     2.678
## 3104 0031761535 3.615 1998     2700      3     2.677
## 5475 0032926069 3.501 1999     2706      2     2.769
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
UniqueAuthors +
##     Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##     k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.5040 -0.3858  0.0307  0.3715  2.7693
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         0.86180    0.03238   26.62  < 2e-16 ***
## FirstAuthorFemale1  0.00553    0.01157    0.48  0.63290
## LastAuthorFemale1   0.00970    0.01343    0.72  0.46980
## UniqueAuthors2      0.21058    0.02646    7.96  1.9e-15 ***
## UniqueAuthors3      0.33234    0.02427   13.69  < 2e-16 ***
## UniqueAuthors4      0.42604    0.02372   17.96  < 2e-16 ***
## UniqueAuthors5      0.63247    0.02099   30.13  < 2e-16 ***
## Year1997           -0.08319    0.03419   -2.43  0.01498 *
```

```
## Year1998           -0.13434    0.04343   -3.09  0.00199 **
## Year1999           -0.47216    0.04322  -10.92  < 2e-16 ***
## Year2000           -0.06740    0.03177   -2.12  0.03392 *
## Year2001           -0.07104    0.03347   -2.12  0.03382 *
## Year2002           -0.09082    0.03125   -2.91  0.00367 **
## Year2003           -0.09604    0.03252   -2.95  0.00315 **
## Year2004           -0.07167    0.03269   -2.19  0.02838 *
## Year2005           -0.11929    0.03347   -3.56  0.00037 ***
## Year2006           -0.05419    0.03268   -1.66  0.09726 .
## Year2007           -0.02760    0.03187   -0.87  0.38654
## Year2008           -0.06768    0.03206   -2.11  0.03479 *
## Year2009           -0.08264    0.03111   -2.66  0.00790 **
## Year2010           -0.11265    0.03175   -3.55  0.00039 ***
## Year2011           -0.05056    0.03084   -1.64  0.10115
## Year2012           -0.06624    0.03174   -2.09  0.03694 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.531
## Multiple R-squared:  0.155,  Adjusted R-squared:  0.153
## Convergence in 17 IRWLS iterations
##
## Robustness weights:
##  3 observations c(1246,1612,2942)
##   are outliers with |weight| = 0 ( < 7.6e-06);
##  1118 weights are ~= 1. The remaining 12050 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0222  0.8460  0.9450  0.8950  0.9840  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          7.59e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.053  1           1.026
## LastAuthorFemale  1.043  1           1.021
## Year              1.026 16           1.001
```



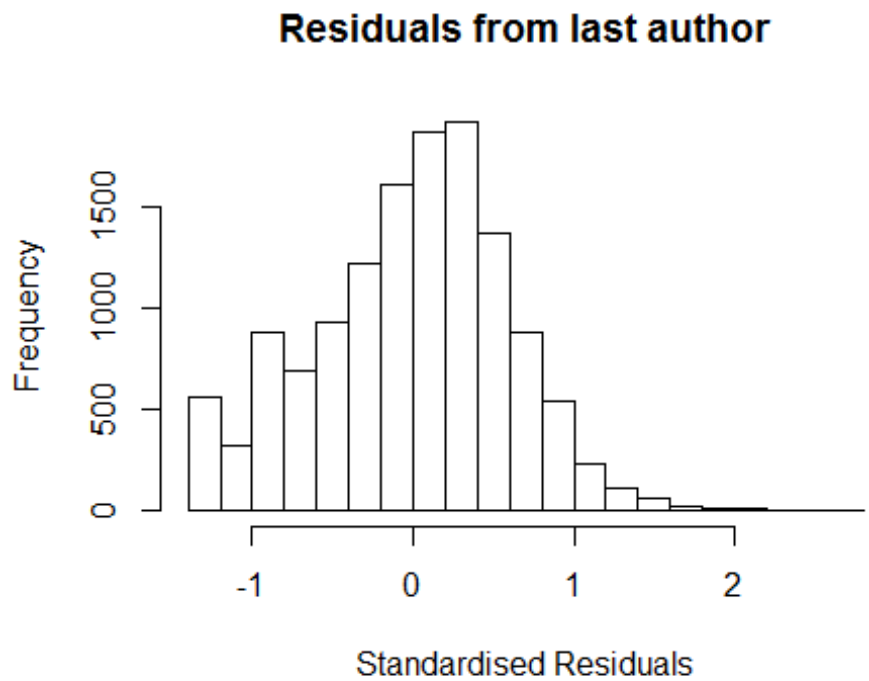

```
## [1] "List of  1 outliers with residuals above  2.5"
##        ScopusId  NLCS Year OneField Fields residuals
## 5475 0032926069 3.501 1999     2706      2     2.738
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.3195 -0.4138  0.0467  0.3927  2.7384
##
## Coefficients:
##                      Estimate Std. Error t value Pr(>|t|)
## (Intercept)          1.273956   0.029796   42.76  < 2e-16 ***
## FirstAuthorFemale1   0.000381   0.012236    0.03  0.97519
## LastAuthorFemale1   -0.055281   0.014304   -3.86  0.00011 ***
## Year1997            -0.083373   0.038983   -2.14  0.03248 *
## Year1998            -0.152623   0.047726   -3.20  0.00139 **
## Year1999            -0.456085   0.045252  -10.08  < 2e-16 ***
## Year2000            -0.020336   0.034931   -0.58  0.56046
## Year2001            -0.019106   0.036939   -0.52  0.60501
## Year2002            -0.042567   0.034907   -1.22  0.22269
## Year2003            -0.032665   0.036368   -0.90  0.36912
## Year2004            -0.023275   0.036629   -0.64  0.52517
## Year2005            -0.068776   0.037104   -1.85  0.06382 .
```

```
## Year2006            0.008595   0.036767    0.23  0.81518
## Year2007            0.045140   0.035384    1.28  0.20207
## Year2008            0.007159   0.035646    0.20  0.84083
## Year2009            0.016497   0.035050    0.47  0.63788
## Year2010           -0.023492   0.035488   -0.66  0.50800
## Year2011            0.031146   0.034171    0.91  0.36206
## Year2012            0.030921   0.035253    0.88  0.38044
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.551
## Multiple R-squared:  0.0459, Adjusted R-squared:  0.0446
## Convergence in 17 IRWLS iterations
##
## Robustness weights:
##  observation 2942 is an outlier with |weight| = 0 ( < 7.6e-06);
##  1017 weights are ~= 1. The remaining 12153 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0046  0.8360  0.9430  0.8920  0.9830  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          7.59e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.018  1           1.009
## Year              1.018 16           1.001
```




```
## [1] "List of  1 outliers with residuals above  2.5"
##        ScopusId  NLCS Year OneField Fields residuals
## 5475 0032926069 3.501 1999     2706      2     2.738
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.3104 -0.4176  0.0462  0.3935  2.6902
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.26740    0.02963   42.77   <2e-16 ***
## FirstAuthorFemale1 -0.00960    0.01208   -0.79   0.4269
## Year1997           -0.08337    0.03894   -2.14   0.0323 *
## Year1998           -0.15252    0.04772   -3.20   0.0014 **
## Year1999           -0.45664    0.04521  -10.10   <2e-16 ***
## Year2000           -0.02059    0.03484   -0.59   0.5546
## Year2001           -0.01964    0.03687   -0.53   0.5941
## Year2002           -0.04363    0.03484   -1.25   0.2105
## Year2003           -0.03190    0.03628   -0.88   0.3792
## Year2004           -0.02478    0.03661   -0.68   0.4986
## Year2005           -0.06961    0.03707   -1.88   0.0604 .
## Year2006            0.00706    0.03677    0.19   0.8478
```

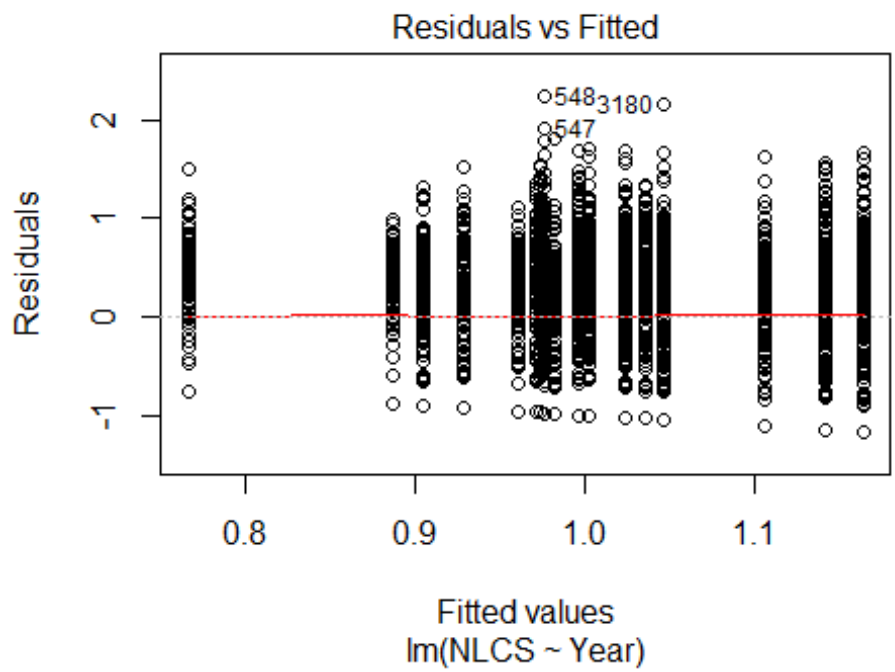
```
## Year2007            0.04297    0.03535    1.22   0.2241
## Year2008            0.00631    0.03557    0.18   0.8592
## Year2009            0.01515    0.03501    0.43   0.6653
## Year2010           -0.02435    0.03548   -0.69   0.4926
## Year2011            0.03091    0.03414    0.91   0.3652
## Year2012            0.02777    0.03520    0.79   0.4302
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.552
## Multiple R-squared:  0.0448, Adjusted R-squared:  0.0435
## Convergence in 17 IRWLS iterations
##
## Robustness weights:
##  observation 2942 is an outlier with |weight| = 0 ( < 7.6e-06);
##  1002 weights are ~= 1. The remaining 12168 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0041  0.8350  0.9430  0.8920  0.9830  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          7.59e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##                   GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.009  1           1.005
## Year             1.009 16           1.000
```



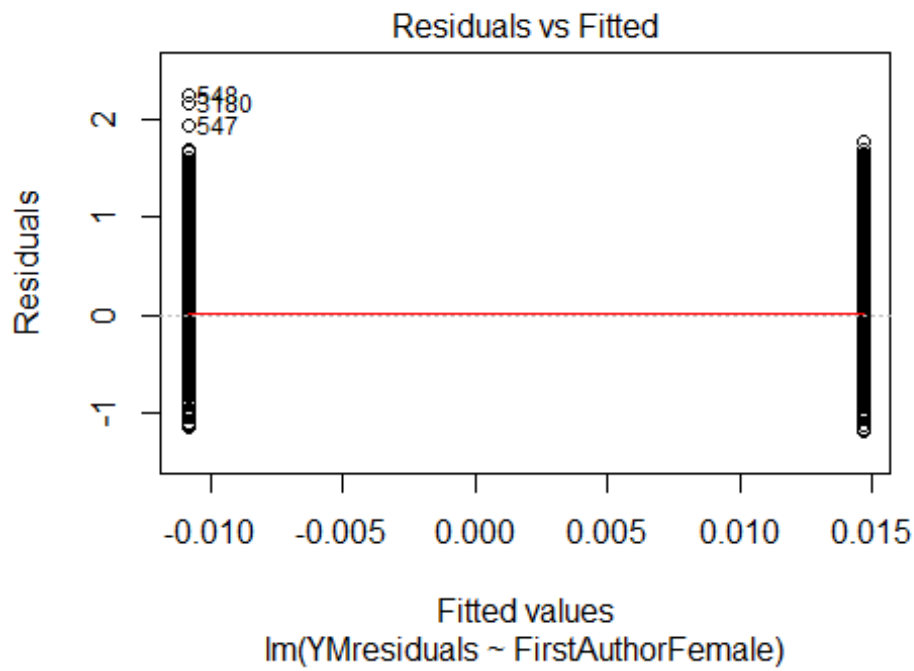

```
## [1] "List of  1 outliers with residuals above  2.5"
##        ScopusId  NLCS Year OneField Fields residuals
## 5475 0032926069 3.501 1999     2706      2     2.738
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##    Min     1Q Median     3Q    Max
## -1.319 -0.414  0.047  0.393  2.738
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.27401    0.02969   42.91  < 2e-16 ***
## LastAuthorFemale1 -0.05519    0.01407   -3.92  8.8e-05 ***
## Year1997          -0.08337    0.03899   -2.14   0.0325 *
## Year1998          -0.15261    0.04773   -3.20   0.0014 **
## Year1999          -0.45606    0.04527  -10.07  < 2e-16 ***
## Year2000          -0.02031    0.03494   -0.58   0.5609
## Year2001          -0.01907    0.03694   -0.52   0.6056
## Year2002          -0.04254    0.03491   -1.22   0.2230
## Year2003          -0.03265    0.03637   -0.90   0.3694
## Year2004          -0.02323    0.03662   -0.63   0.5258
## Year2005          -0.06874    0.03709   -1.85   0.0639 .
## Year2006           0.00863    0.03676    0.23   0.8145
```

```
## Year2007           0.04518    0.03537    1.28   0.2015
## Year2008           0.00721    0.03560    0.20   0.8395
## Year2009           0.01654    0.03505    0.47   0.6371
## Year2010          -0.02344    0.03547   -0.66   0.5087
## Year2011           0.03120    0.03416    0.91   0.3612
## Year2012           0.03098    0.03523    0.88   0.3792
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.551
## Multiple R-squared:  0.0459, Adjusted R-squared:  0.0446
## Convergence in 17 IRWLS iterations
##
## Robustness weights:
##  observation 2942 is an outlier with |weight| = 0 ( < 7.6e-06);
##  1016 weights are ~= 1. The remaining 12154 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0046  0.8360  0.9430  0.8920  0.9830  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          7.59e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Sample size for the above analysis:  13171"
## [1] ""
## [1] ""
## [1] "####################################"
## [1] "Analysis of AJSC 2707"
## [1] "####################################"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  221  253  211  235  238  273  291  253  246  311  337  343  399  411  438
## 2011 2012
##  467  455
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  156  150  128  138  124  124  217  191  191  240  261  253  317  320  345
```

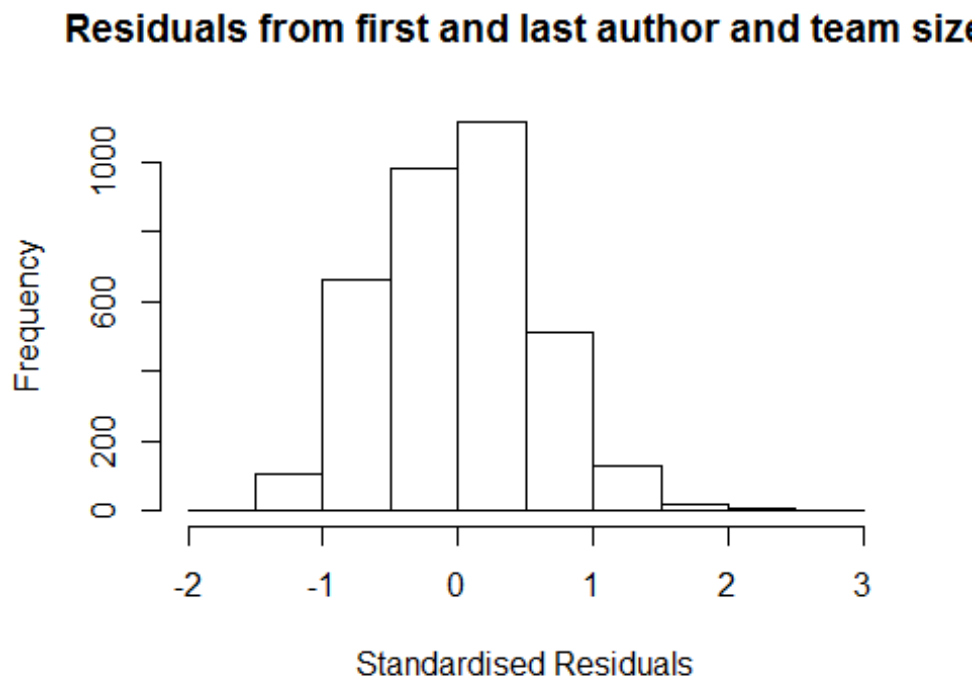
```
## 2011 2012
##  363  364
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  142  142  121  123  115  117  196  169  173  211  240  229  293  281  326
## 2011 2012
##  330  327
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
##  Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 40, df = 16, p-value = 8e-04
```




```
##
##  Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 3.1, df = 1, p-value = 0.08
```

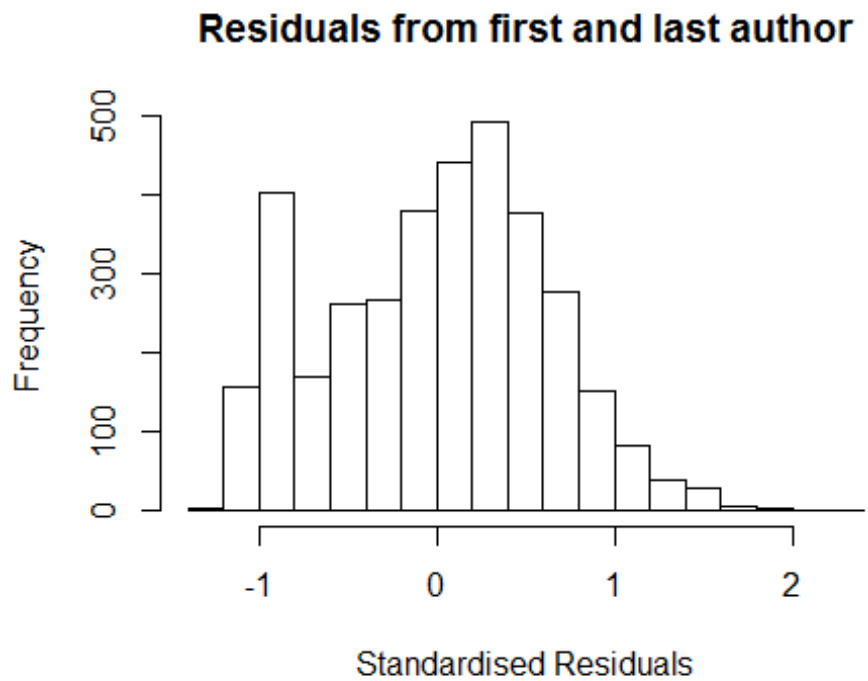



```
## [1] "Female first author team size 2018 geometric mean: 2.97176019340211"
## [1] "Male first author team size 2018 geometric mean: 2.54129907446022"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 13000, p-value = 0.08
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 2.84425488258167"
## [1] "Male last author team size 2018 geometric mean: 2.67927930030127"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 12000, p-value = 0.5
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.197  1           1.094
## LastAuthorFemale  1.211  1           1.100
## UniqueAuthors     1.177  4           1.021
## Year              1.187 16           1.005
```



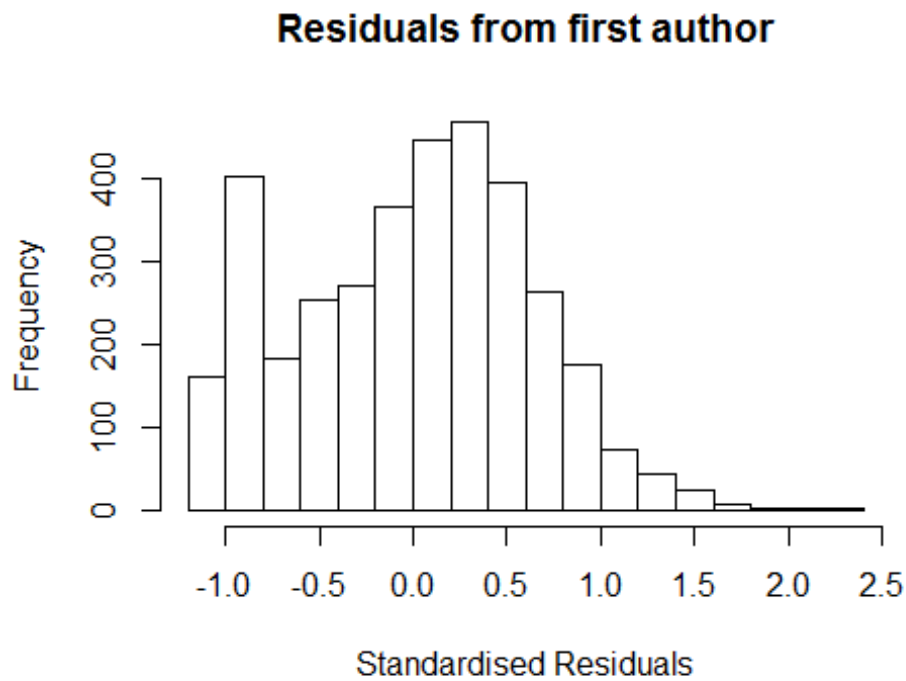

```
## [1] "List of  1 outliers with residuals above  2.5"
##       ScopusId  NLCS Year OneField Fields residuals
## 548 0032239998 3.215 1998     2700      2     2.627
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
UniqueAuthors +
##     Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##     k.max = 1000))
##  \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.52772 -0.41402  0.00506  0.39370  2.62749
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         0.59571    0.05410   11.01   <2e-16 ***
## FirstAuthorFemale1  0.04869    0.02129    2.29    0.022 *
## LastAuthorFemale1  -0.01220    0.02203   -0.55    0.580
## UniqueAuthors2      0.36868    0.03268   11.28   <2e-16 ***
## UniqueAuthors3      0.55631    0.03295   16.88   <2e-16 ***
## UniqueAuthors4      0.62419    0.03436   18.16   <2e-16 ***
## UniqueAuthors5      0.77417    0.02605   29.72   <2e-16 ***
## Year1997            0.02765    0.07002    0.39    0.693
## Year1998           -0.00820    0.08052   -0.10    0.919
## Year1999           -0.12589    0.06733   -1.87    0.062 .
```

```
## Year2000            0.02854    0.07301    0.39    0.696
## Year2001           -0.15698    0.07997   -1.96    0.050 *
## Year2002           -0.11136    0.06441   -1.73    0.084 .
## Year2003           -0.05414    0.06781   -0.80    0.425
## Year2004            0.04507    0.06700    0.67    0.501
## Year2005            0.02263    0.06758    0.33    0.738
## Year2006            0.00558    0.06489    0.09    0.931
## Year2007            0.10915    0.06827    1.60    0.110
## Year2008            0.05723    0.06235    0.92    0.359
## Year2009           -0.03118    0.06151   -0.51    0.612
## Year2010           -0.05680    0.06079   -0.93    0.350
## Year2011           -0.03639    0.05866   -0.62    0.535
## Year2012           -0.08843    0.06199   -1.43    0.154
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.539
## Multiple R-squared:  0.253,  Adjusted R-squared:  0.248
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
##  observation 305 is an outlier with |weight| = 0 ( < 2.8e-05);
##  269 weights are ~= 1. The remaining 3265 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0267  0.8760  0.9400  0.9020  0.9840  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          2.83e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.187  1           1.090
## LastAuthorFemale  1.172  1           1.083
## Year              1.050 16           1.002
```



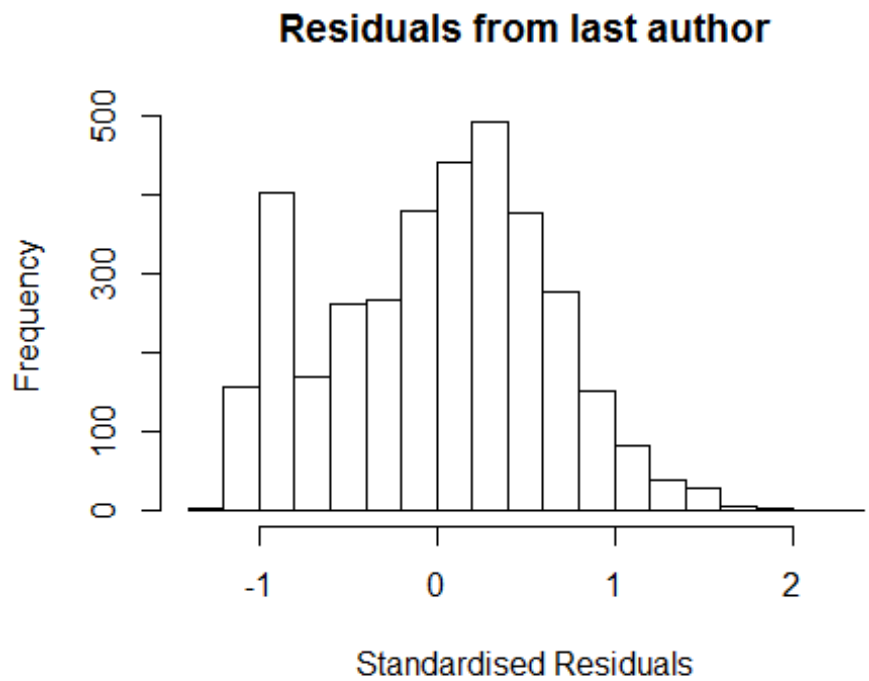

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.2272 -0.5030  0.0612  0.4397  2.2778
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)           0.9138     0.0548   16.66  < 2e-16 ***
## FirstAuthorFemale1    0.0654     0.0246    2.65  0.00798 **
## LastAuthorFemale1    -0.0959     0.0253   -3.80  0.00015 ***
## Year1997              0.0432     0.0764    0.56  0.57227
## Year1998              0.0235     0.0914    0.26  0.79740
## Year1999             -0.0514     0.0731   -0.70  0.48187
## Year2000              0.0598     0.0766    0.78  0.43532
## Year2001             -0.1631     0.0829   -1.97  0.04912 *
## Year2002             -0.0139     0.0715   -0.19  0.84572
## Year2003              0.0455     0.0732    0.62  0.53431
## Year2004              0.1970     0.0724    2.72  0.00654 **
## Year2005              0.1298     0.0725    1.79  0.07355 .
```

```
## Year2006             0.1154     0.0716    1.61  0.10715
## Year2007             0.2481     0.0753    3.29  0.00100 **
## Year2008             0.2150     0.0663    3.24  0.00119 **
## Year2009             0.0768     0.0677    1.14  0.25624
## Year2010             0.0800     0.0670    1.19  0.23265
## Year2011             0.0662     0.0648    1.02  0.30697
## Year2012             0.0421     0.0671    0.63  0.53080
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.65
## Multiple R-squared:  0.0238, Adjusted R-squared:  0.0188
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  285 weights are ~= 1. The remaining 3250 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.194   0.854   0.944   0.912   0.984   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          2.83e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.032  1           1.016
## Year              1.032 16           1.001
```




```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.1719 -0.5127  0.0652  0.4438  2.2926
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)           0.9020     0.0551   16.36   <2e-16 ***
## FirstAuthorFemale1    0.0231     0.0233    0.99   0.3209
## Year1997              0.0390     0.0765    0.51   0.6103
## Year1998              0.0204     0.0916    0.22   0.8233
## Year1999             -0.0568     0.0739   -0.77   0.4421
## Year2000              0.0465     0.0771    0.60   0.5469
## Year2001             -0.1775     0.0835   -2.13   0.0335 *
## Year2002             -0.0279     0.0716   -0.39   0.6971
## Year2003              0.0290     0.0733    0.40   0.6928
## Year2004              0.1928     0.0730    2.64   0.0083 **
## Year2005              0.1236     0.0725    1.71   0.0882 .
## Year2006              0.1074     0.0718    1.50   0.1346
```

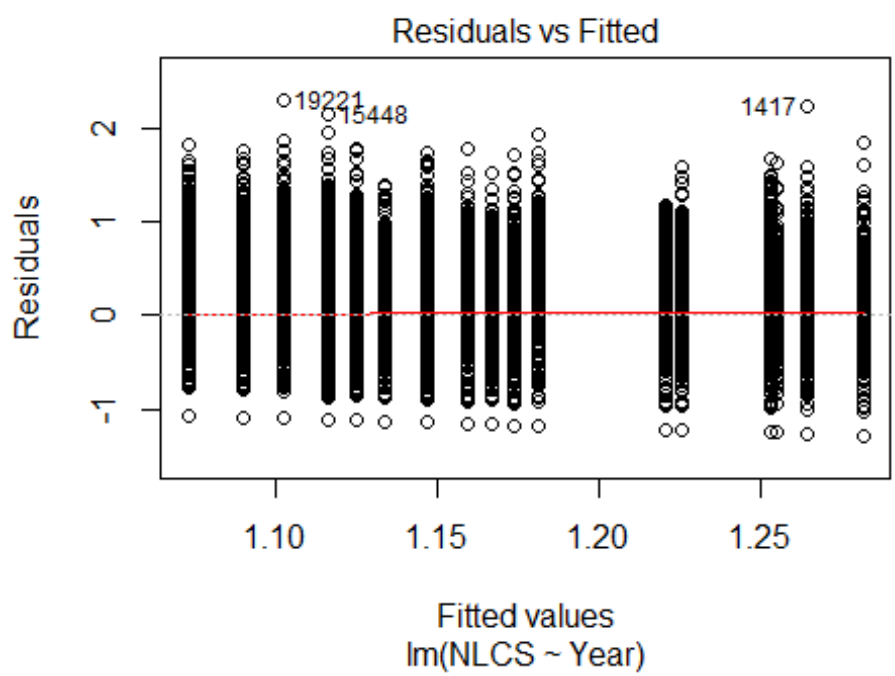
```
## Year2007             0.2468     0.0754    3.27   0.0011 **
## Year2008             0.2144     0.0666    3.22   0.0013 **
## Year2009             0.0719     0.0679    1.06   0.2902
## Year2010             0.0697     0.0675    1.03   0.3019
## Year2011             0.0601     0.0653    0.92   0.3575
## Year2012             0.0368     0.0674    0.55   0.5857
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.654
## Multiple R-squared:  0.0198, Adjusted R-squared:  0.0151
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  277 weights are ~= 1. The remaining 3258 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.194   0.853   0.944   0.914   0.985   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          2.83e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##                   GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.019  1           1.009
## Year             1.019 16           1.001
```



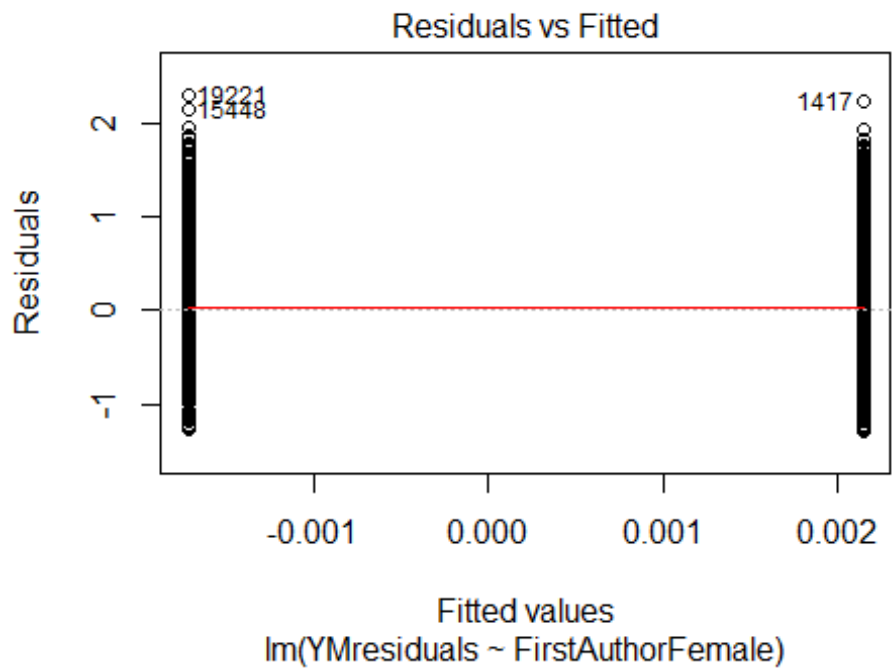

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.1785 -0.5067  0.0624  0.4399  2.2633
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)          0.9236     0.0546   16.92  < 2e-16 ***
## LastAuthorFemale1   -0.0658     0.0236   -2.79  0.00524 **
## Year1997             0.0423     0.0764    0.55  0.57978
## Year1998             0.0280     0.0918    0.31  0.76036
## Year1999            -0.0484     0.0730   -0.66  0.50697
## Year2000             0.0628     0.0764    0.82  0.41121
## Year2001            -0.1551     0.0826   -1.88  0.06047 .
## Year2002            -0.0136     0.0717   -0.19  0.84981
## Year2003             0.0549     0.0732    0.75  0.45285
## Year2004             0.2086     0.0722    2.89  0.00387 **
## Year2005             0.1400     0.0722    1.94  0.05259 .
## Year2006             0.1210     0.0716    1.69  0.09119 .
```

```
## Year2007            0.2549     0.0755    3.38  0.00074 ***
## Year2008            0.2236     0.0661    3.38  0.00073 ***
## Year2009            0.0845     0.0676    1.25  0.21121
## Year2010            0.0869     0.0669    1.30  0.19415
## Year2011            0.0746     0.0648    1.15  0.24971
## Year2012            0.0509     0.0670    0.76  0.44735
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.651
## Multiple R-squared:  0.022,  Adjusted R-squared:  0.0172
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  287 weights are ~= 1. The remaining 3248 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.202   0.853   0.943   0.913   0.985   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          2.83e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Sample size for the above analysis:  3535"
## [1] ""
## [1] ""
## [1] "####################################"
## [1] "Analysis of AJSC 2708"
## [1] "####################################"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 1195 1018  977  955 1104 1135 1032  926  963 1021 1159 1258 1563 1493 1509
## 2011 2012
## 1560 1539
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  753  465  439  488  711  742  797  712  762  819  930 1030 1248 1167 1154
## 2011 2012
```

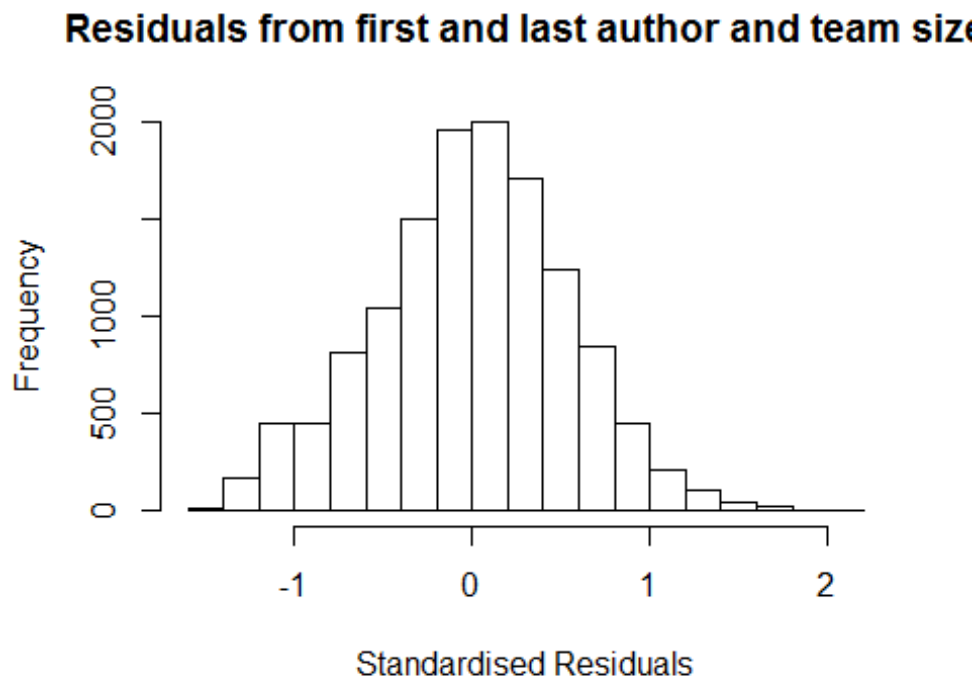
```
## 1242 1218
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  655  414  392  424  630  657  703  616  679  714  826  936 1117 1026 1027
## 2011 2012
## 1093 1089
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
##  Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 120, df = 16, p-value <2e-16
```




```
##
##  Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 0.5, df = 1, p-value = 0.5
```

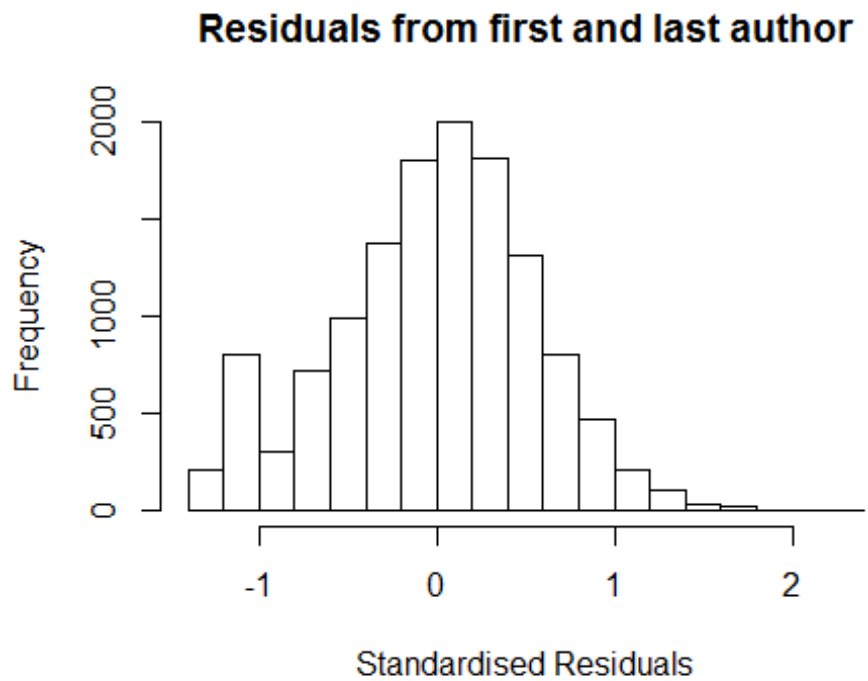



```
## [1] "Female first author team size 2018 geometric mean: 3.89170395862243"
## [1] "Male first author team size 2018 geometric mean: 3.70551169708213"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 210000, p-value = 0.2
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 3.7934642991187"
## [1] "Male last author team size 2018 geometric mean: 3.8098872684655"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 190000, p-value = 0.8
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.062  1           1.031
## LastAuthorFemale  1.049  1           1.024
## UniqueAuthors     1.073  4           1.009
## Year              1.077 16           1.002
```



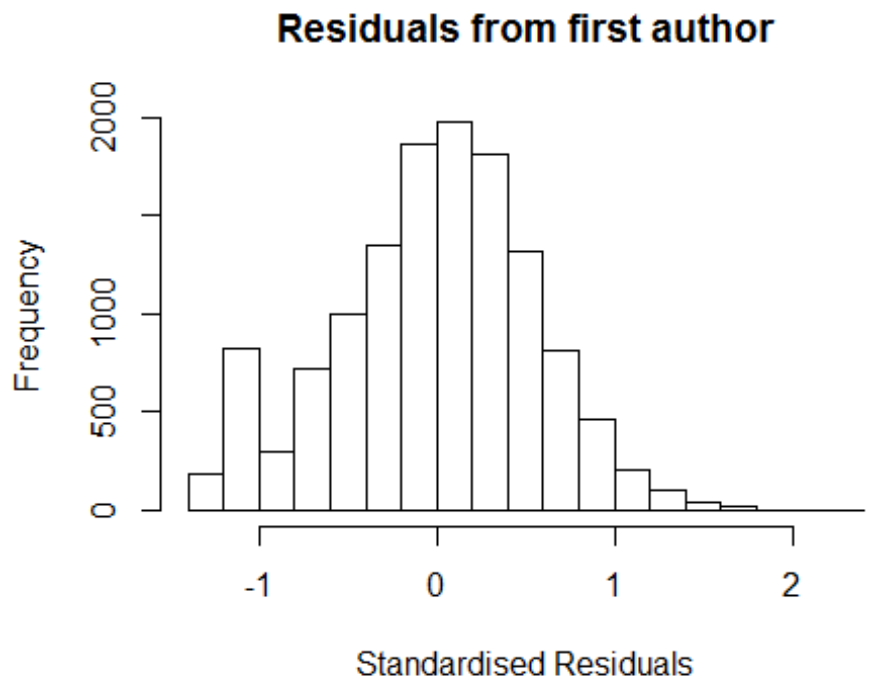

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
UniqueAuthors +
##     Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##     k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.5332 -0.3522  0.0146  0.3570  2.1970
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)           0.9877     0.0310   31.86  < 2e-16 ***
## FirstAuthorFemale1   -0.0337     0.0101   -3.32  0.00089 ***
## LastAuthorFemale1    -0.0250     0.0109   -2.29  0.02227 *
## UniqueAuthors2        0.2889     0.0215   13.46  < 2e-16 ***
## UniqueAuthors3        0.2974     0.0206   14.43  < 2e-16 ***
## UniqueAuthors4        0.3791     0.0210   18.07  < 2e-16 ***
## UniqueAuthors5        0.5455     0.0195   27.96  < 2e-16 ***
## Year1997             -0.0391     0.0394   -0.99  0.32119
## Year1998             -0.0500     0.0372   -1.35  0.17858
## Year1999             -0.0207     0.0359   -0.57  0.56531
```

```
## Year2000            -0.0698     0.0336   -2.08  0.03764 *
## Year2001            -0.0740     0.0325   -2.28  0.02269 *
## Year2002            -0.1040     0.0336   -3.10  0.00194 **
## Year2003            -0.1841     0.0337   -5.47  4.6e-08 ***
## Year2004            -0.1513     0.0331   -4.58  4.8e-06 ***
## Year2005            -0.1369     0.0335   -4.09  4.4e-05 ***
## Year2006            -0.1357     0.0340   -3.99  6.6e-05 ***
## Year2007            -0.1697     0.0328   -5.17  2.3e-07 ***
## Year2008            -0.1937     0.0315   -6.16  7.6e-10 ***
## Year2009            -0.1910     0.0322   -5.92  3.2e-09 ***
## Year2010            -0.2126     0.0325   -6.53  6.6e-11 ***
## Year2011            -0.2412     0.0318   -7.58  3.6e-14 ***
## Year2012            -0.2758     0.0329   -8.37  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.526
## Multiple R-squared:  0.098,  Adjusted R-squared:  0.0965
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
##  1095 weights are ~= 1. The remaining 11903 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0417  0.8600  0.9500  0.9010  0.9860  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          7.69e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.027  1           1.013
## LastAuthorFemale  1.034  1           1.017
## Year              1.031 16           1.001
```



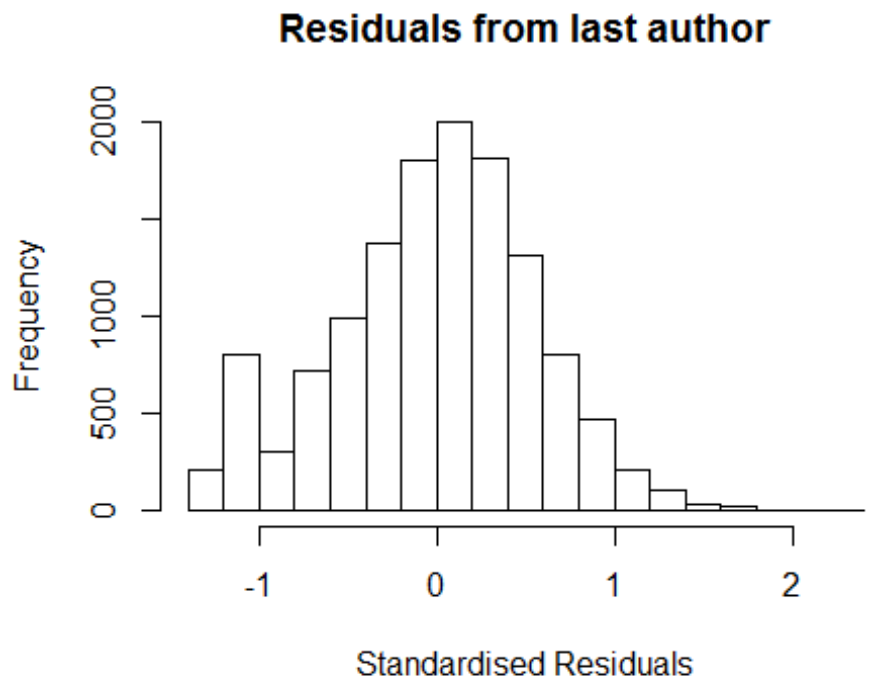

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.2964 -0.3608  0.0267  0.3642  2.2780
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)          1.28408    0.02716   47.27  < 2e-16 ***
## FirstAuthorFemale1  0.00791    0.01042    0.76  0.44788
## LastAuthorFemale1  -0.02700    0.01144   -2.36  0.01833 *
## Year1997           -0.00151    0.04035   -0.04  0.97018
## Year1998           -0.01794    0.03720   -0.48  0.62962
## Year1999            0.00440    0.03592    0.12  0.90248
## Year2000           -0.04884    0.03391   -1.44  0.14987
## Year2001           -0.06168    0.03300   -1.87  0.06165 .
## Year2002           -0.08884    0.03488   -2.55  0.01088 *
## Year2003           -0.14729    0.03442   -4.28  1.9e-05 ***
## Year2004           -0.11331    0.03418   -3.32  0.00092 ***
## Year2005           -0.09221    0.03438   -2.68  0.00733 **
```

```
## Year2006           -0.08963    0.03533   -2.54  0.01120 *
## Year2007           -0.13242    0.03326   -3.98  6.9e-05 ***
## Year2008           -0.15649    0.03181   -4.92  8.8e-07 ***
## Year2009           -0.15384    0.03275   -4.70  2.7e-06 ***
## Year2010           -0.16607    0.03291   -5.05  4.6e-07 ***
## Year2011           -0.18328    0.03232   -5.67  1.5e-08 ***
## Year2012           -0.21064    0.03342   -6.30  3.0e-10 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.536
## Multiple R-squared:  0.0132, Adjusted R-squared:  0.0118
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
##  1100 weights are ~= 1. The remaining 11898 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0318  0.8580  0.9490  0.8970  0.9860  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          7.69e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.013  1           1.007
## Year              1.013 16           1.000
```




```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.2856 -0.3586  0.0276  0.3621  2.2852
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)          1.27955    0.02707   47.26  < 2e-16 ***
## FirstAuthorFemale1  0.00346    0.01038    0.33  0.73886
## Year1997           -0.00092    0.04028   -0.02  0.98177
## Year1998           -0.01808    0.03717   -0.49  0.62660
## Year1999            0.00264    0.03588    0.07  0.94137
## Year2000           -0.04872    0.03388   -1.44  0.15047
## Year2001           -0.06230    0.03297   -1.89  0.05882 .
## Year2002           -0.09052    0.03481   -2.60  0.00932 **
## Year2003           -0.14879    0.03435   -4.33  1.5e-05 ***
## Year2004           -0.11400    0.03414   -3.34  0.00084 ***
## Year2005           -0.09451    0.03431   -2.75  0.00588 **
## Year2006           -0.09160    0.03526   -2.60  0.00939 **
```

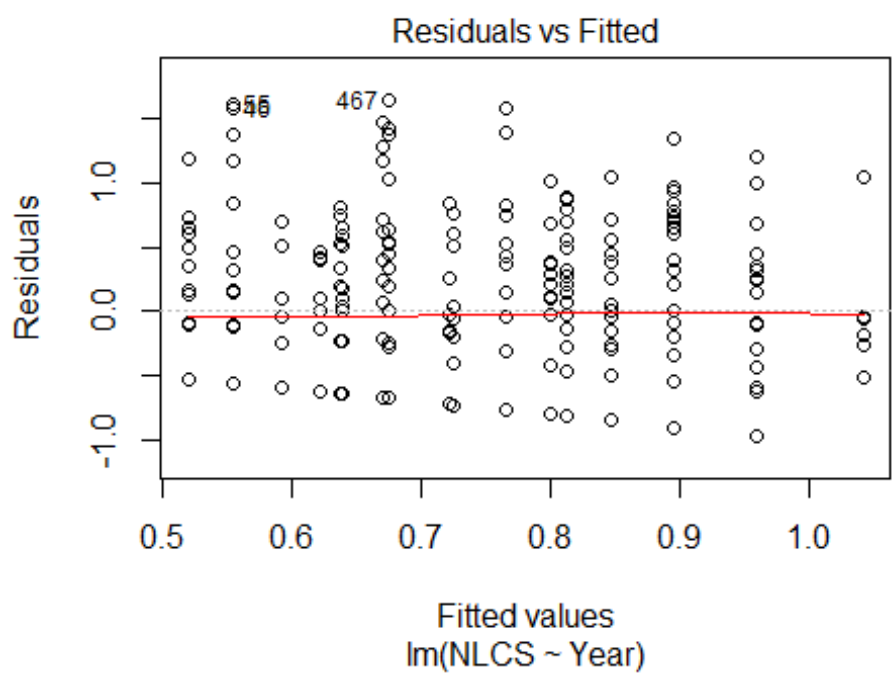
```
## Year2007           -0.13532    0.03321   -4.07  4.6e-05 ***
## Year2008           -0.15986    0.03171   -5.04  4.7e-07 ***
## Year2009           -0.15628    0.03269   -4.78  1.8e-06 ***
## Year2010           -0.16873    0.03285   -5.14  2.8e-07 ***
## Year2011           -0.18563    0.03225   -5.76  8.8e-09 ***
## Year2012           -0.21376    0.03334   -6.41  1.5e-10 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.536
## Multiple R-squared:  0.0127, Adjusted R-squared:  0.0114
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
##  1093 weights are ~= 1. The remaining 11905 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0299  0.8580  0.9490  0.8970  0.9860  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          7.69e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##                  GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.02  1           1.010
## Year             1.02 16           1.001
```




```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.2909 -0.3632  0.0278  0.3628  2.2746
##
## Coefficients:
##                   Estimate Std. Error t value Pr(>|t|)
## (Intercept)        1.28626    0.02695   47.73  < 2e-16 ***
## LastAuthorFemale1 -0.02550    0.01139   -2.24    0.025 *
## Year1997          -0.00110    0.04036   -0.03    0.978
## Year1998          -0.01766    0.03719   -0.47    0.635
## Year1999           0.00463    0.03594    0.13    0.897
## Year2000          -0.04871    0.03392   -1.44    0.151
## Year2001          -0.06109    0.03301   -1.85    0.064 .
## Year2002          -0.08808    0.03489   -2.52    0.012 *
## Year2003          -0.14659    0.03442   -4.26  2.1e-05 ***
## Year2004          -0.11224    0.03417   -3.28    0.001 **
## Year2005          -0.09126    0.03438   -2.65    0.008 **
## Year2006          -0.08887    0.03534   -2.51    0.012 *
```

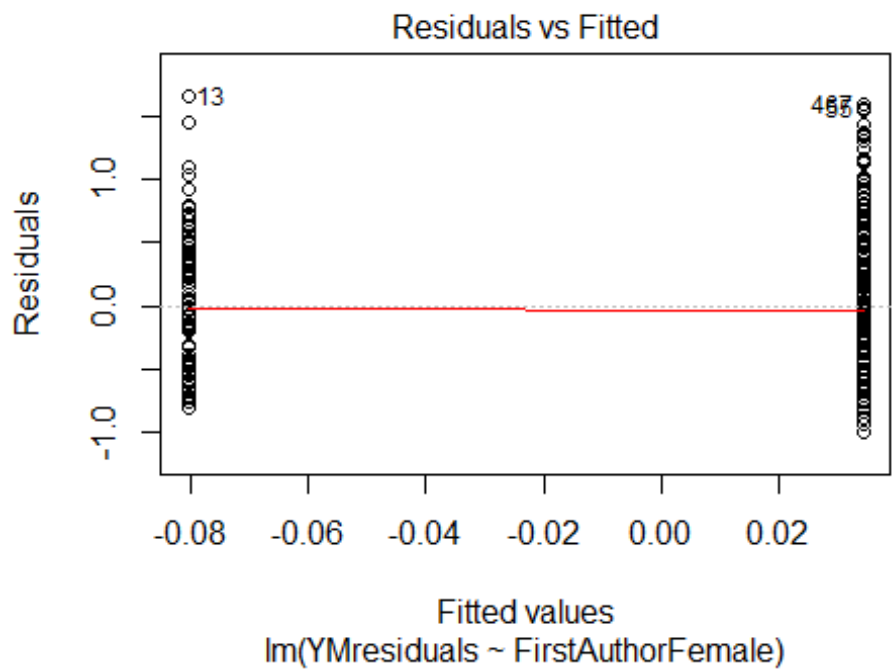
```
## Year2007          -0.13128    0.03324   -3.95  7.9e-05 ***
## Year2008          -0.15525    0.03179   -4.88  1.1e-06 ***
## Year2009          -0.15291    0.03277   -4.67  3.1e-06 ***
## Year2010          -0.16482    0.03291   -5.01  5.6e-07 ***
## Year2011          -0.18196    0.03230   -5.63  1.8e-08 ***
## Year2012          -0.20932    0.03342   -6.26  3.9e-10 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.536
## Multiple R-squared:  0.0131, Adjusted R-squared:  0.0118
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
##  1086 weights are ~= 1. The remaining 11912 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0327  0.8590  0.9490  0.8970  0.9860  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          7.69e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Sample size for the above analysis:  12998"
## [1] ""
## [1] ""
## [1] "####################################"
## [1] "Analysis of AJSC 2709"
## [1] "####################################"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##   19   47   28   38   63   45   12   12   10   11    9   30   25   18   14
## 2011 2012
##   40   32
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##   18   35   21   31   24   18   12    9    6   11    7   25   19   17   12
## 2011 2012
```

```
##   32   23
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##   18   32   21   29   22   16   12    9    6    9    6   22   18   12   10
## 2011 2012
##   28   20
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
##  Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 22, df = 16, p-value = 0.1
```

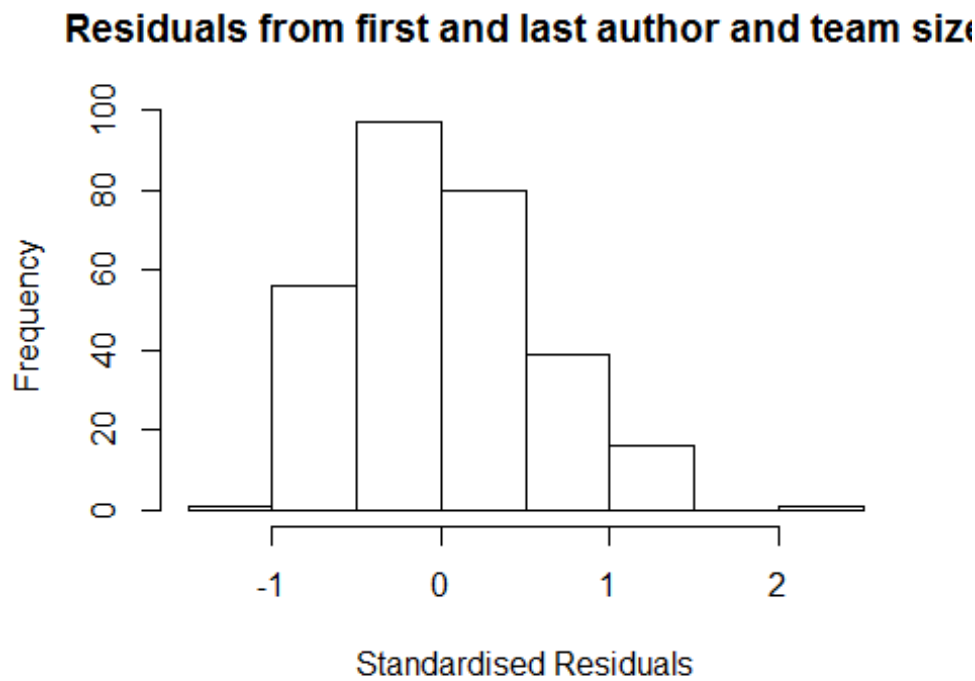



```
##
##  Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 2.3, df = 1, p-value = 0.1

## [1] "Female first author team size 2018 geometric mean: 2"
## [1] "Male first author team size 2018 geometric mean: 2"

## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties
```

```
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 0.5, p-value = NA
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 2"
## [1] "Male last author team size 2018 geometric mean: 2"

## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties
```

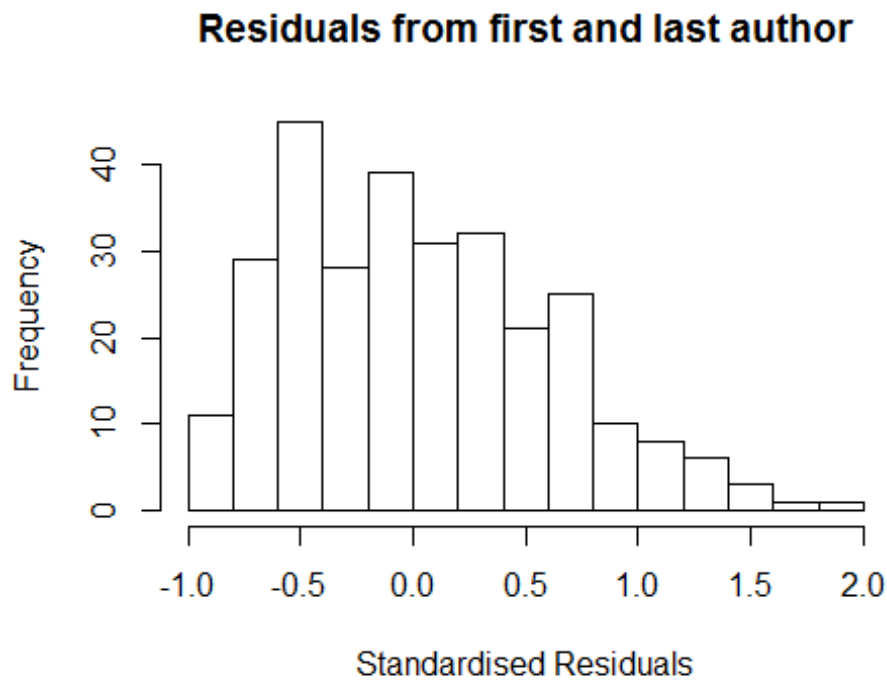



```
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 0.5, p-value = NA
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##                     GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.637  1           1.279
## LastAuthorFemale  1.679  1           1.296
## UniqueAuthors     2.339  4           1.112
## Year              3.249 16           1.038
```



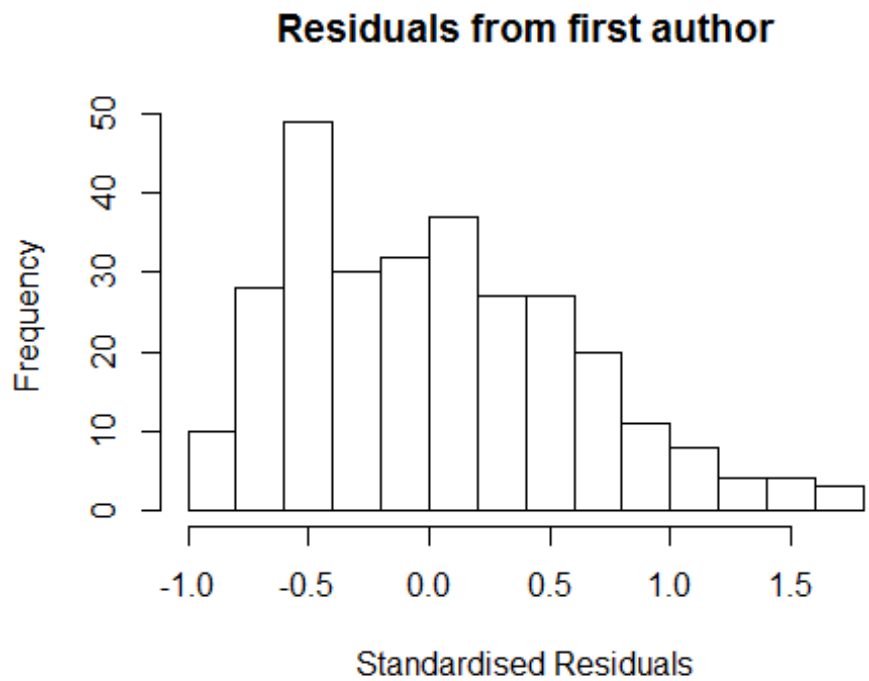

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
UniqueAuthors +
##     Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##     k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.2002 -0.4004 -0.0382  0.3819  2.0551
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)          0.5100     0.2080    2.45   0.0149 *
## FirstAuthorFemale1  -0.0647     0.0844   -0.77   0.4439
## LastAuthorFemale1   -0.1613     0.0909   -1.77   0.0772 .
## UniqueAuthors2       0.2865     0.0916    3.13   0.0020 **
## UniqueAuthors3       0.2933     0.1200    2.44   0.0152 *
## UniqueAuthors4       0.3702     0.1303    2.84   0.0048 **
## UniqueAuthors5       0.3753     0.1207    3.11   0.0021 **
## Year1997            -0.0714     0.2396   -0.30   0.7661
## Year1998            -0.1095     0.2360   -0.46   0.6430
## Year1999             0.3149     0.2461    1.28   0.2019
```

```
## Year2000             0.2184     0.2386    0.92   0.3610
## Year2001             0.2718     0.2642    1.03   0.3046
## Year2002            -0.0998     0.2228   -0.45   0.6545
## Year2003            -0.0477     0.2375   -0.20   0.8409
## Year2004             0.2604     0.2893    0.90   0.3689
## Year2005             0.1337     0.2356    0.57   0.5709
## Year2006             0.0390     0.2860    0.14   0.8917
## Year2007             0.0439     0.2408    0.18   0.8554
## Year2008             0.0907     0.2379    0.38   0.7035
## Year2009             0.1512     0.2507    0.60   0.5469
## Year2010            -0.1621     0.2766   -0.59   0.5584
## Year2011             0.0373     0.2496    0.15   0.8813
## Year2012            -0.1120     0.2463   -0.45   0.6497
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.584
## Multiple R-squared:  0.128,  Adjusted R-squared:  0.0564
## Convergence in 15 IRWLS iterations
##
## Robustness weights:
##  26 weights are ~= 1. The remaining 264 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.190   0.869   0.949   0.913   0.984   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          3.45e-04          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.554  1           1.247
## LastAuthorFemale  1.737  1           1.318
## Year              1.607 16           1.015
```



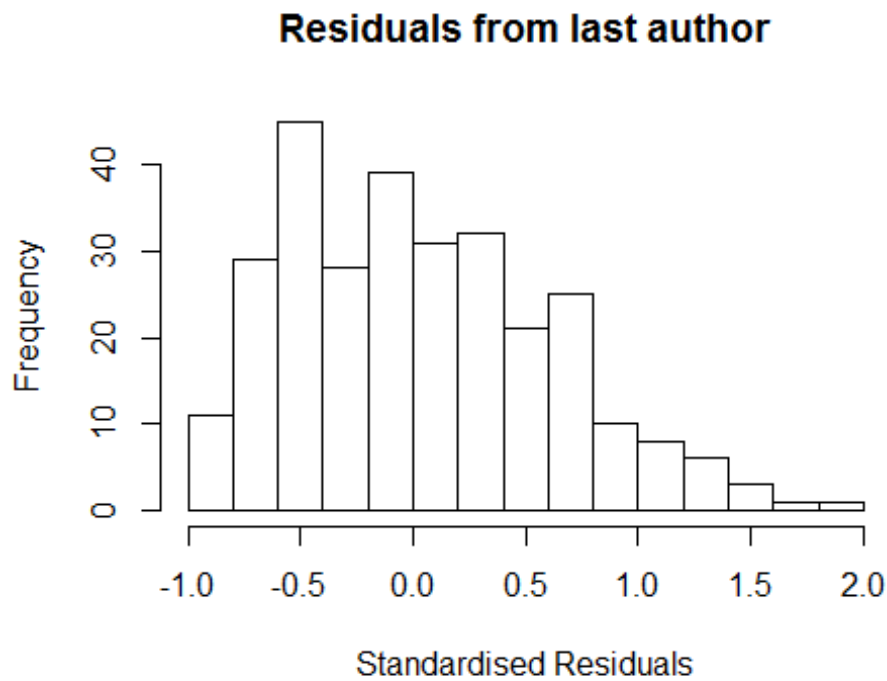

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -0.9801 -0.4646 -0.0162  0.4166  1.8186
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         0.68560    0.21827    3.14   0.0019 **
## FirstAuthorFemale1 -0.04677    0.08607   -0.54   0.5873
## LastAuthorFemale1  -0.11842    0.09228   -1.28   0.2005
## Year1997           -0.10235    0.25219   -0.41   0.6852
## Year1998           -0.14323    0.24911   -0.57   0.5658
## Year1999            0.29446    0.26978    1.09   0.2760
## Year2000            0.19128    0.25105    0.76   0.4468
## Year2001            0.24188    0.27687    0.87   0.3831
## Year2002           -0.06728    0.23809   -0.28   0.7777
## Year2003            0.01673    0.25839    0.06   0.9484
## Year2004            0.36898    0.29134    1.27   0.2064
## Year2005            0.14238    0.25371    0.56   0.5751
```

```
## Year2006            0.04698    0.29744    0.16   0.8746
## Year2007            0.12559    0.25331    0.50   0.6204
## Year2008            0.10945    0.24273    0.45   0.6524
## Year2009            0.07826    0.25070    0.31   0.7551
## Year2010           -0.11622    0.30772   -0.38   0.7060
## Year2011            0.00764    0.25974    0.03   0.9766
## Year2012           -0.03615    0.25942   -0.14   0.8893
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.591
## Multiple R-squared:  0.0669, Adjusted R-squared:  0.0049
## Convergence in 15 IRWLS iterations
##
## Robustness weights:
##  21 weights are ~= 1. The remaining 269 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.324   0.881   0.942   0.911   0.983   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          3.45e-04          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.241  1           1.114
## Year              1.241 16           1.007
```




```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -0.9663 -0.4465 -0.0272  0.4200  1.7409
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)           0.6935     0.2213    3.13   0.0019 **
## FirstAuthorFemale1   -0.0953     0.0777   -1.23   0.2209
## Year1997             -0.1565     0.2473   -0.63   0.5272
## Year1998             -0.1785     0.2503   -0.71   0.4763
## Year1999              0.2728     0.2665    1.02   0.3069
## Year2000              0.1382     0.2526    0.55   0.5848
## Year2001              0.1962     0.2752    0.71   0.4765
## Year2002             -0.0848     0.2425   -0.35   0.7267
## Year2003             -0.0225     0.2598   -0.09   0.9310
## Year2004              0.3749     0.2919    1.28   0.2001
## Year2005              0.1328     0.2583    0.51   0.6077
## Year2006              0.0354     0.2881    0.12   0.9024
```

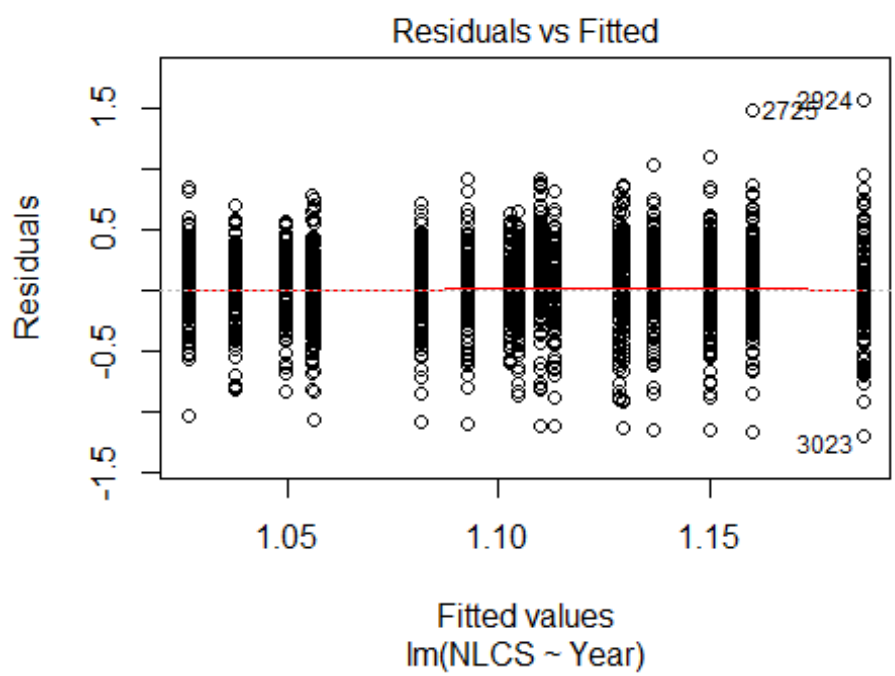
```
## Year2007             0.0964     0.2539    0.38   0.7043
## Year2008             0.0923     0.2469    0.37   0.7088
## Year2009             0.0324     0.2511    0.13   0.8973
## Year2010            -0.1616     0.3023   -0.53   0.5935
## Year2011            -0.0378     0.2561   -0.15   0.8828
## Year2012            -0.0475     0.2631   -0.18   0.8569
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.594
## Multiple R-squared:  0.062,  Adjusted R-squared:  0.00339
## Convergence in 15 IRWLS iterations
##
## Robustness weights:
##  22 weights are ~= 1. The remaining 268 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.371   0.888   0.938   0.911   0.982   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          3.45e-04          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##                   GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.375  1           1.173
## Year             1.375 16           1.010
```



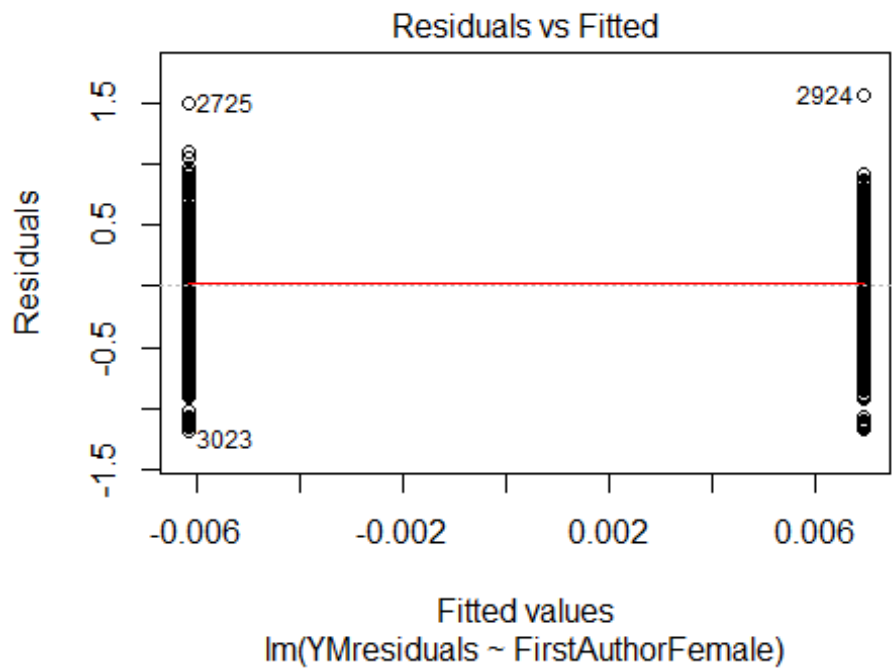

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -0.9724 -0.4391 -0.0242  0.4344  1.8011
##
## Coefficients:
##                   Estimate Std. Error t value Pr(>|t|)
## (Intercept)        0.67644    0.21556    3.14   0.0019 **
## LastAuthorFemale1 -0.13849    0.08268   -1.67   0.0951 .
## Year1997          -0.09885    0.25026   -0.39   0.6932
## Year1998          -0.13775    0.24768   -0.56   0.5785
## Year1999           0.29595    0.26809    1.10   0.2706
## Year2000           0.19436    0.24948    0.78   0.4366
## Year2001           0.24521    0.27559    0.89   0.3744
## Year2002          -0.06571    0.23573   -0.28   0.7807
## Year2003           0.01405    0.25777    0.05   0.9566
## Year2004           0.34609    0.28680    1.21   0.2286
## Year2005           0.14263    0.25348    0.56   0.5741
## Year2006           0.05135    0.30121    0.17   0.8648
```

```
## Year2007           0.12876    0.25231    0.51   0.6102
## Year2008           0.10723    0.24172    0.44   0.6577
## Year2009           0.07705    0.24866    0.31   0.7569
## Year2010          -0.11730    0.30983   -0.38   0.7053
## Year2011           0.00114    0.25877    0.00   0.9965
## Year2012          -0.03443    0.25859   -0.13   0.8942
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.588
## Multiple R-squared:  0.066,  Adjusted R-squared:  0.00763
## Convergence in 15 IRWLS iterations
##
## Robustness weights:
##  21 weights are ~= 1. The remaining 269 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.328   0.883   0.940   0.910   0.983   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          3.45e-04          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Sample size for the above analysis:  290"
## [1] ""
## [1] ""
## [1] "####################################"
## [1] "Analysis of AJSC 2710"
## [1] "####################################"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  169  206  199  208  176  172  172  193  191  165  167  163  212  183  175
## 2011 2012
##  163  172
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##   85  129  112  136  106   91  116  112  118  106   98  104  151  129  126
## 2011 2012
```

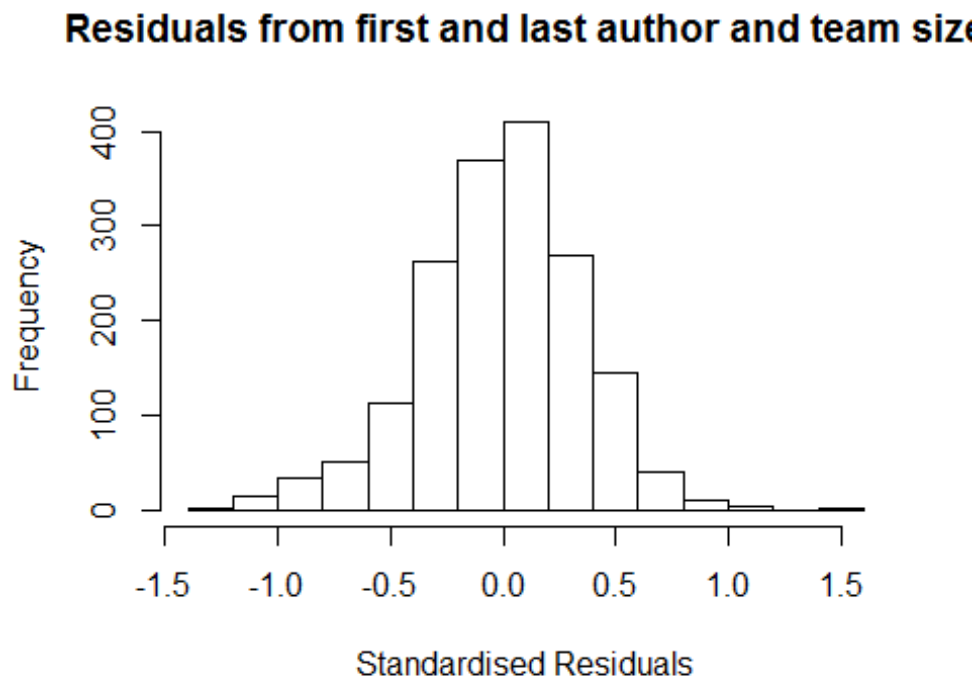
```
##  125  130
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##   73  115   97  117   91   82  101  106  101   93   84   93  133  104  116
## 2011 2012
##  107  114
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
##  Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 100, df = 16, p-value = 1e-14
```




```
##
##  Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 6.4, df = 1, p-value = 0.01
```

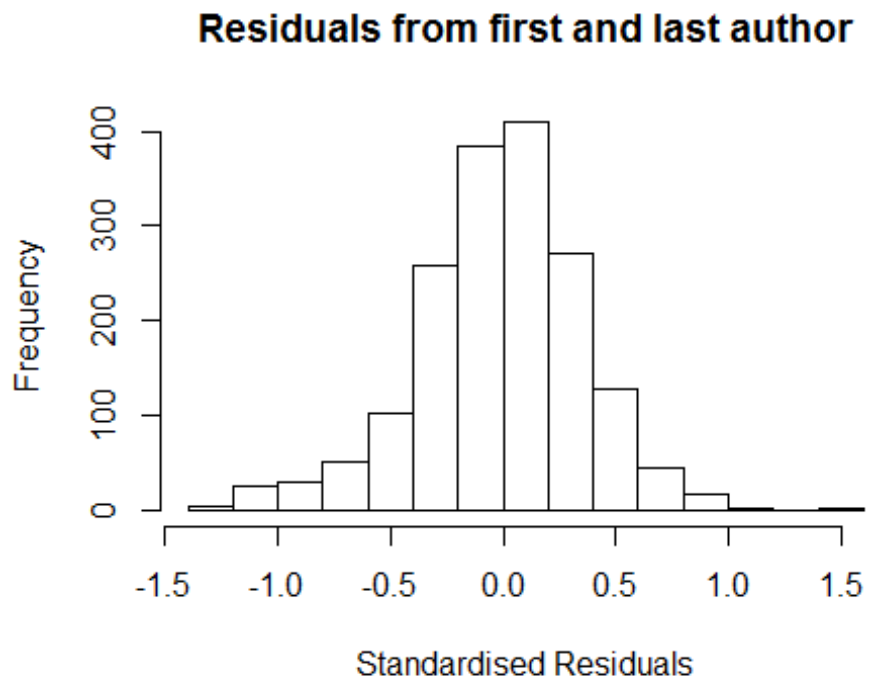



```
## [1] "Female first author team size 2018 geometric mean: 4.77575572552103"
## [1] "Male first author team size 2018 geometric mean: 4.35779145865135"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 1400, p-value = 0.5
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 4.41382855702515"
## [1] "Male last author team size 2018 geometric mean: 4.79890611851148"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 1200, p-value = 0.5
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.083  1           1.041
## LastAuthorFemale  1.051  1           1.025
## UniqueAuthors     1.282  4           1.032
## Year              1.406 16           1.011
```



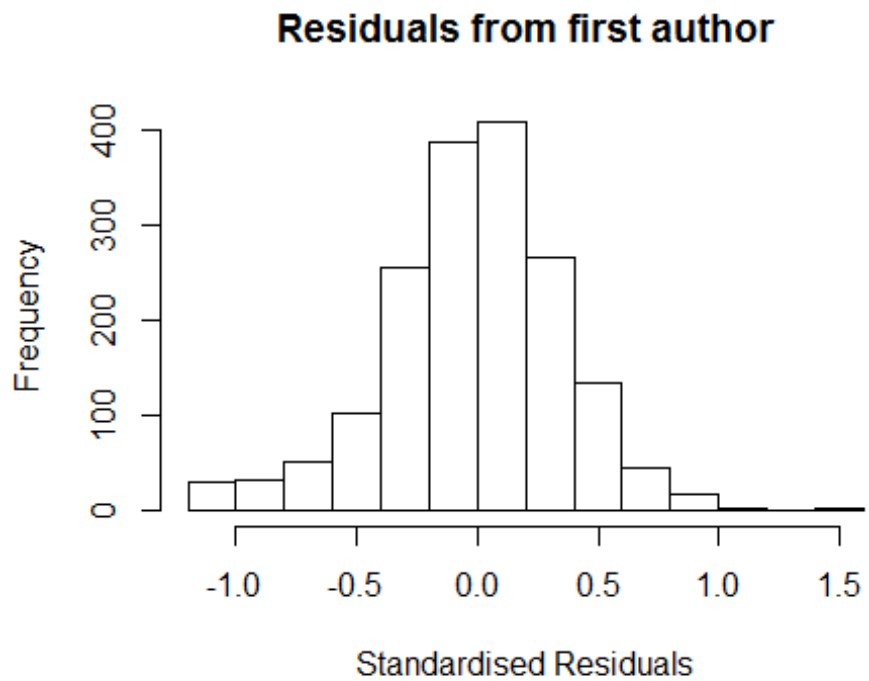

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
UniqueAuthors +
##     Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##     k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.2237 -0.2230  0.0064  0.2183  1.5172
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         0.95824    0.06527   14.68  < 2e-16 ***
## FirstAuthorFemale1  0.01300    0.01726    0.75  0.45155
## LastAuthorFemale1   0.02745    0.01826    1.50  0.13287
## UniqueAuthors2      0.17250    0.05161    3.34  0.00085 ***
## UniqueAuthors3      0.19017    0.05075    3.75  0.00019 ***
## UniqueAuthors4      0.22839    0.05141    4.44  9.4e-06 ***
## UniqueAuthors5      0.25858    0.04975    5.20  2.3e-07 ***
## Year1997           -0.03372    0.05701   -0.59  0.55428
## Year1998           -0.05984    0.05774   -1.04  0.30024
## Year1999           -0.08992    0.05554   -1.62  0.10564
```

```
## Year2000           -0.14505    0.05838   -2.48  0.01306 *
## Year2001           -0.11988    0.05937   -2.02  0.04362 *
## Year2002           -0.14314    0.05540   -2.58  0.00986 **
## Year2003           -0.11468    0.05515   -2.08  0.03774 *
## Year2004           -0.13714    0.05561   -2.47  0.01375 *
## Year2005           -0.07318    0.05629   -1.30  0.19378
## Year2006           -0.04442    0.06192   -0.72  0.47322
## Year2007           -0.09005    0.06364   -1.42  0.15725
## Year2008           -0.00344    0.05917   -0.06  0.95367
## Year2009           -0.00496    0.06250   -0.08  0.93678
## Year2010            0.01101    0.06058    0.18  0.85577
## Year2011           -0.01141    0.05943   -0.19  0.84773
## Year2012           -0.03351    0.06640   -0.50  0.61382
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.334
## Multiple R-squared:  0.0577, Adjusted R-squared:  0.0455
## Convergence in 15 IRWLS iterations
##
## Robustness weights:
##  153 weights are ~= 1. The remaining 1574 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0035  0.8660  0.9510  0.8950  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          5.79e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.081  1           1.040
## LastAuthorFemale  1.046  1           1.023
## Year              1.120 16           1.004
```



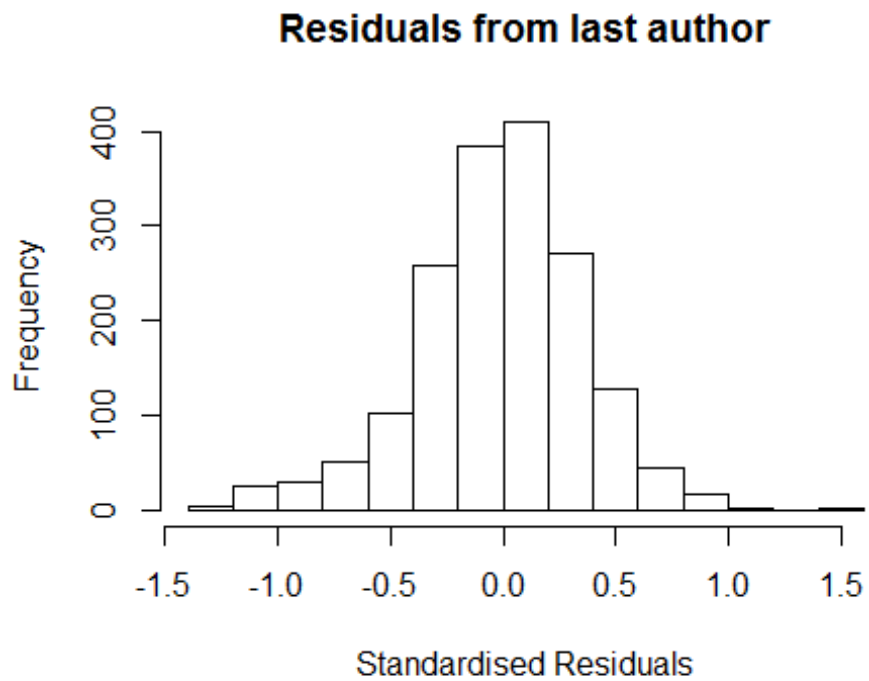

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.21096 -0.22080  0.00194  0.22562  1.56394
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.15212    0.04886   23.58   <2e-16 ***
## FirstAuthorFemale1  0.02213    0.01742    1.27    0.204
## LastAuthorFemale1   0.02186    0.01842    1.19    0.236
## Year1997           -0.02807    0.05809   -0.48    0.629
## Year1998           -0.06261    0.05928   -1.06    0.291
## Year1999           -0.08623    0.05664   -1.52    0.128
## Year2000           -0.13346    0.06019   -2.22    0.027 *
## Year2001           -0.10524    0.06081   -1.73    0.084 .
## Year2002           -0.14019    0.05681   -2.47    0.014 *
## Year2003           -0.10691    0.05664   -1.89    0.059 .
## Year2004           -0.11926    0.05643   -2.11    0.035 *
## Year2005           -0.06133    0.05720   -1.07    0.284
```

```
## Year2006           -0.03130    0.06252   -0.50    0.617
## Year2007           -0.07795    0.06430   -1.21    0.226
## Year2008            0.00702    0.06140    0.11    0.909
## Year2009            0.01485    0.06506    0.23    0.819
## Year2010            0.01981    0.06061    0.33    0.744
## Year2011            0.00309    0.06129    0.05    0.960
## Year2012           -0.02463    0.06642   -0.37    0.711
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.337
## Multiple R-squared:  0.0251, Adjusted R-squared:  0.0149
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
##  149 weights are ~= 1. The remaining 1578 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0004  0.8720  0.9510  0.8930  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          5.79e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.079  1           1.039
## Year              1.079 16           1.002
```




```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.19817 -0.22206  0.00269  0.22406  1.55643
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.15415    0.04884   23.63   <2e-16 ***
## FirstAuthorFemale1  0.02472    0.01745    1.42    0.157
## Year1997           -0.02516    0.05793   -0.43    0.664
## Year1998           -0.06223    0.05929   -1.05    0.294
## Year1999           -0.08314    0.05647   -1.47    0.141
## Year2000           -0.13040    0.06003   -2.17    0.030 *
## Year2001           -0.10194    0.06059   -1.68    0.093 .
## Year2002           -0.13601    0.05643   -2.41    0.016 *
## Year2003           -0.10327    0.05640   -1.83    0.067 .
## Year2004           -0.11487    0.05624   -2.04    0.041 *
## Year2005           -0.05764    0.05701   -1.01    0.312
## Year2006           -0.02802    0.06247   -0.45    0.654
```

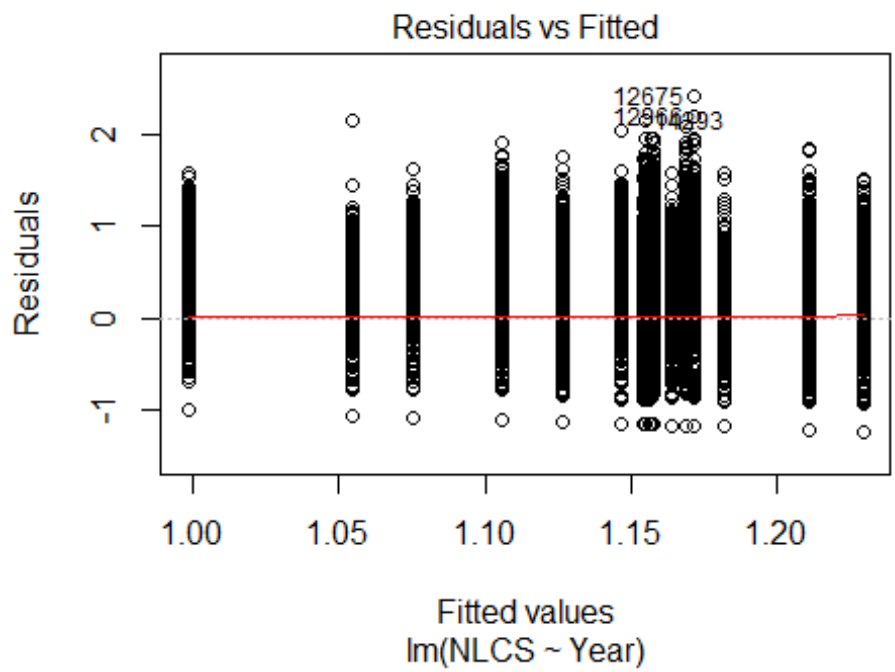
```
## Year2007           -0.07347    0.06408   -1.15    0.252
## Year2008            0.00960    0.06122    0.16    0.875
## Year2009            0.01930    0.06468    0.30    0.765
## Year2010            0.02270    0.06035    0.38    0.707
## Year2011            0.00788    0.06111    0.13    0.897
## Year2012           -0.01856    0.06608   -0.28    0.779
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.337
## Multiple R-squared:  0.0244, Adjusted R-squared:  0.0147
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
##  151 weights are ~= 1. The remaining 1576 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0008  0.8710  0.9500  0.8930  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          5.79e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##                   GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.046  1           1.023
## Year             1.046 16           1.001
```



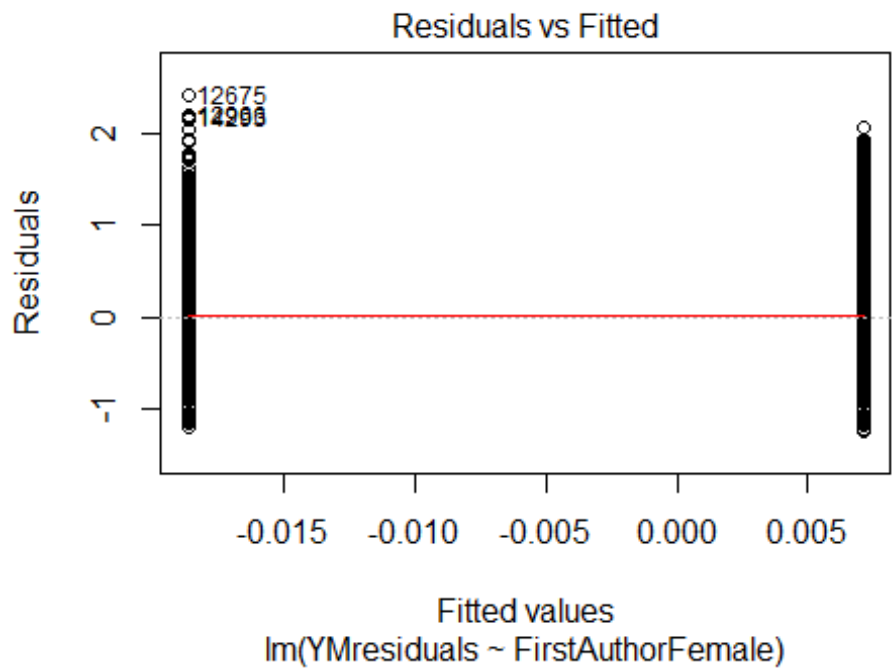

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##    Min     1Q Median     3Q    Max
## -1.207 -0.223  0.001  0.225  1.576
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.16005    0.04847   23.93   <2e-16 ***
## LastAuthorFemale1   0.02495    0.01845    1.35    0.176
## Year1997           -0.02849    0.05816   -0.49    0.624
## Year1998           -0.06023    0.05922   -1.02    0.309
## Year1999           -0.08437    0.05656   -1.49    0.136
## Year2000           -0.13442    0.06016   -2.23    0.026 *
## Year2001           -0.10386    0.06081   -1.71    0.088 .
## Year2002           -0.14064    0.05679   -2.48    0.013 *
## Year2003           -0.10369    0.05649   -1.84    0.067 .
## Year2004           -0.11605    0.05617   -2.07    0.039 *
## Year2005           -0.05821    0.05689   -1.02    0.306
## Year2006           -0.02902    0.06243   -0.46    0.642
```

```
## Year2007          -0.07457    0.06405   -1.16    0.244
## Year2008           0.01023    0.06114    0.17    0.867
## Year2009           0.01735    0.06493    0.27    0.789
## Year2010           0.02214    0.06035    0.37    0.714
## Year2011           0.00572    0.06107    0.09    0.925
## Year2012          -0.02195    0.06631   -0.33    0.741
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.338
## Multiple R-squared:  0.0243, Adjusted R-squared:  0.0146
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
##  observation 1400 is an outlier with |weight| <= 5.5e-05 ( < 5.8e-05);
##  145 weights are ~= 1. The remaining 1581 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0156  0.8730  0.9510  0.8940  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          5.79e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Sample size for the above analysis:  1727"
## [1] ""
## [1] ""
## [1] "####################################"
## [1] "Analysis of AJSC 2711"
## [1] "####################################"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  843  923  897 1003 1033 1022  926  801  863  802 1021 1114 1221 1291 1398
## 2011 2012
## 1488 1529
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  460  549  519  610  775  673  765  671  733  705  881  967 1060 1141 1218
```

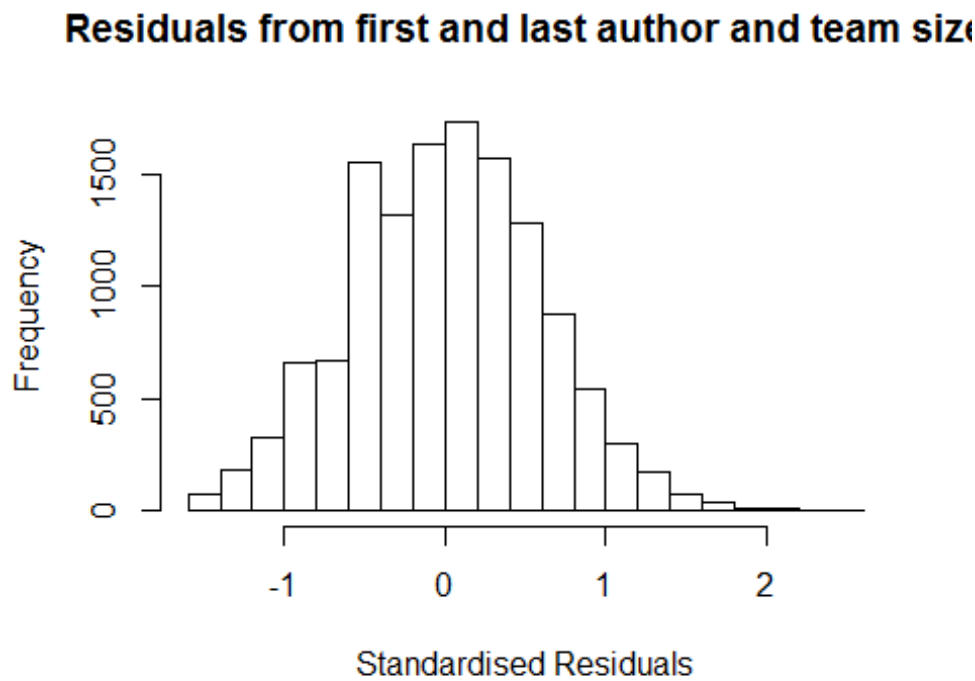
```
## 2011 2012
## 1299 1303
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  416  504  466  567  709  607  700  608  681  639  793  876  971 1040 1104
## 2011 2012
## 1178 1162
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
##  Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 290, df = 16, p-value <2e-16
```




```
##
##  Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 8.5, df = 1, p-value = 0.003
```

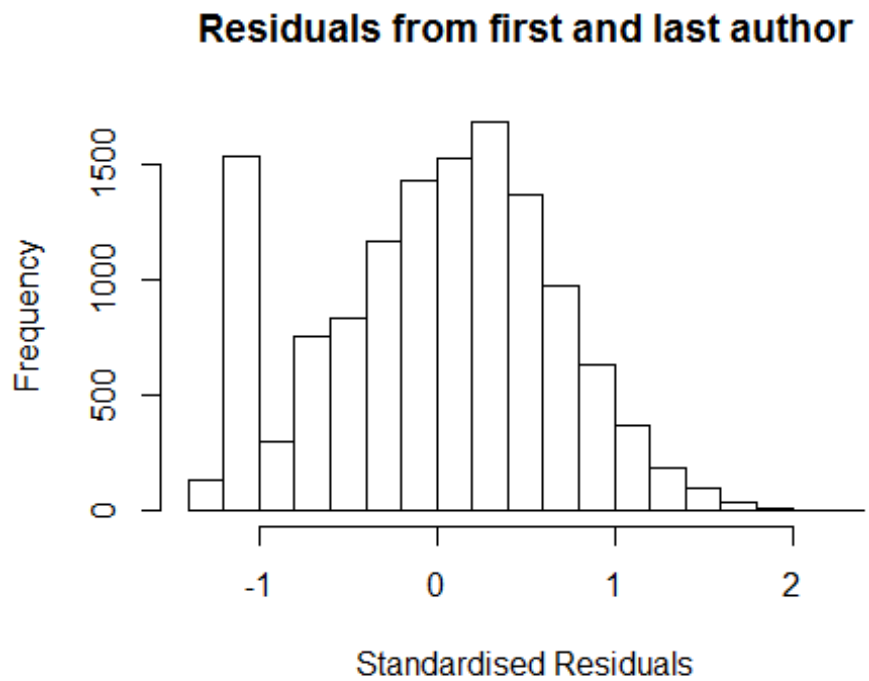



```
## [1] "Female first author team size 2018 geometric mean: 3.85766322410294"
## [1] "Male first author team size 2018 geometric mean: 4.0599174574845"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 210000, p-value = 0.2
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 3.59636119153992"
## [1] "Male last author team size 2018 geometric mean: 4.17927765936503"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 180000, p-value = 3e-05
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##                     GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.058  1           1.028
## LastAuthorFemale  1.051  1           1.025
## UniqueAuthors     1.077  4           1.009
## Year              1.102 16           1.003
```



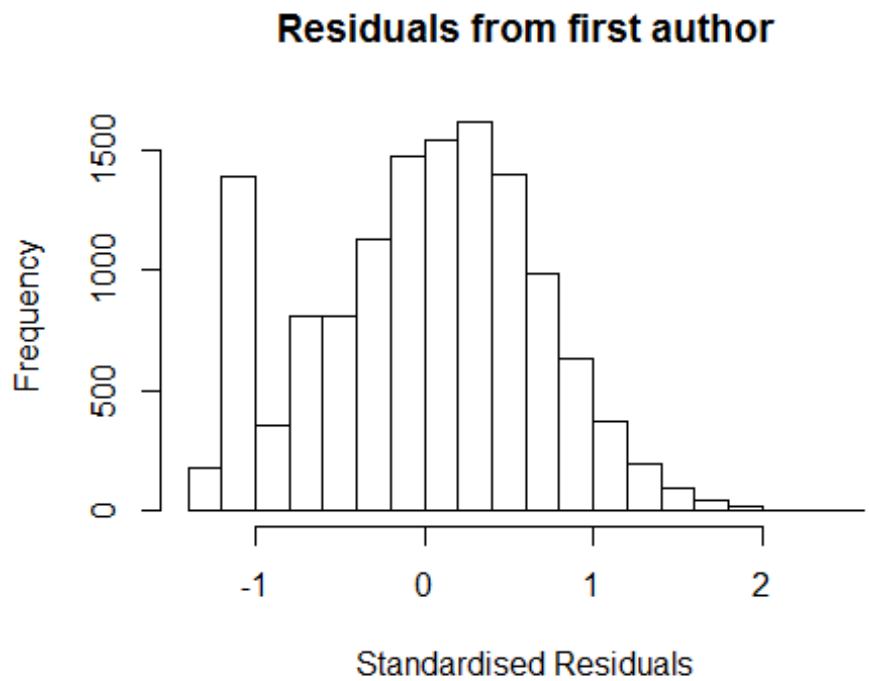

```
## [1] "List of  1 outliers with residuals above  2.5"
##          ScopusId  NLCS Year OneField Fields residuals
## 12787 48549092221 3.093 2008     2711      2     2.577
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
UniqueAuthors +
##     Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##     k.max = 1000))
##  \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.56141 -0.43755  0.00828  0.40422  2.57708
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         0.54359    0.03030   17.94  < 2e-16 ***
## FirstAuthorFemale1 -0.03023    0.01225   -2.47   0.0136 *
## LastAuthorFemale1  -0.02706    0.01324   -2.04   0.0410 *
## UniqueAuthors2      0.42588    0.02041   20.86  < 2e-16 ***
## UniqueAuthors3      0.64527    0.02016   32.01  < 2e-16 ***
## UniqueAuthors4      0.77831    0.02034   38.27  < 2e-16 ***
## UniqueAuthors5      0.95713    0.01797   53.27  < 2e-16 ***
## Year1997            0.02231    0.03554    0.63   0.5302
## Year1998            0.03206    0.03525    0.91   0.3631
## Year1999            0.05875    0.03525    1.67   0.0956 .
```

```
## Year2000            0.04281    0.03316    1.29   0.1967
## Year2001            0.00163    0.03504    0.05   0.9628
## Year2002            0.00467    0.03450    0.14   0.8922
## Year2003           -0.10605    0.03390   -3.13   0.0018 **
## Year2004           -0.14657    0.03647   -4.02  5.9e-05 ***
## Year2005           -0.04811    0.03618   -1.33   0.1837
## Year2006            0.06069    0.03488    1.74   0.0819 .
## Year2007            0.04672    0.03347    1.40   0.1627
## Year2008            0.02962    0.03384    0.88   0.3815
## Year2009           -0.00885    0.03354   -0.26   0.7919
## Year2010           -0.07205    0.03321   -2.17   0.0301 *
## Year2011           -0.03987    0.03267   -1.22   0.2225
## Year2012           -0.07336    0.03236   -2.27   0.0234 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.569
## Multiple R-squared:  0.233,  Adjusted R-squared:  0.232
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
##  1043 weights are ~= 1. The remaining 11978 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0045  0.8670  0.9410  0.9010  0.9840  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          7.68e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.036  1           1.018
## LastAuthorFemale  1.026  1           1.013
## Year              1.038 16           1.001
```



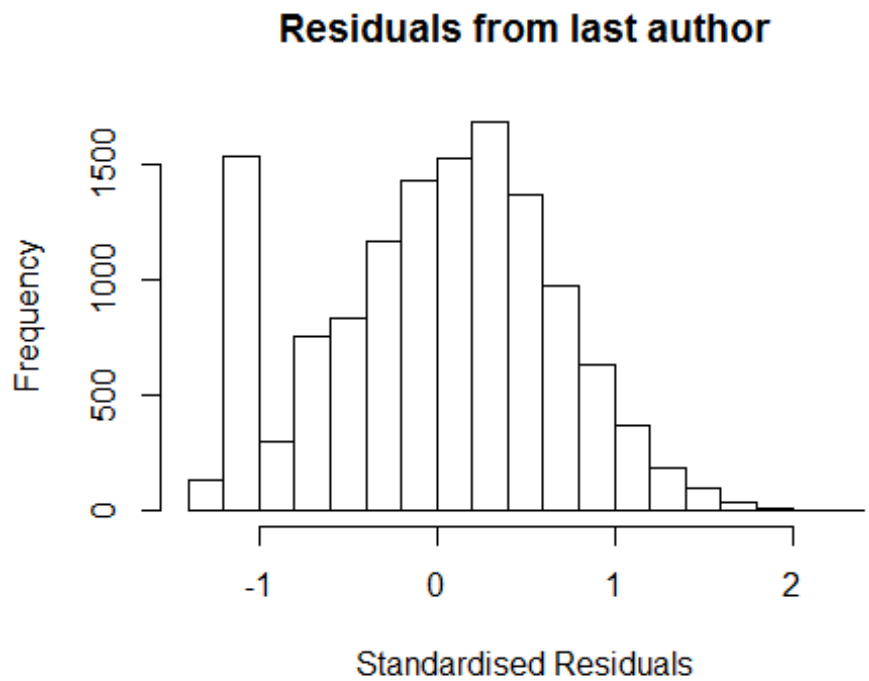

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.2693 -0.4703  0.0491  0.4593  2.3989
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)          1.15602    0.02535   45.61  < 2e-16 ***
## FirstAuthorFemale1 -0.01162    0.01363   -0.85  0.39379
## LastAuthorFemale1  -0.07217    0.01478   -4.88  1.1e-06 ***
## Year1997            0.02735    0.03487    0.78  0.43295
## Year1998            0.04370    0.03397    1.29  0.19828
## Year1999            0.03729    0.03677    1.01  0.31054
## Year2000            0.02287    0.03429    0.67  0.50472
## Year2001           -0.05684    0.03803   -1.49  0.13503
## Year2002           -0.01465    0.03569   -0.41  0.68145
## Year2003           -0.08669    0.03539   -2.45  0.01433 *
## Year2004           -0.14584    0.03862   -3.78  0.00016 ***
## Year2005           -0.00975    0.03838   -0.25  0.79947
```

```
## Year2006            0.11328    0.03547    3.19  0.00141 **
## Year2007            0.08614    0.03455    2.49  0.01268 *
## Year2008            0.03166    0.03550    0.89  0.37256
## Year2009           -0.00412    0.03577   -0.12  0.90818
## Year2010           -0.04329    0.03546   -1.22  0.22216
## Year2011            0.02050    0.03390    0.60  0.54543
## Year2012            0.01491    0.03294    0.45  0.65088
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.666
## Multiple R-squared:  0.00923,    Adjusted R-squared:  0.00786
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
##  1074 weights are ~= 1. The remaining 11947 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.167   0.856   0.948   0.907   0.985   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          7.68e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.026  1           1.013
## Year              1.026 16           1.001
```




```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.2544 -0.4694  0.0469  0.4605  2.4287
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.14616    0.02514   45.60  < 2e-16 ***
## FirstAuthorFemale1 -0.02545    0.01373   -1.85  0.06376 .
## Year1997            0.02700    0.03485    0.77  0.43858
## Year1998            0.04400    0.03383    1.30  0.19344
## Year1999            0.03657    0.03676    0.99  0.31978
## Year2000            0.02166    0.03427    0.63  0.52739
## Year2001           -0.05774    0.03805   -1.52  0.12923
## Year2002           -0.01578    0.03564   -0.44  0.65792
## Year2003           -0.08857    0.03533   -2.51  0.01218 *
## Year2004           -0.14833    0.03849   -3.85  0.00012 ***
## Year2005           -0.01294    0.03827   -0.34  0.73525
## Year2006            0.10828    0.03544    3.06  0.00226 **
```

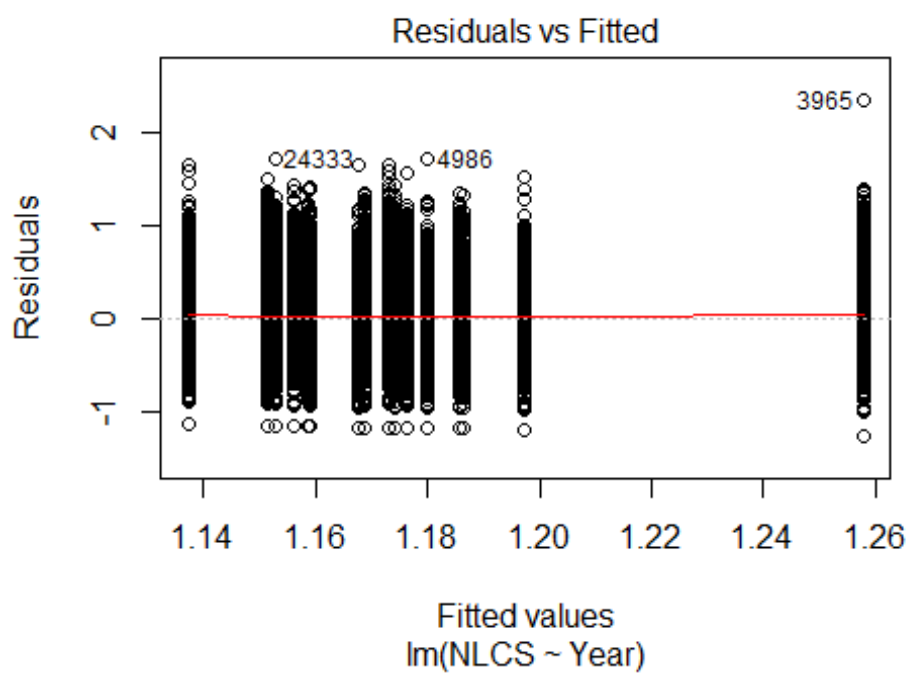
```
## Year2007            0.08430    0.03453    2.44  0.01464 *
## Year2008            0.02559    0.03547    0.72  0.47054
## Year2009           -0.00702    0.03574   -0.20  0.84433
## Year2010           -0.04686    0.03543   -1.32  0.18599
## Year2011            0.01758    0.03384    0.52  0.60341
## Year2012            0.00958    0.03287    0.29  0.77059
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.667
## Multiple R-squared:  0.00736,    Adjusted R-squared:  0.00606
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  1087 weights are ~= 1. The remaining 11934 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.157   0.857   0.948   0.907   0.984   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          7.68e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##                   GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.017  1           1.008
## Year             1.017 16           1.001
```



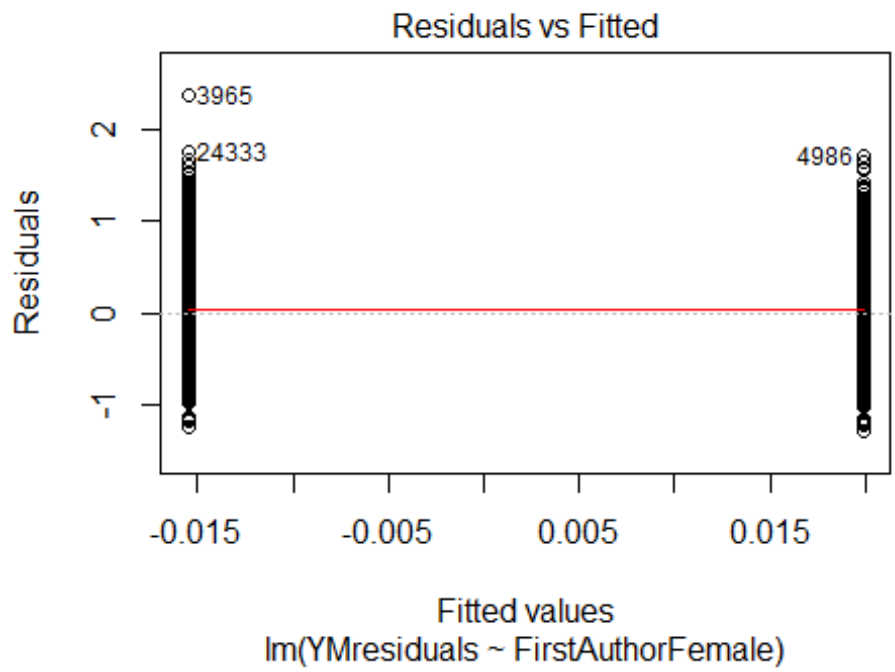

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.2665 -0.4719  0.0504  0.4609  2.3904
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.15391    0.02522   45.75  < 2e-16 ***
## LastAuthorFemale1 -0.07471    0.01480   -5.05  4.5e-07 ***
## Year1997            0.02731    0.03489    0.78  0.43381
## Year1998            0.04358    0.03398    1.28  0.19967
## Year1999            0.03737    0.03680    1.02  0.30992
## Year2000            0.02294    0.03430    0.67  0.50366
## Year2001           -0.05688    0.03805   -1.50  0.13494
## Year2002           -0.01511    0.03569   -0.42  0.67200
## Year2003           -0.08702    0.03540   -2.46  0.01397 *
## Year2004           -0.14581    0.03861   -3.78  0.00016 ***
## Year2005           -0.00972    0.03839   -0.25  0.80015
## Year2006            0.11257    0.03544    3.18  0.00150 **
```

```
## Year2007           0.08543    0.03455    2.47  0.01342 *
## Year2008           0.03066    0.03546    0.86  0.38732
## Year2009          -0.00530    0.03570   -0.15  0.88208
## Year2010          -0.04407    0.03542   -1.24  0.21352
## Year2011           0.01906    0.03383    0.56  0.57318
## Year2012           0.01364    0.03291    0.41  0.67864
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.667
## Multiple R-squared:  0.00916,    Adjusted R-squared:  0.00787
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  1070 weights are ~= 1. The remaining 11951 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.172   0.857   0.948   0.907   0.985   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          7.68e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Sample size for the above analysis:  13021"
## [1] ""
## [1] ""
## [1] "####################################"
## [1] "Analysis of AJSC 2712"
## [1] "####################################"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 1606 2060 1433 1542 1638 1788 1803 1661 1848 2016 2081 2327 2424 2293 2562
## 2011 2012
## 2748 2713
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  941  980  861  925  819  873 1131 1100 1263 1381 1418 1620 1706 1603 1795
## 2011 2012
```

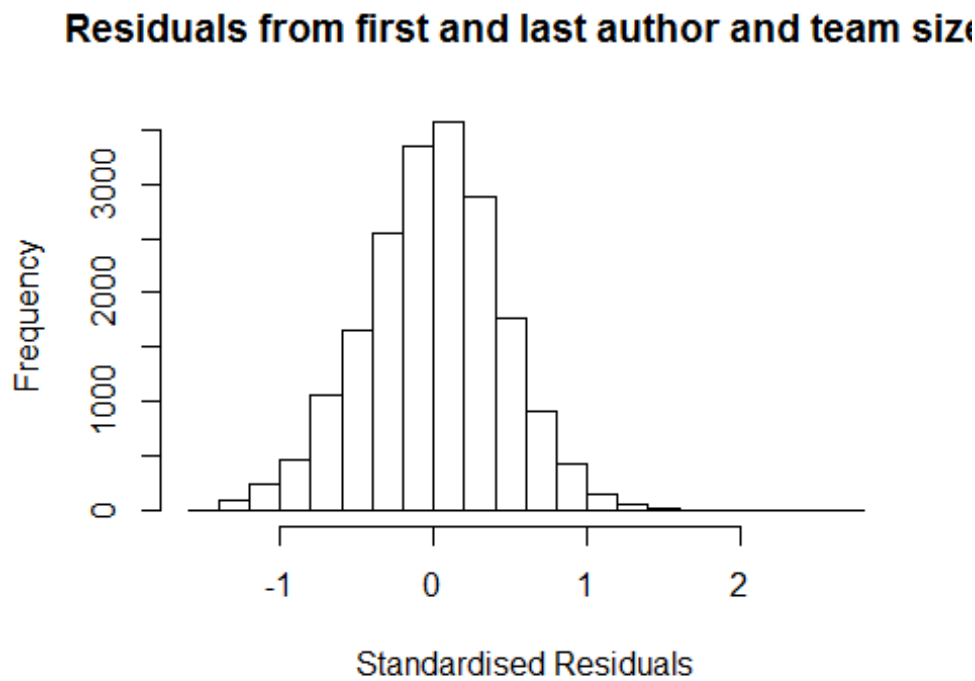
```
## 1905 1873
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  824  875  756  795  703  758  970  943 1092 1180 1202 1398 1480 1385 1541
## 2011 2012
## 1654 1619
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
##  Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 160, df = 16, p-value <2e-16
```




```
##
##  Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 15, df = 1, p-value = 8e-05
```

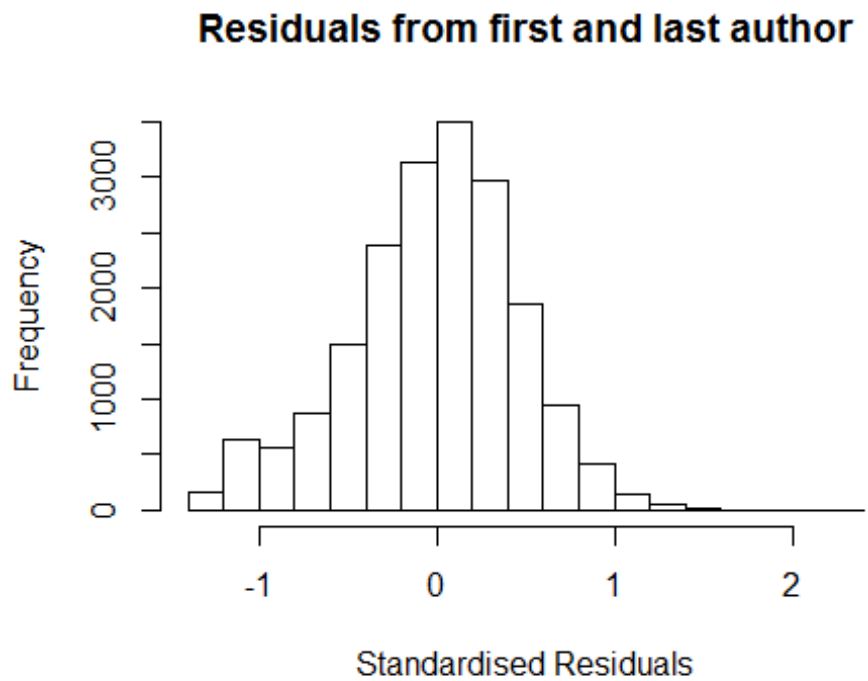



```
## [1] "Female first author team size 2018 geometric mean: 5.67721628929419"
## [1] "Male first author team size 2018 geometric mean: 5.33611223658003"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 240000, p-value = 0.4
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 5.43657474604497"
## [1] "Male last author team size 2018 geometric mean: 5.58728406223748"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 220000, p-value = 0.2
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##                     GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.044  1           1.022
## LastAuthorFemale  1.035  1           1.017
## UniqueAuthors     1.047  4           1.006
## Year              1.067 16           1.002
```



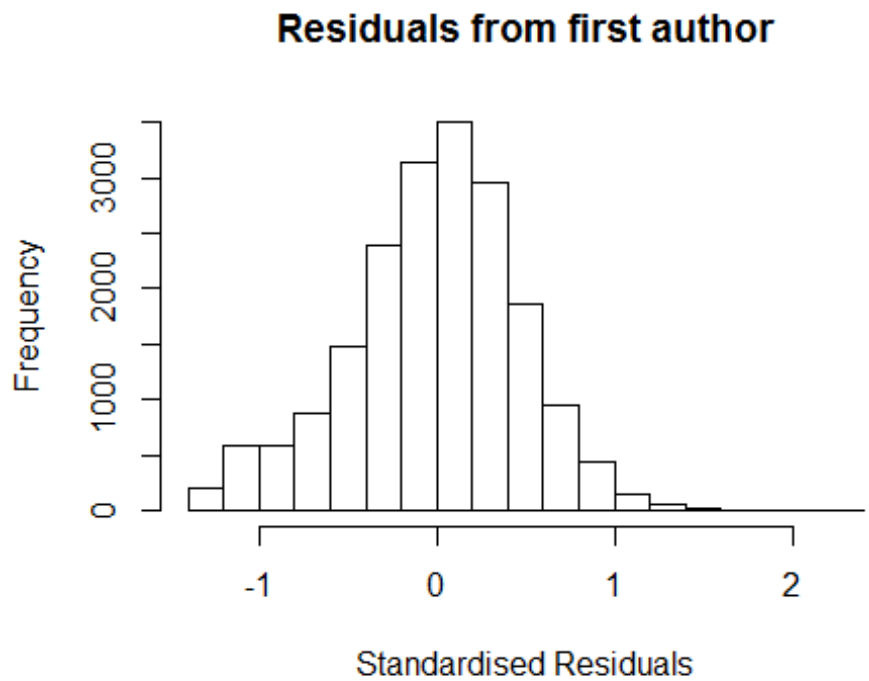

```
## [1] "List of  1 outliers with residuals above  2.5"
##        ScopusId  NLCS Year OneField Fields residuals
## 3965 1842334456 3.603 1997     2712      3     2.706
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
UniqueAuthors +
##     Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##     k.max = 1000))
##  \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.49513 -0.29155  0.00855  0.29000  2.70566
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         0.79188    0.02405   32.93  < 2e-16 ***
## FirstAuthorFemale1  0.01482    0.00669    2.21  0.02682 *
## LastAuthorFemale1  -0.00612    0.00741   -0.83  0.40852
## UniqueAuthors2      0.24680    0.01923   12.84  < 2e-16 ***
## UniqueAuthors3      0.36731    0.01791   20.51  < 2e-16 ***
## UniqueAuthors4      0.45351    0.01755   25.84  < 2e-16 ***
## UniqueAuthors5      0.58297    0.01606   36.30  < 2e-16 ***
## Year1997            0.10546    0.02654    3.97  7.1e-05 ***
## Year1998           -0.03330    0.02529   -1.32  0.18802
## Year1999           -0.02098    0.02519   -0.83  0.40497
```

```
## Year2000           -0.02911    0.02472   -1.18  0.23894
## Year2001           -0.02896    0.02429   -1.19  0.23333
## Year2002           -0.05321    0.02405   -2.21  0.02695 *
## Year2003           -0.07152    0.02353   -3.04  0.00238 **
## Year2004           -0.07681    0.02314   -3.32  0.00090 ***
## Year2005           -0.07968    0.02290   -3.48  0.00050 ***
## Year2006           -0.08291    0.02267   -3.66  0.00026 ***
## Year2007           -0.08364    0.02267   -3.69  0.00023 ***
## Year2008           -0.10386    0.02261   -4.59  4.4e-06 ***
## Year2009           -0.07201    0.02320   -3.10  0.00191 **
## Year2010           -0.10111    0.02278   -4.44  9.1e-06 ***
## Year2011           -0.08647    0.02232   -3.87  0.00011 ***
## Year2012           -0.10804    0.02242   -4.82  1.5e-06 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.429
## Multiple R-squared:  0.148,  Adjusted R-squared:  0.148
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
##  2 observations c(1558,17201) are outliers with |weight| = 0 ( < 5.2e-06);
##  1685 weights are ~= 1. The remaining 17488 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0694  0.8670  0.9500  0.9000  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          5.22e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.028  1           1.014
## LastAuthorFemale  1.028  1           1.014
## Year              1.032 16           1.001
```



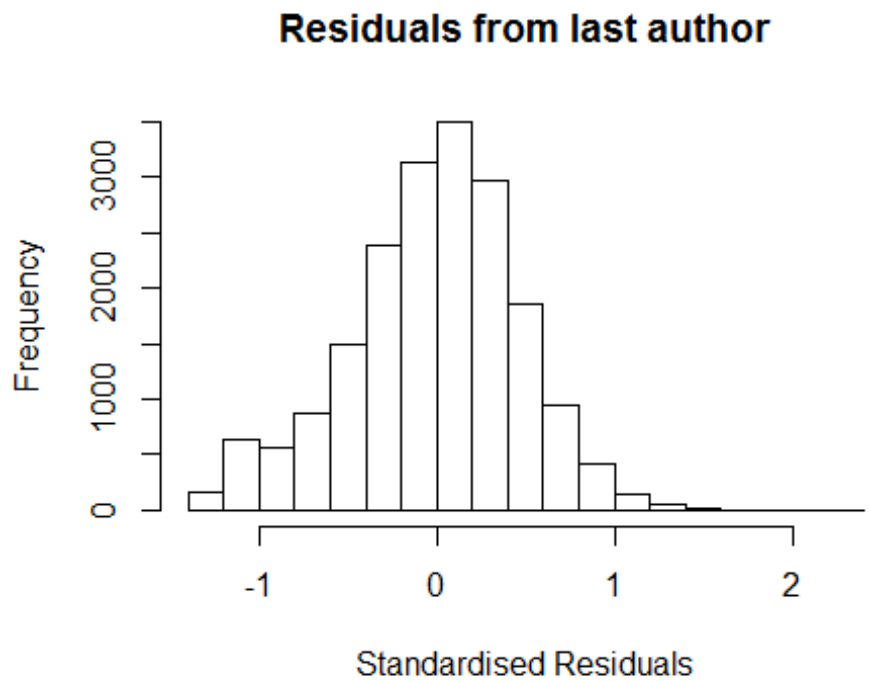

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.3327 -0.3032  0.0185  0.2996  2.3152
##
## Coefficients:
##                      Estimate Std. Error t value Pr(>|t|)
## (Intercept)          1.181026   0.020104   58.75  < 2e-16 ***
## FirstAuthorFemale1   0.044913   0.007042    6.38  1.8e-10 ***
## LastAuthorFemale1   -0.024439   0.007870   -3.11  0.00190 **
## Year1997             0.106765   0.027653    3.86  0.00011 ***
## Year1998             0.000313   0.026473    0.01  0.99056
## Year1999             0.007958   0.026029    0.31  0.75980
## Year2000             0.015216   0.025751    0.59  0.55460
## Year2001             0.010406   0.025626    0.41  0.68470
## Year2002            -0.005511   0.025427   -0.22  0.82842
## Year2003            -0.006740   0.024723   -0.27  0.78515
## Year2004            -0.009832   0.024362   -0.40  0.68653
## Year2005            -0.014962   0.023918   -0.63  0.53162
```

```
## Year2006           -0.024576   0.023789   -1.03  0.30159
## Year2007           -0.019324   0.023711   -0.81  0.41509
## Year2008           -0.046928   0.023789   -1.97  0.04854 *
## Year2009            0.003024   0.024163    0.13  0.90039
## Year2010           -0.017144   0.023777   -0.72  0.47090
## Year2011           -0.013627   0.023165   -0.59  0.55637
## Year2012           -0.035270   0.023371   -1.51  0.13128
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.445
## Multiple R-squared:  0.00604,    Adjusted R-squared:  0.00511
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
##  observation 1558 is an outlier with |weight| = 0 ( < 5.2e-06);
##  1607 weights are ~= 1. The remaining 17567 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0935  0.8630  0.9500  0.8940  0.9860  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          5.22e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.018  1           1.009
## Year              1.018 16           1.001
```




```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##    Min     1Q Median     3Q    Max
## -1.324 -0.304  0.019  0.299  2.320
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.18e+00   2.01e-02   58.59  < 2e-16 ***
## FirstAuthorFemale1  4.11e-02   7.03e-03    5.85  5.1e-09 ***
## Year1997            1.06e-01   2.77e-02    3.82  0.00013 ***
## Year1998           -3.49e-05   2.65e-02    0.00  0.99895
## Year1999            6.62e-03   2.60e-02    0.25  0.79930
## Year2000            1.48e-02   2.58e-02    0.57  0.56564
## Year2001            9.44e-03   2.56e-02    0.37  0.71270
## Year2002           -6.43e-03   2.54e-02   -0.25  0.80042
## Year2003           -7.41e-03   2.47e-02   -0.30  0.76430
## Year2004           -1.11e-02   2.43e-02   -0.46  0.64838
## Year2005           -1.55e-02   2.39e-02   -0.65  0.51760
## Year2006           -2.61e-02   2.38e-02   -1.10  0.27223
```

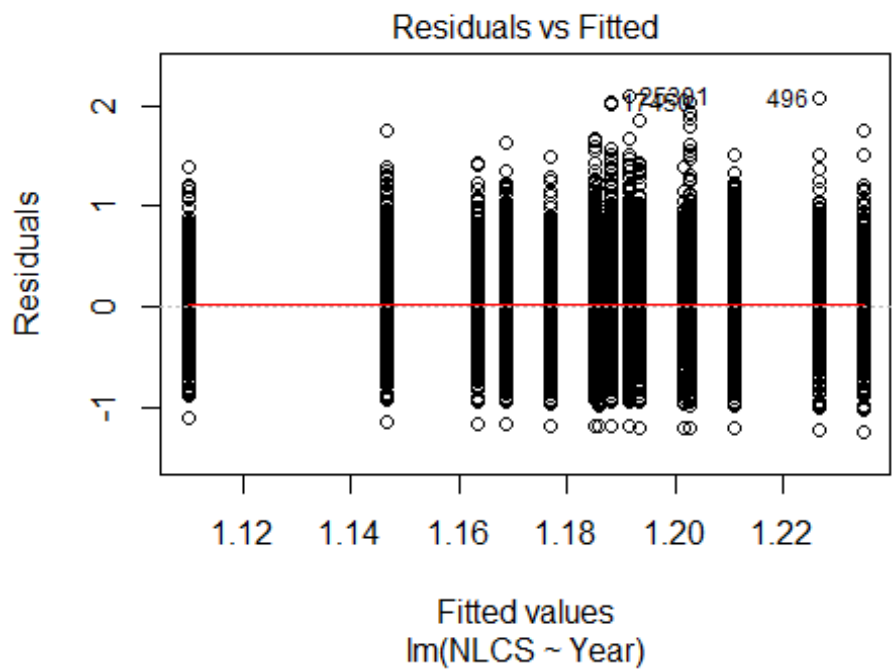
```
## Year2007           -2.13e-02   2.37e-02   -0.90  0.36784
## Year2008           -4.86e-02   2.38e-02   -2.05  0.04078 *
## Year2009            1.58e-03   2.42e-02    0.07  0.94774
## Year2010           -1.98e-02   2.37e-02   -0.84  0.40348
## Year2011           -1.62e-02   2.31e-02   -0.70  0.48374
## Year2012           -3.82e-02   2.33e-02   -1.64  0.10194
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.445
## Multiple R-squared:  0.00557,    Adjusted R-squared:  0.00469
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
##  observation 1558 is an outlier with |weight| = 0 ( < 5.2e-06);
##  1608 weights are ~= 1. The remaining 17566 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.103   0.863   0.950   0.894   0.986   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          5.22e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##                   GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.018  1           1.009
## Year             1.018 16           1.001
```



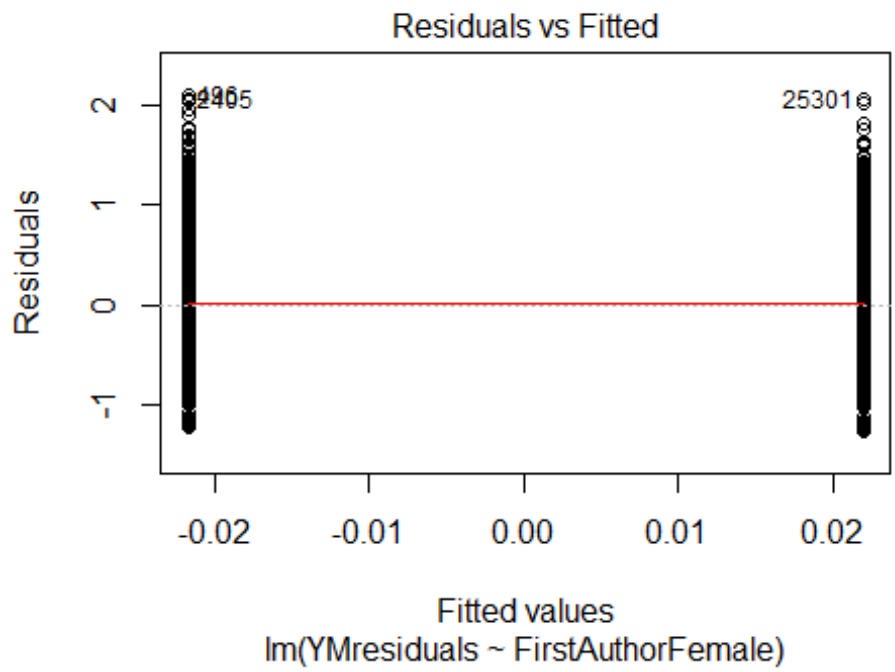

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.3022 -0.3036  0.0183  0.3005  2.3008
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.19615    0.01995   59.97  < 2e-16 ***
## LastAuthorFemale1 -0.01587    0.00785   -2.02  0.04312 *
## Year1997            0.10606    0.02762    3.84  0.00012 ***
## Year1998           -0.00121    0.02646   -0.05  0.96364
## Year1999            0.00865    0.02598    0.33  0.73908
## Year2000            0.01452    0.02573    0.56  0.57253
## Year2001            0.01033    0.02564    0.40  0.68717
## Year2002           -0.00436    0.02541   -0.17  0.86370
## Year2003           -0.00480    0.02470   -0.19  0.84598
## Year2004           -0.00768    0.02435   -0.32  0.75246
## Year2005           -0.01199    0.02390   -0.50  0.61573
## Year2006           -0.02154    0.02379   -0.91  0.36517
```

```
## Year2007          -0.01640    0.02367   -0.69  0.48835
## Year2008          -0.04339    0.02379   -1.82  0.06820 .
## Year2009           0.00669    0.02413    0.28  0.78156
## Year2010          -0.01300    0.02372   -0.55  0.58368
## Year2011          -0.00791    0.02311   -0.34  0.73213
## Year2012          -0.02967    0.02332   -1.27  0.20334
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.446
## Multiple R-squared:  0.00391,    Adjusted R-squared:  0.00302
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
##  observation 1558 is an outlier with |weight| = 0 ( < 5.2e-06);
##  1621 weights are ~= 1. The remaining 17553 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.101   0.863   0.950   0.894   0.985   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          5.22e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                    psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Sample size for the above analysis:  19175"
## [1] ""
## [1] ""
## [1] "####################################"
## [1] "Analysis of AJSC 2713"
## [1] "####################################"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  914 1065  893  963 1116 1179 1212 1115 1142 1313 1492 1627 1738 1788 1910
## 2011 2012
## 2146 2161
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  671  709  600  755  724  750  944  868  892 1009 1164 1270 1321 1383 1464
```

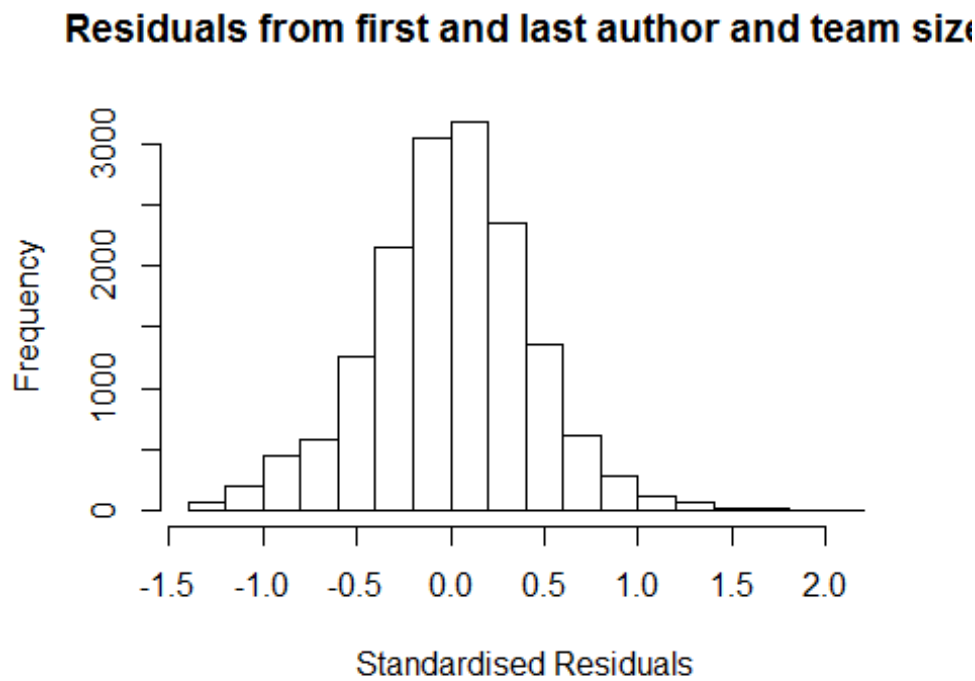
```
## 2011 2012
## 1671 1661
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  601  649  541  667  662  659  844  754  783  885 1020 1118 1150 1222 1287
## 2011 2012
## 1443 1468
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
##  Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 140, df = 16, p-value <2e-16
```




```
##
##  Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 130, df = 1, p-value <2e-16
```

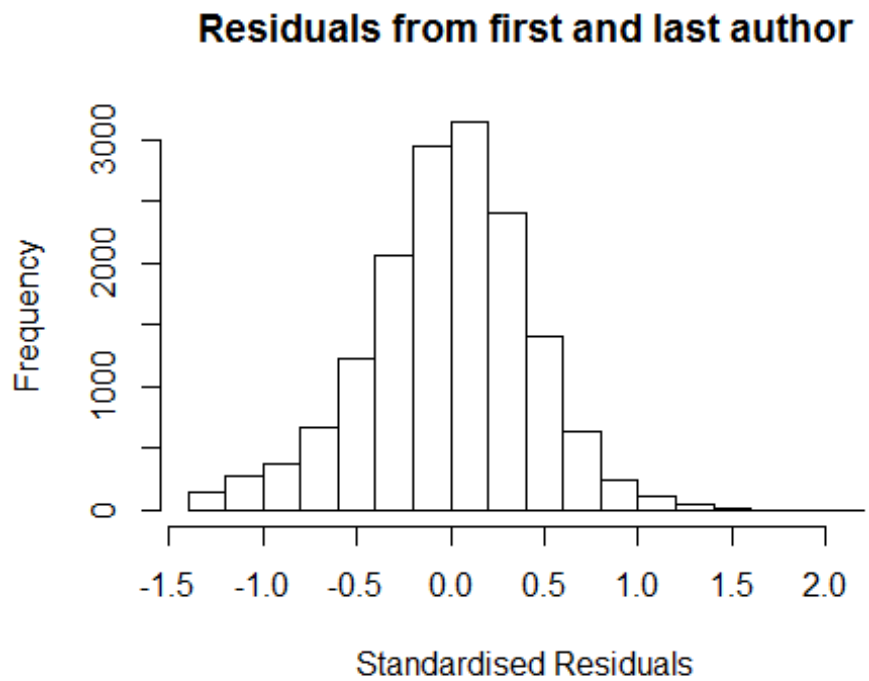



```
## [1] "Female first author team size 2018 geometric mean: 5.36606109244337"
## [1] "Male first author team size 2018 geometric mean: 4.55730130638984"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 310000, p-value = 5e-05
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 5.17163631473886"
## [1] "Male last author team size 2018 geometric mean: 4.92965916786018"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 3e+05, p-value = 0.4
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##                     GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.057  1           1.028
## LastAuthorFemale  1.034  1           1.017
## UniqueAuthors     1.071  4           1.009
## Year              1.061 16           1.002
```



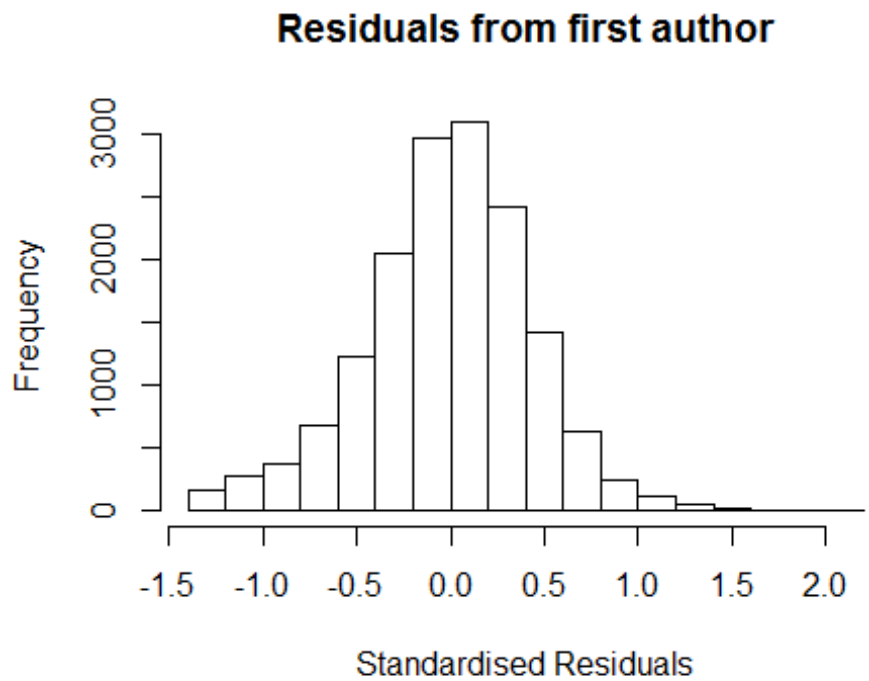

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
UniqueAuthors +
##     Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##     k.max = 1000))
##  \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.36092 -0.25986  0.00679  0.26545  2.17601
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         0.94729    0.02335   40.57  < 2e-16 ***
## FirstAuthorFemale1  0.01239    0.00691    1.79  0.07289 .
## LastAuthorFemale1  -0.01292    0.00698   -1.85  0.06428 .
## UniqueAuthors2      0.20361    0.01985   10.26  < 2e-16 ***
## UniqueAuthors3      0.28554    0.01856   15.38  < 2e-16 ***
## UniqueAuthors4      0.32868    0.01831   17.95  < 2e-16 ***
## UniqueAuthors5      0.40124    0.01682   23.86  < 2e-16 ***
## Year1997            0.00161    0.02534    0.06  0.94929
## Year1998           -0.05130    0.02461   -2.08  0.03713 *
## Year1999            0.00532    0.02404    0.22  0.82497
```

```
## Year2000           -0.02877    0.02383   -1.21  0.22730
## Year2001           -0.04807    0.02373   -2.03  0.04276 *
## Year2002           -0.07673    0.02218   -3.46  0.00054 ***
## Year2003           -0.08623    0.02238   -3.85  0.00012 ***
## Year2004           -0.11062    0.02214   -5.00  5.9e-07 ***
## Year2005           -0.14182    0.02231   -6.36  2.1e-10 ***
## Year2006           -0.11817    0.02183   -5.41  6.3e-08 ***
## Year2007           -0.09092    0.02197   -4.14  3.5e-05 ***
## Year2008           -0.07436    0.02240   -3.32  0.00090 ***
## Year2009           -0.07723    0.02188   -3.53  0.00042 ***
## Year2010           -0.07245    0.02152   -3.37  0.00076 ***
## Year2011           -0.08621    0.02177   -3.96  7.5e-05 ***
## Year2012           -0.08467    0.02161   -3.92  8.9e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.393
## Multiple R-squared:  0.0827, Adjusted R-squared:  0.0814
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
##  8 observations c(273,1325,1340,9858,10210,11243,11557,14221)
##   are outliers with |weight| = 0 ( < 6.3e-06);
##  1366 weights are ~= 1. The remaining 14379 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.000   0.862   0.951   0.893   0.986   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          6.35e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.031  1           1.015
## LastAuthorFemale  1.026  1           1.013
## Year              1.017 16           1.001
```



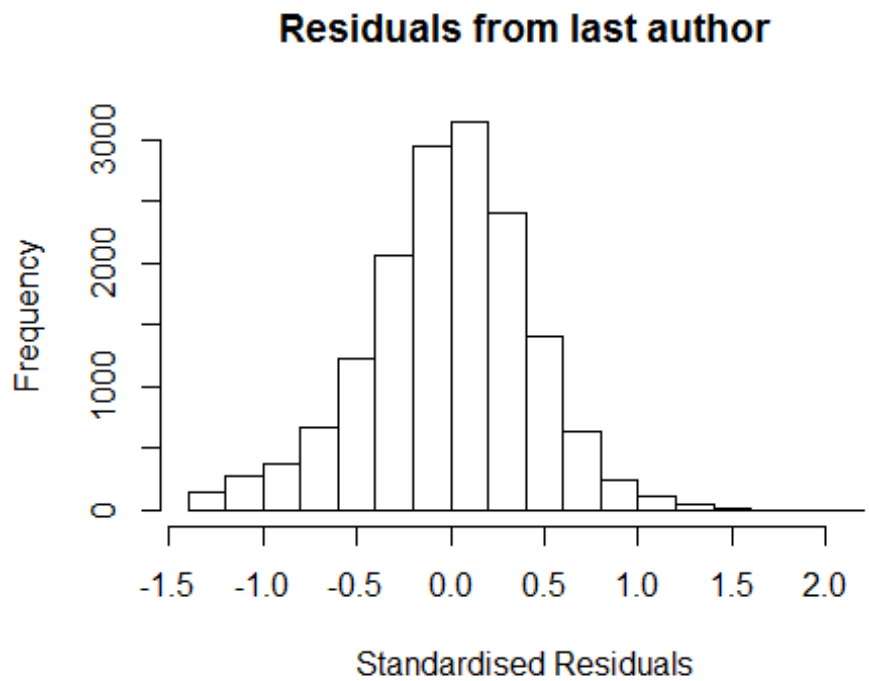

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.27678 -0.27193  0.00975  0.26994  2.08207
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)          1.22093    0.01941   62.89  < 2e-16 ***
## FirstAuthorFemale1   0.04858    0.00706    6.88  6.1e-12 ***
## LastAuthorFemale1   -0.01244    0.00715   -1.74  0.08186 .
## Year1997            -0.00971    0.02649   -0.37  0.71383
## Year1998            -0.02945    0.02518   -1.17  0.24222
## Year1999             0.00728    0.02519    0.29  0.77265
## Year2000            -0.01218    0.02518   -0.48  0.62863
## Year2001            -0.03448    0.02492   -1.38  0.16639
## Year2002            -0.04452    0.02338   -1.90  0.05684 .
## Year2003            -0.04723    0.02367   -2.00  0.04598 *
## Year2004            -0.06399    0.02328   -2.75  0.00599 **
## Year2005            -0.10304    0.02347   -4.39  1.1e-05 ***
```

```
## Year2006           -0.08602    0.02299   -3.74  0.00018 ***
## Year2007           -0.05372    0.02298   -2.34  0.01941 *
## Year2008           -0.04274    0.02349   -1.82  0.06881 .
## Year2009           -0.03663    0.02313   -1.58  0.11318
## Year2010           -0.03126    0.02277   -1.37  0.16997
## Year2011           -0.04854    0.02292   -2.12  0.03419 *
## Year2012           -0.04643    0.02259   -2.06  0.03983 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.402
## Multiple R-squared:  0.00651,    Adjusted R-squared:  0.00538
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
##  6 observations c(273,1325,1506,9858,10210,14221)
##   are outliers with |weight| = 0 ( < 6.3e-06);
##  1345 weights are ~= 1. The remaining 14402 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.000   0.863   0.950   0.890   0.986   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          6.35e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.012  1           1.006
## Year              1.012 16           1.000
```




```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.27145 -0.27096  0.00998  0.27091  2.08481
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)          1.21819    0.01931   63.07  < 2e-16 ***
## FirstAuthorFemale1   0.04631    0.00700    6.61  3.9e-11 ***
## Year1997            -0.01017    0.02649   -0.38  0.70093
## Year1998            -0.02959    0.02519   -1.17  0.24025
## Year1999             0.00696    0.02519    0.28  0.78238
## Year2000            -0.01222    0.02519   -0.49  0.62755
## Year2001            -0.03482    0.02491   -1.40  0.16212
## Year2002            -0.04522    0.02337   -1.93  0.05305 .
## Year2003            -0.04810    0.02367   -2.03  0.04217 *
## Year2004            -0.06445    0.02328   -2.77  0.00565 **
## Year2005            -0.10384    0.02346   -4.43  9.7e-06 ***
## Year2006            -0.08667    0.02299   -3.77  0.00016 ***
```

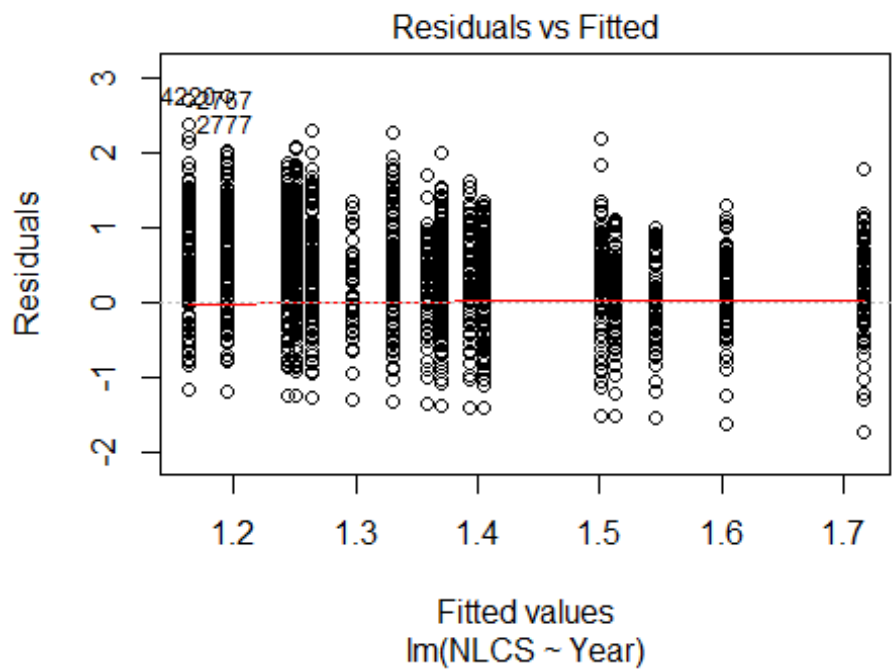
```
## Year2007           -0.05454    0.02299   -2.37  0.01767 *
## Year2008           -0.04354    0.02349   -1.85  0.06382 .
## Year2009           -0.03774    0.02313   -1.63  0.10285
## Year2010           -0.03204    0.02278   -1.41  0.15968
## Year2011           -0.04946    0.02292   -2.16  0.03097 *
## Year2012           -0.04754    0.02259   -2.10  0.03535 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.402
## Multiple R-squared:  0.00634,    Adjusted R-squared:  0.00527
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
##  7 observations c(273,1325,1340,1506,9858,10210,14221)
##   are outliers with |weight| <= 3.5e-06 ( < 6.3e-06);
##  1329 weights are ~= 1. The remaining 14417 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0058  0.8620  0.9500  0.8900  0.9860  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          6.35e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##                   GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.007  1           1.004
## Year             1.007 16           1.000
```



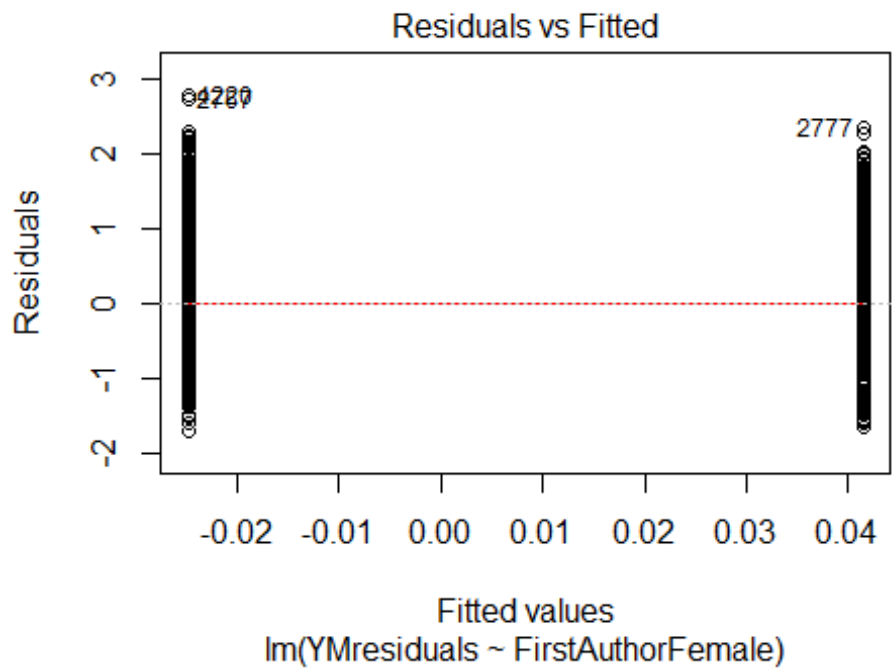

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.2490 -0.2727  0.0109  0.2711  2.0890
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.23748    0.01928   64.18  < 2e-16 ***
## LastAuthorFemale1 -0.00293    0.00710   -0.41  0.67922
## Year1997          -0.00816    0.02657   -0.31  0.75882
## Year1998          -0.02519    0.02527   -1.00  0.31889
## Year1999           0.01152    0.02525    0.46  0.64804
## Year2000          -0.00842    0.02522   -0.33  0.73863
## Year2001          -0.03167    0.02494   -1.27  0.20414
## Year2002          -0.04041    0.02343   -1.72  0.08465 .
## Year2003          -0.04267    0.02374   -1.80  0.07227 .
## Year2004          -0.05754    0.02334   -2.47  0.01370 *
## Year2005          -0.09487    0.02349   -4.04  5.4e-05 ***
## Year2006          -0.08099    0.02304   -3.52  0.00044 ***
```

```
## Year2007          -0.04901    0.02305   -2.13  0.03349 *
## Year2008          -0.03677    0.02349   -1.57  0.11749
## Year2009          -0.02973    0.02317   -1.28  0.19947
## Year2010          -0.02480    0.02285   -1.09  0.27773
## Year2011          -0.04258    0.02297   -1.85  0.06382 .
## Year2012          -0.03919    0.02262   -1.73  0.08321 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.403
## Multiple R-squared:  0.00337,    Adjusted R-squared:  0.00229
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
##  6 observations c(273,1325,1506,9858,10210,14221)
##   are outliers with |weight| = 0 ( < 6.3e-06);
##  1354 weights are ~= 1. The remaining 14393 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0008  0.8630  0.9500  0.8900  0.9860  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          6.35e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                  psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Sample size for the above analysis:  15753"
## [1] ""
## [1] ""
## [1] "##################################"
## [1] "Analysis of AJSC 2714"
## [1] "##################################"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  148  169  162  174  145  155  193  126  129  155  222  232  271  326  436
## 2011 2012
##  434  440
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
```

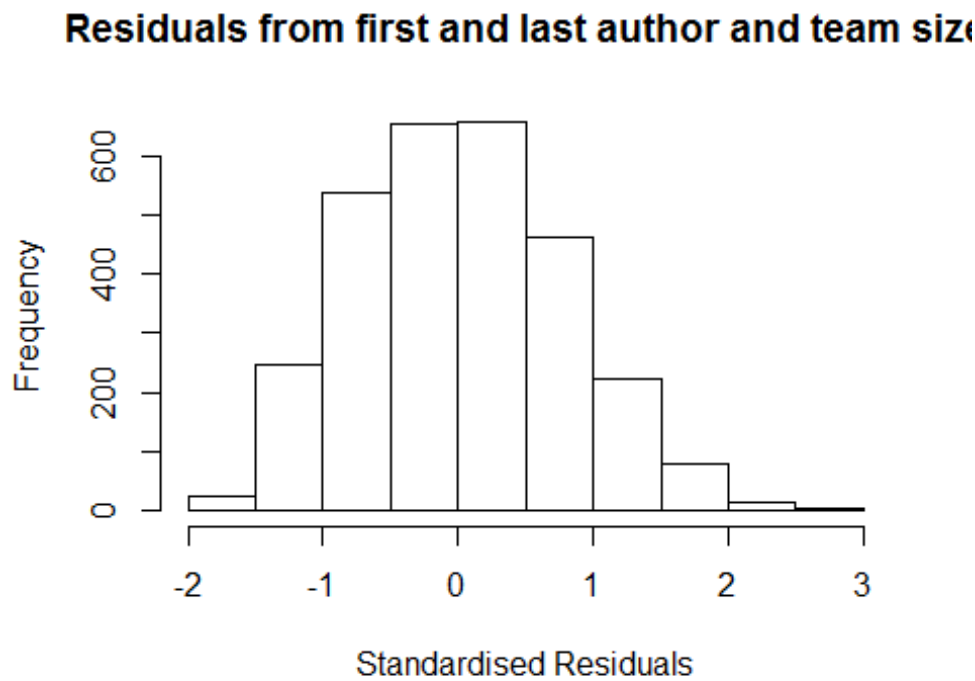
```
##  106  107  109  114   62   54  170  113  117  131  207  198  235  284  368
## 2011 2012
##  393  390
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##   95   92  101  105   56   48  155  105  110  118  190  186  222  265  347
## 2011 2012
##  362  345
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
##  Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 140, df = 16, p-value <2e-16
```




```
##
##  Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 0.03, df = 1, p-value = 0.9
```

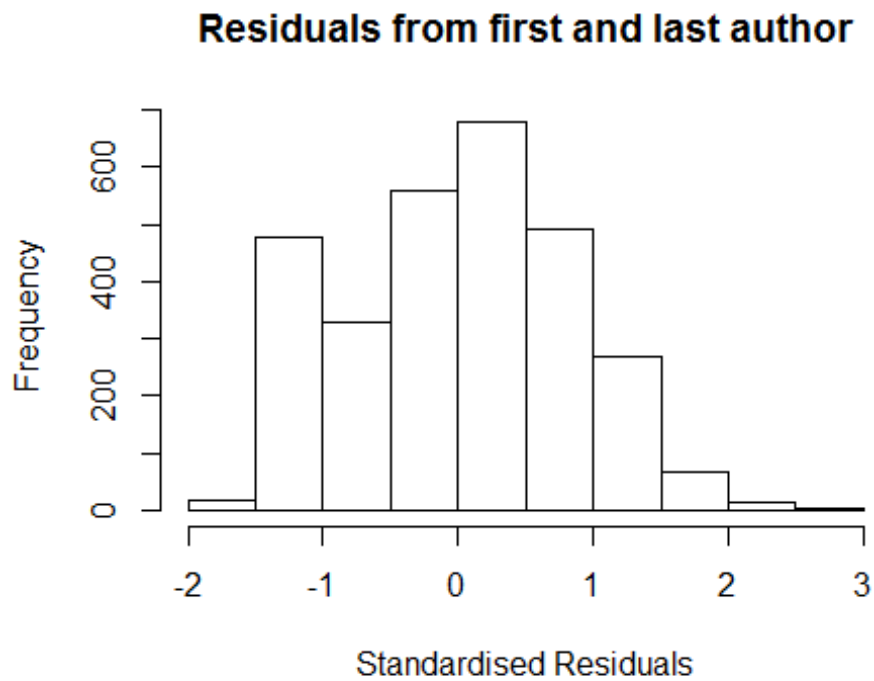



```
## [1] "Female first author team size 2018 geometric mean: 3.22913483535697"
## [1] "Male first author team size 2018 geometric mean: 2.62133286721163"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 18000, p-value = 0.004
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 2.86990917392614"
## [1] "Male last author team size 2018 geometric mean: 2.93042379458021"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 14000, p-value = 0.8
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.140  1           1.068
## LastAuthorFemale  1.151  1           1.073
## UniqueAuthors     1.194  4           1.022
## Year              1.204 16           1.006
```



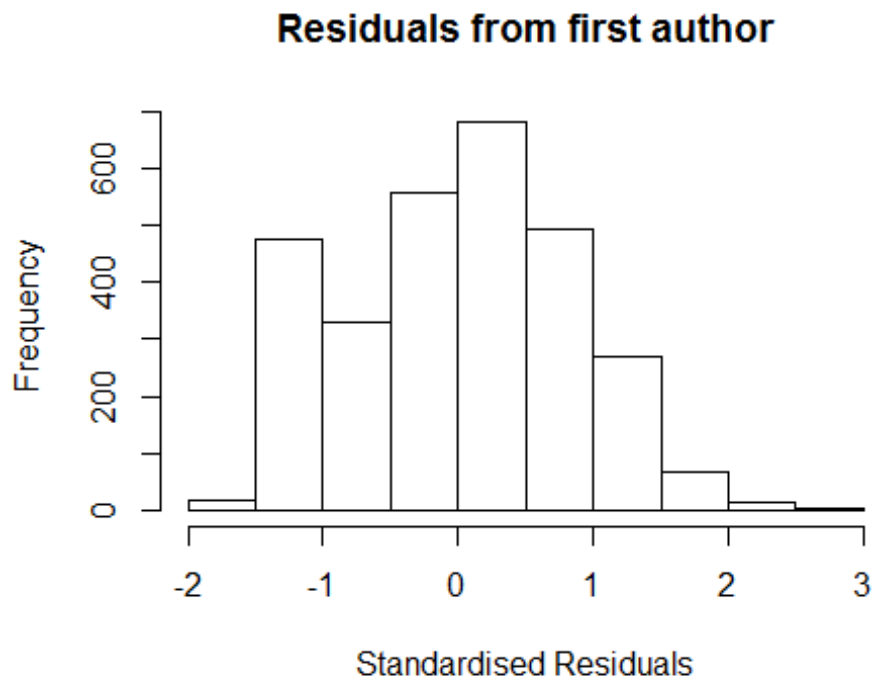

```
## [1] "List of  3 outliers with residuals above  2.5"
##         ScopusId  NLCS Year OneField Fields residuals
## 2870 78149400477 3.305 2010     2714      2     2.552
## 4031 84858246427 3.181 2012     2714      1     2.628
## 4097 84855551703 3.129 2012     2714      1     2.576
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
UniqueAuthors +
##     Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##     k.max = 1000))
##  \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.86240 -0.58511 -0.00943  0.54435  2.62837
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         0.80981    0.07030   11.52   <2e-16 ***
## FirstAuthorFemale1  0.04124    0.03284    1.26   0.2092
## LastAuthorFemale1   0.00187    0.03324    0.06   0.9552
## UniqueAuthors2      0.50405    0.05043   10.00   <2e-16 ***
## UniqueAuthors3      0.66961    0.05010   13.37   <2e-16 ***
## UniqueAuthors4      0.77063    0.05237   14.72   <2e-16 ***
## UniqueAuthors5      0.96012    0.04247   22.61   <2e-16 ***
## Year1997            0.26534    0.09100    2.92   0.0036 **
```

```
## Year1998            0.19059    0.08264    2.31   0.0212 *
## Year1999            0.11020    0.08355    1.32   0.1873
## Year2000            0.23885    0.10398    2.30   0.0217 *
## Year2001           -0.03263    0.13145   -0.25   0.8040
## Year2002            0.01803    0.08671    0.21   0.8353
## Year2003            0.03747    0.09408    0.40   0.6904
## Year2004           -0.04799    0.09377   -0.51   0.6088
## Year2005            0.08525    0.09735    0.88   0.3813
## Year2006           -0.05141    0.07881   -0.65   0.5143
## Year2007           -0.12518    0.08760   -1.43   0.1531
## Year2008           -0.15200    0.07889   -1.93   0.0541 .
## Year2009           -0.22451    0.08002   -2.81   0.0051 **
## Year2010           -0.09975    0.08200   -1.22   0.2239
## Year2011           -0.17198    0.08081   -2.13   0.0334 *
## Year2012           -0.25719    0.08134   -3.16   0.0016 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.758
## Multiple R-squared:  0.191,  Adjusted R-squared:  0.185
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
##  229 weights are ~= 1. The remaining 2673 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.205   0.867   0.944   0.906   0.985   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          3.45e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##                     GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.086  1           1.042
## LastAuthorFemale  1.093  1           1.045
## Year              1.037 16           1.001
```



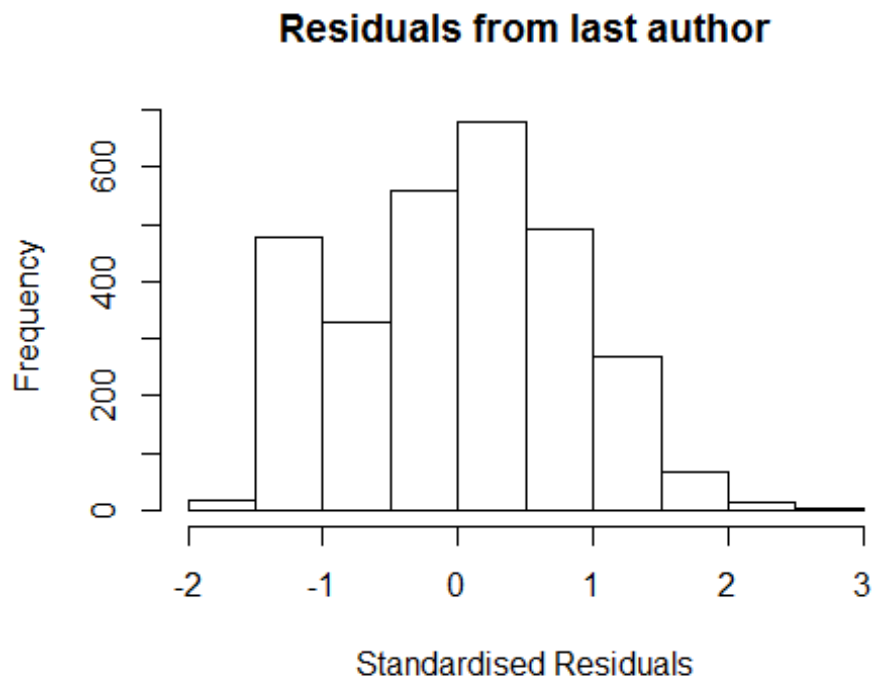

```
## [1] "List of  2 outliers with residuals above  2.5"
##         ScopusId NLCS Year OneField Fields residuals
## 2767 67649610853 3.85 2009     2714      1     2.729
## 4220 84869207969 3.94 2012     2714      1     2.822
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.6793 -0.6258  0.0454  0.6016  2.8220
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.325340   0.062596   21.17  < 2e-16 ***
## FirstAuthorFemale1  0.093660   0.035074    2.67  0.00762 **
## LastAuthorFemale1   0.001532   0.035220    0.04  0.96532
## Year1997            0.352417   0.091528    3.85  0.00012 ***
## Year1998            0.250377   0.084773    2.95  0.00317 **
## Year1999            0.148528   0.079488    1.87  0.06179 .
## Year2000            0.183669   0.103334    1.78  0.07560 .
## Year2001           -0.067382   0.128895   -0.52  0.60118
## Year2002            0.028319   0.089092    0.32  0.75061
## Year2003            0.000787   0.103412    0.01  0.99393
## Year2004            0.044452   0.100697    0.44  0.65892
```

```
## Year2005            0.168779   0.100401    1.68  0.09286 .
## Year2006            0.044293   0.079826    0.55  0.57903
## Year2007           -0.039952   0.091799   -0.44  0.66344
## Year2008           -0.087706   0.085878   -1.02  0.30721
## Year2009           -0.205532   0.085850   -2.39  0.01673 *
## Year2010           -0.135444   0.084080   -1.61  0.10731
## Year2011           -0.126872   0.081576   -1.56  0.11999
## Year2012           -0.207303   0.082085   -2.53  0.01161 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.887
## Multiple R-squared:  0.0305, Adjusted R-squared:  0.0244
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  236 weights are ~= 1. The remaining 2666 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.290   0.860   0.948   0.918   0.985   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          3.45e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.015  1           1.007
## Year              1.015 16           1.000
```




```
## [1] "List of  2 outliers with residuals above  2.5"
##         ScopusId NLCS Year OneField Fields residuals
## 2767 67649610853 3.85 2009     2714      1     2.729
## 4220 84869207969 3.94 2012     2714      1     2.822
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##    Min     1Q Median     3Q    Max
## -1.678 -0.626  0.045  0.602  2.822
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.325735   0.062219   21.31  < 2e-16 ***
## FirstAuthorFemale1  0.094107   0.033920    2.77  0.00557 **
## Year1997            0.352365   0.091531    3.85  0.00012 ***
## Year1998            0.250418   0.084767    2.95  0.00316 **
## Year1999            0.148484   0.079527    1.87  0.06199 .
## Year2000            0.183853   0.103179    1.78  0.07487 .
## Year2001           -0.067408   0.128908   -0.52  0.60107
## Year2002            0.028365   0.089095    0.32  0.75023
## Year2003            0.000759   0.103431    0.01  0.99415
## Year2004            0.044492   0.100703    0.44  0.65866
## Year2005            0.168934   0.100202    1.69  0.09192 .
```

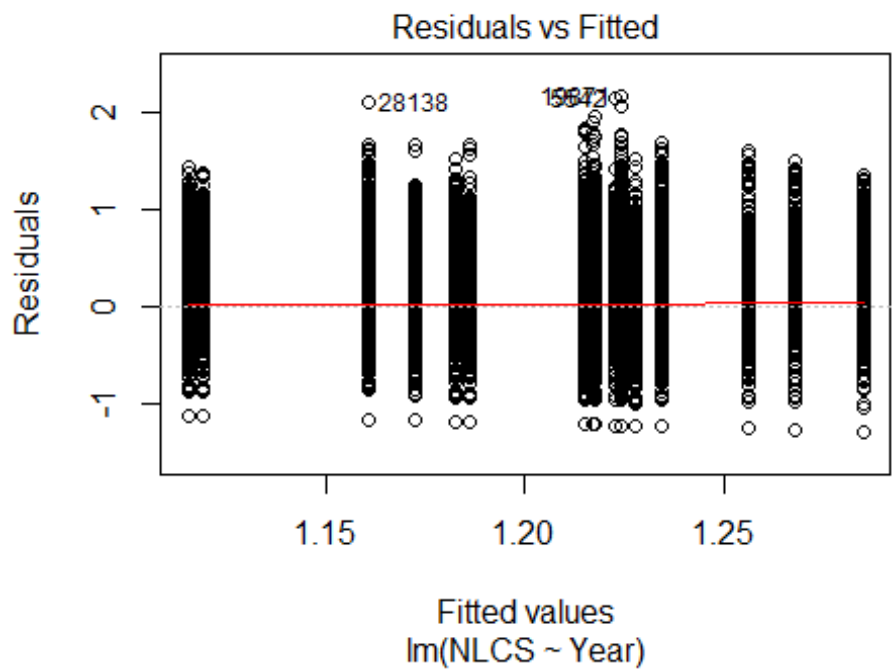
```
## Year2006            0.044230   0.079845    0.55  0.57966
## Year2007           -0.040055   0.091888   -0.44  0.66293
## Year2008           -0.087765   0.085919   -1.02  0.30711
## Year2009           -0.205586   0.085919   -2.39  0.01679 *
## Year2010           -0.135429   0.084074   -1.61  0.10733
## Year2011           -0.126803   0.081542   -1.56  0.12004
## Year2012           -0.207243   0.082035   -2.53  0.01158 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.887
## Multiple R-squared:  0.0305, Adjusted R-squared:  0.0247
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  236 weights are ~= 1. The remaining 2666 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.290   0.860   0.948   0.918   0.985   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          3.45e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##                   GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.021  1           1.010
## Year             1.021 16           1.001
```



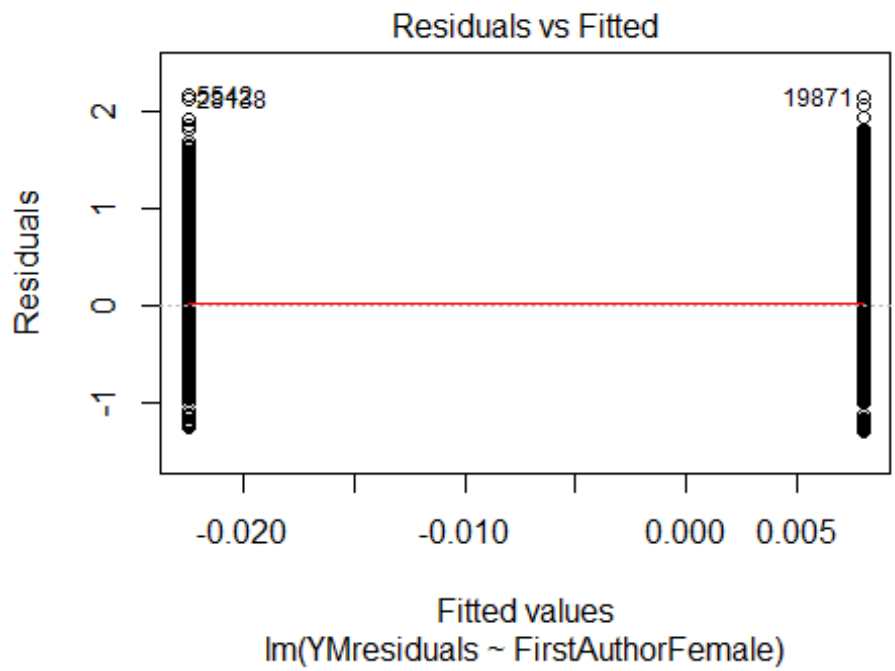

```
## [1] "List of  2 outliers with residuals above  2.5"
##         ScopusId NLCS Year OneField Fields residuals
## 2767 67649610853 3.85 2009     2714      1     2.729
## 4220 84869207969 3.94 2012     2714      1     2.822
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##    Min     1Q Median     3Q    Max
## -1.733 -0.648  0.044  0.599  2.794
##
## Coefficients:
##                   Estimate Std. Error t value Pr(>|t|)
## (Intercept)        1.34692    0.06125   21.99   <2e-16 ***
## LastAuthorFemale1  0.02879    0.03394    0.85   0.3963
## Year1997           0.35752    0.09120    3.92    9e-05 ***
## Year1998           0.24714    0.08438    2.93   0.0034 **
## Year1999           0.15341    0.07903    1.94   0.0523 .
## Year2000           0.18737    0.10421    1.80   0.0723 .
## Year2001          -0.06164    0.12998   -0.47   0.6353
## Year2002           0.03067    0.08862    0.35   0.7293
## Year2003           0.00169    0.10346    0.02   0.9870
## Year2004           0.03736    0.10003    0.37   0.7088
## Year2005           0.17807    0.10011    1.78   0.0754 .
```

```
## Year2006           0.05061    0.07939    0.64   0.5239
## Year2007          -0.03469    0.09142   -0.38   0.7044
## Year2008          -0.08429    0.08583   -0.98   0.3261
## Year2009          -0.20122    0.08565   -2.35   0.0189 *
## Year2010          -0.13220    0.08370   -1.58   0.1143
## Year2011          -0.12194    0.08099   -1.51   0.1323
## Year2012          -0.20141    0.08112   -2.48   0.0131 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.898
## Multiple R-squared:  0.0278, Adjusted R-squared:  0.0221
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  238 weights are ~= 1. The remaining 2664 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.312   0.857   0.951   0.920   0.985   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          3.45e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Sample size for the above analysis:  2902"
## [1] ""
## [1] ""
## [1] "####################################"
## [1] "Analysis of AJSC 2715"
## [1] "####################################"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 1554 1815 1561 1330 1338 1360 1347 1026 1052 1175 1332 1605 1694 1784 1807
## 2011 2012
## 1835 1848
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  526  617  552  644  816  726  957  723  731  832  961 1149 1220 1285 1268
```

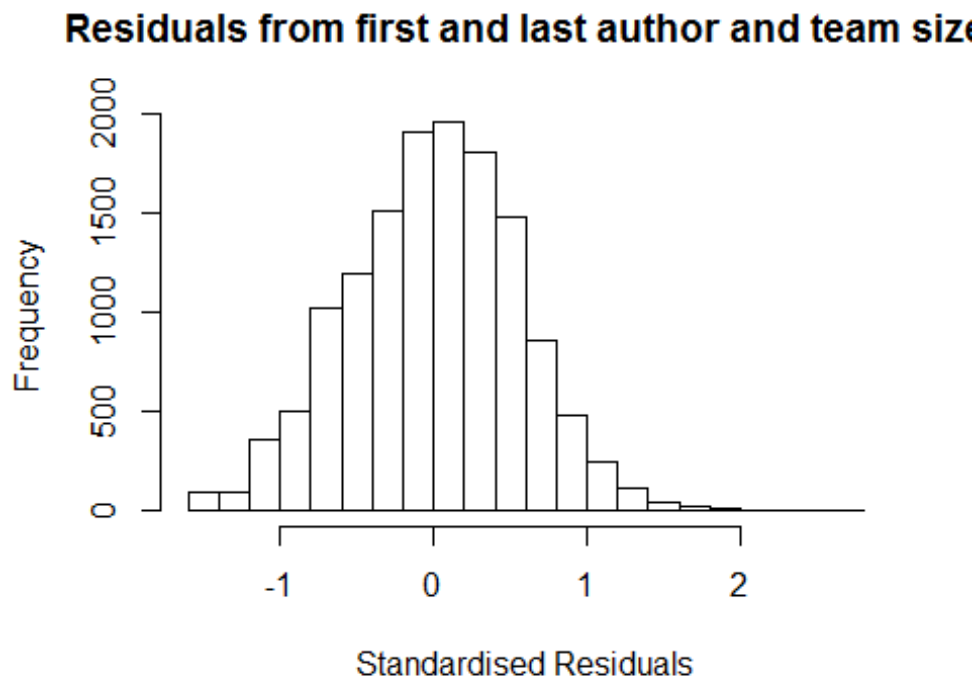
```
## 2011 2012
## 1328 1290
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  467  552  486  585  714  638  818  645  649  720  850  985 1080 1113 1100
## 2011 2012
## 1151 1110
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
##  Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 98, df = 16, p-value = 8e-14
```




```
##
##  Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 0.44, df = 1, p-value = 0.5
```

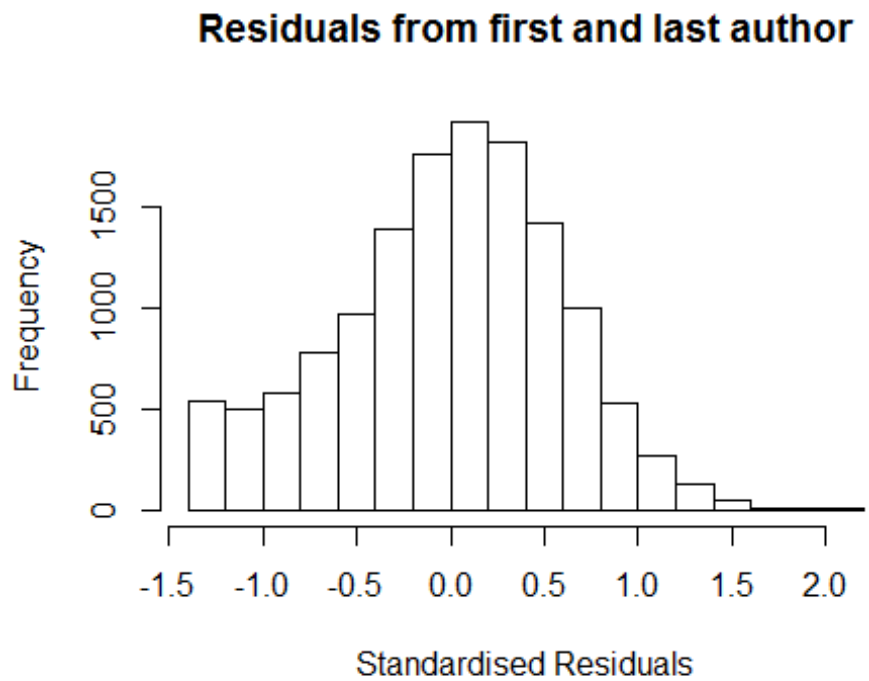



```
## [1] "Female first author team size 2018 geometric mean: 4.68244121839783"
## [1] "Male first author team size 2018 geometric mean: 4.62910756478675"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 190000, p-value = 0.8
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 4.32246997541873"
## [1] "Male last author team size 2018 geometric mean: 4.73513428054028"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 120000, p-value = 0.2
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.056  1           1.028
## LastAuthorFemale  1.072  1           1.035
## UniqueAuthors     1.078  4           1.009
## Year              1.080 16           1.002
```



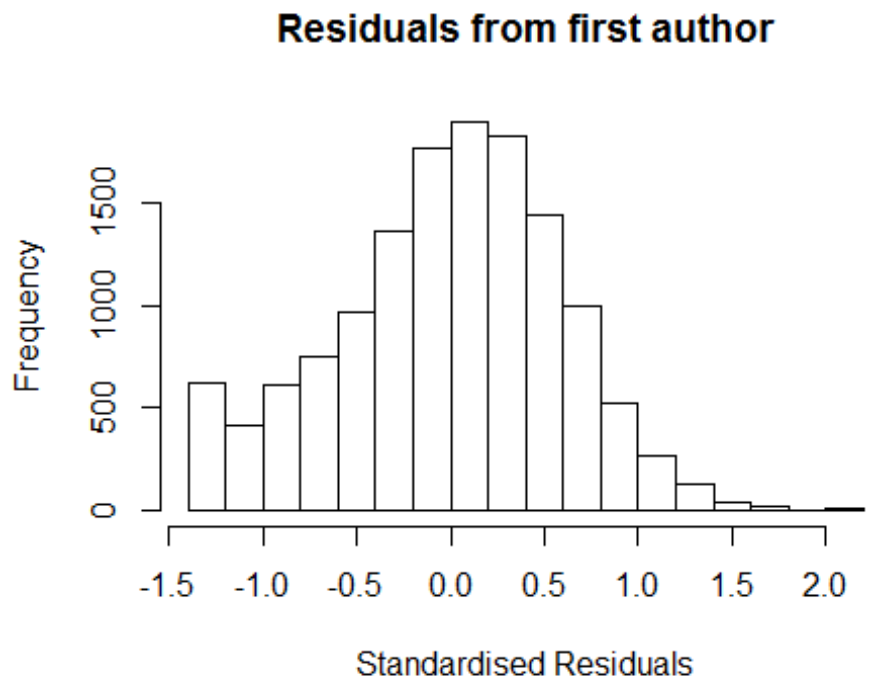

```
## [1] "List of  1 outliers with residuals above  2.5"
##          ScopusId  NLCS Year OneField Fields residuals
## 19871 42949163491 3.382 2008     2715      1     2.603
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
UniqueAuthors +
##     Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##     k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.5678 -0.3767  0.0135  0.3789  2.6031
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)           0.8332     0.0348   23.97  < 2e-16 ***
## FirstAuthorFemale1   -0.0377     0.0111   -3.41  0.00065 ***
## LastAuthorFemale1    -0.0431     0.0131   -3.28  0.00105 **
## UniqueAuthors2        0.2431     0.0218   11.18  < 2e-16 ***
## UniqueAuthors3        0.3341     0.0204   16.35  < 2e-16 ***
## UniqueAuthors4        0.4816     0.0199   24.15  < 2e-16 ***
## UniqueAuthors5        0.6840     0.0172   39.78  < 2e-16 ***
## Year1997              0.0419     0.0414    1.01  0.31119
## Year1998              0.0506     0.0422    1.20  0.23063
## Year1999             -0.0329     0.0375   -0.88  0.37996
```

```
## Year2000             0.0492     0.0368    1.34  0.18115
## Year2001            -0.0203     0.0367   -0.55  0.58081
## Year2002            -0.0803     0.0366   -2.20  0.02813 *
## Year2003            -0.1141     0.0371   -3.07  0.00213 **
## Year2004            -0.0480     0.0377   -1.28  0.20230
## Year2005            -0.1380     0.0369   -3.74  0.00019 ***
## Year2006            -0.0829     0.0368   -2.25  0.02426 *
## Year2007            -0.0588     0.0365   -1.61  0.10677
## Year2008            -0.0543     0.0355   -1.53  0.12626
## Year2009            -0.0452     0.0359   -1.26  0.20792
## Year2010            -0.0771     0.0356   -2.16  0.03045 *
## Year2011            -0.0557     0.0355   -1.57  0.11656
## Year2012            -0.1336     0.0353   -3.78  0.00016 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.55
## Multiple R-squared:  0.168,  Adjusted R-squared:  0.167
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  observation 9033 is an outlier with |weight| = 0 ( < 7.3e-06);
##  1134 weights are ~= 1. The remaining 12528 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0424  0.8690  0.9490  0.9070  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          7.32e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.038  1           1.019
## LastAuthorFemale  1.033  1           1.016
## Year              1.029 16           1.001
```



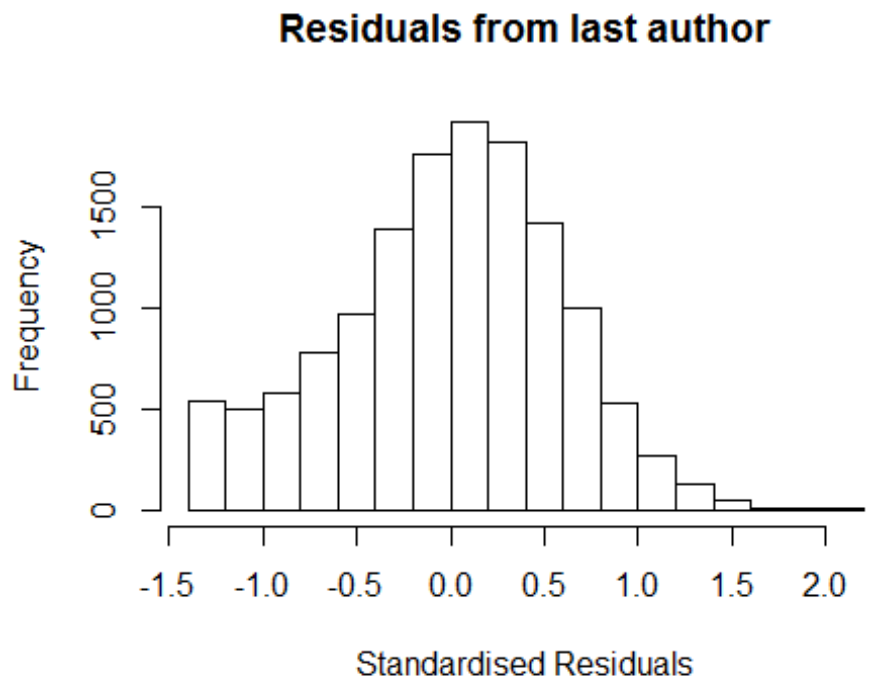

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.3222 -0.3917  0.0354  0.3998  2.1366
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)          1.24797    0.03150   39.61   <2e-16 ***
## FirstAuthorFemale1 -0.02373    0.01210   -1.96    0.050 *
## LastAuthorFemale1  -0.11401    0.01466   -7.78    8e-15 ***
## Year1997            0.04105    0.04229    0.97    0.332
## Year1998            0.06275    0.04299    1.46    0.144
## Year1999           -0.00183    0.03837   -0.05    0.962
## Year2000            0.07424    0.03793    1.96    0.050 .
## Year2001            0.01034    0.03750    0.28    0.783
## Year2002           -0.05265    0.03822   -1.38    0.168
## Year2003           -0.09585    0.03967   -2.42    0.016 *
## Year2004           -0.02350    0.03955   -0.59    0.552
## Year2005           -0.09892    0.03888   -2.54    0.011 *
```

```
## Year2006           -0.02608    0.03854   -0.68    0.498
## Year2007            0.00594    0.03836    0.15    0.877
## Year2008            0.00694    0.03718    0.19    0.852
## Year2009            0.02741    0.03739    0.73    0.464
## Year2010            0.01321    0.03711    0.36    0.722
## Year2011            0.01879    0.03659    0.51    0.608
## Year2012           -0.05217    0.03643   -1.43    0.152
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.58
## Multiple R-squared:  0.0119, Adjusted R-squared:  0.0106
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  1155 weights are ~= 1. The remaining 12508 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.146   0.861   0.949   0.901   0.985   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          7.32e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.021  1           1.010
## Year              1.021 16           1.001
```




```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.3112 -0.3931  0.0335  0.4009  2.1728
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)          1.23666    0.03149   39.27  < 2e-16 ***
## FirstAuthorFemale1 -0.04588    0.01221   -3.76  0.00017 ***
## Year1997             0.04195    0.04241    0.99  0.32257
## Year1998             0.06102    0.04308    1.42  0.15670
## Year1999            -0.00456    0.03841   -0.12  0.90552
## Year2000             0.07450    0.03794    1.96  0.04957 *
## Year2001             0.01013    0.03746    0.27  0.78696
## Year2002            -0.05710    0.03827   -1.49  0.13575
## Year2003            -0.09986    0.03961   -2.52  0.01171 *
## Year2004            -0.02469    0.03962   -0.62  0.53315
## Year2005            -0.10411    0.03888   -2.68  0.00742 **
## Year2006            -0.03207    0.03850   -0.83  0.40483
```

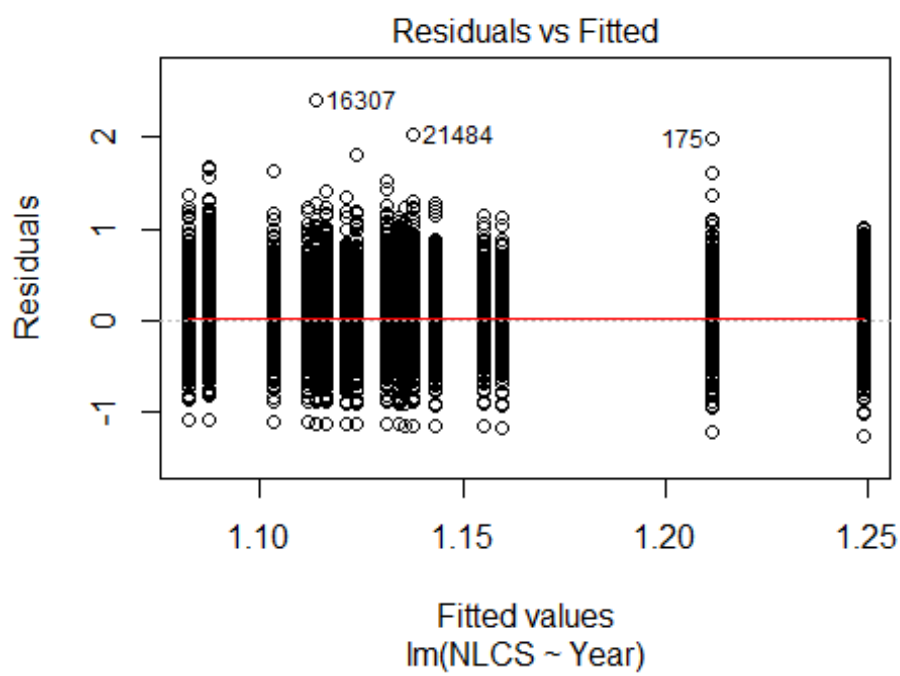
```
## Year2007            0.00185    0.03829    0.05  0.96151
## Year2008            0.00524    0.03720    0.14  0.88799
## Year2009            0.02082    0.03734    0.56  0.57711
## Year2010            0.00910    0.03708    0.25  0.80611
## Year2011            0.01653    0.03657    0.45  0.65129
## Year2012           -0.05354    0.03640   -1.47  0.14132
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.581
## Multiple R-squared:  0.0069, Adjusted R-squared:  0.00566
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  1160 weights are ~= 1. The remaining 12503 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.132   0.860   0.949   0.901   0.985   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          7.32e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##                   GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.014  1           1.007
## Year             1.014 16           1.000
```



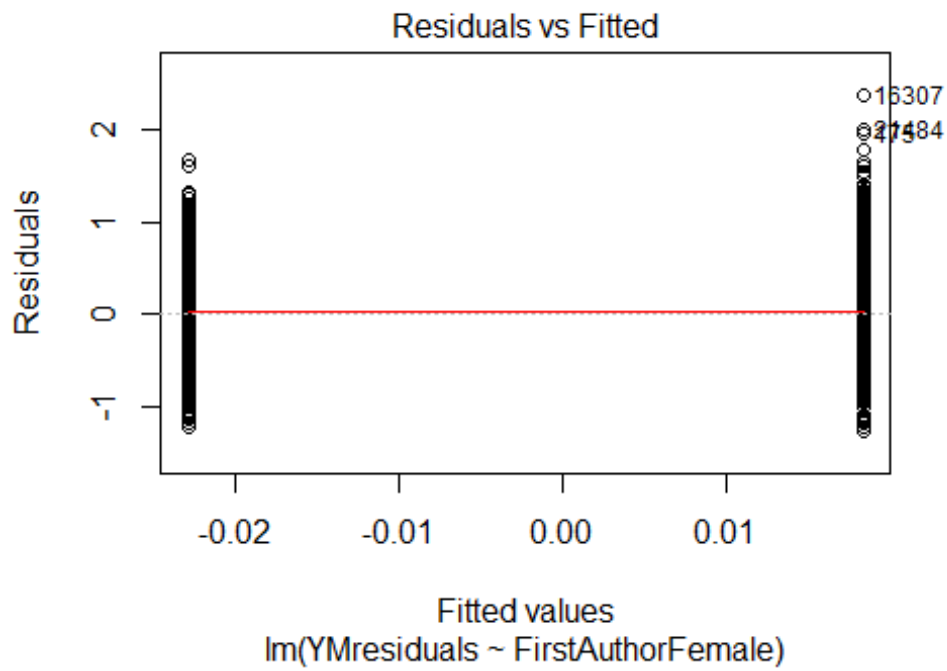

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.3176 -0.3935  0.0356  0.4014  2.1329
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.24442    0.03153   39.47   <2e-16 ***
## LastAuthorFemale1 -0.12029    0.01463   -8.22   <2e-16 ***
## Year1997            0.03996    0.04236    0.94   0.3456
## Year1998            0.06262    0.04307    1.45   0.1460
## Year1999           -0.00156    0.03846   -0.04   0.9677
## Year2000            0.07314    0.03801    1.92   0.0543 .
## Year2001            0.00937    0.03756    0.25   0.8031
## Year2002           -0.05389    0.03830   -1.41   0.1594
## Year2003           -0.09753    0.03972   -2.46   0.0141 *
## Year2004           -0.02543    0.03959   -0.64   0.5207
## Year2005           -0.10115    0.03891   -2.60   0.0093 **
## Year2006           -0.02747    0.03859   -0.71   0.4766
```

```
## Year2007           0.00413    0.03840    0.11   0.9144
## Year2008           0.00468    0.03720    0.13   0.8999
## Year2009           0.02527    0.03742    0.68   0.4995
## Year2010           0.01117    0.03716    0.30   0.7636
## Year2011           0.01606    0.03662    0.44   0.6611
## Year2012          -0.05529    0.03646   -1.52   0.1294
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.581
## Multiple R-squared:  0.0115, Adjusted R-squared:  0.0103
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
##  1131 weights are ~= 1. The remaining 12532 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.149   0.860   0.949   0.902   0.985   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          7.32e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Sample size for the above analysis:  13663"
## [1] ""
## [1] ""
## [1] "####################################"
## [1] "Analysis of AJSC 2716"
## [1] "####################################"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  979 1113 1270 1055 1300 1254 1056 1021 1106 1342 1292 1356 1348 1415 1431
## 2011 2012
## 1723 1881
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  602  652  342  649  324  568  662  671  685  837  837  888  921  920  952
## 2011 2012
```

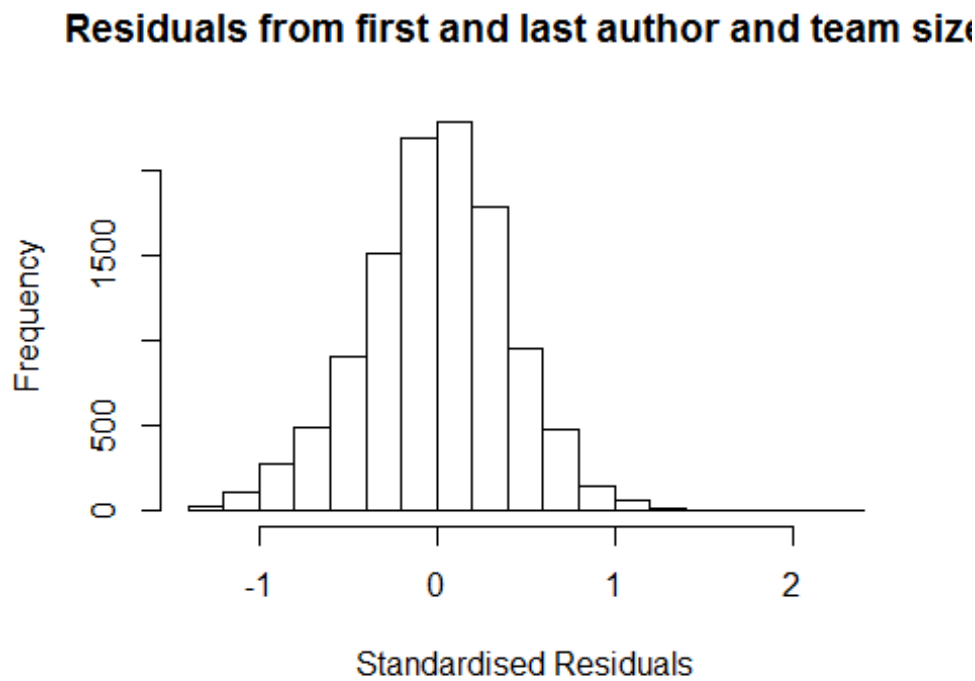
```
## 1182 1272
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  526  589  302  565  288  484  576  561  575  710  726  778  795  785  829
## 2011 2012
## 1018 1110
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
##  Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 110, df = 16, p-value = 1e-15
```




```
##
##  Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 54, df = 1, p-value = 2e-13
```

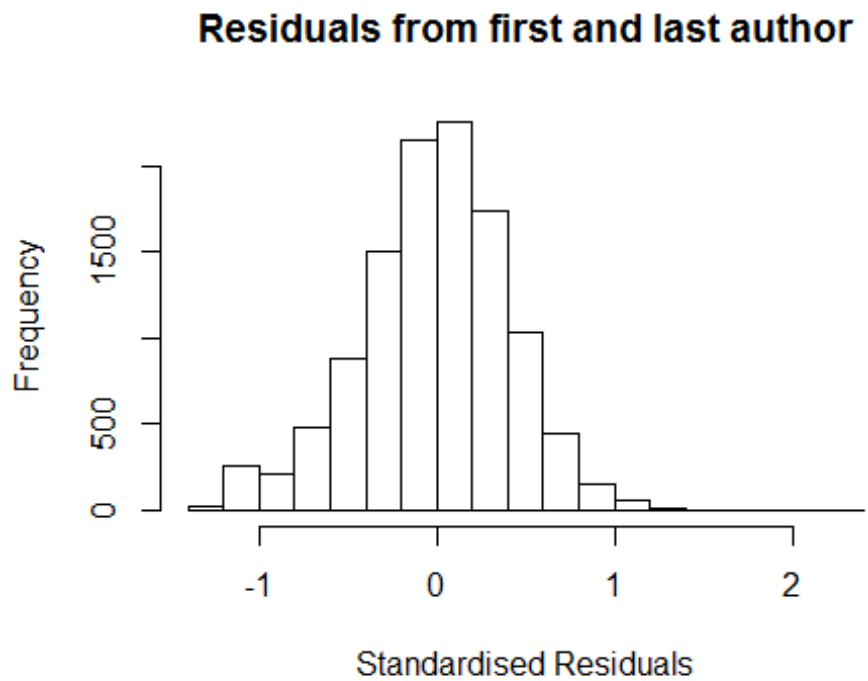



```
## [1] "Female first author team size 2018 geometric mean: 5.67998979577716"
## [1] "Male first author team size 2018 geometric mean: 5.17890869002327"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 140000, p-value = 0.02
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 5.54839619787746"
## [1] "Male last author team size 2018 geometric mean: 5.39671605124862"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 130000, p-value = 0.3
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.040  1           1.020
## LastAuthorFemale  1.031  1           1.016
## UniqueAuthors     1.092  4           1.011
## Year              1.096 16           1.003
```



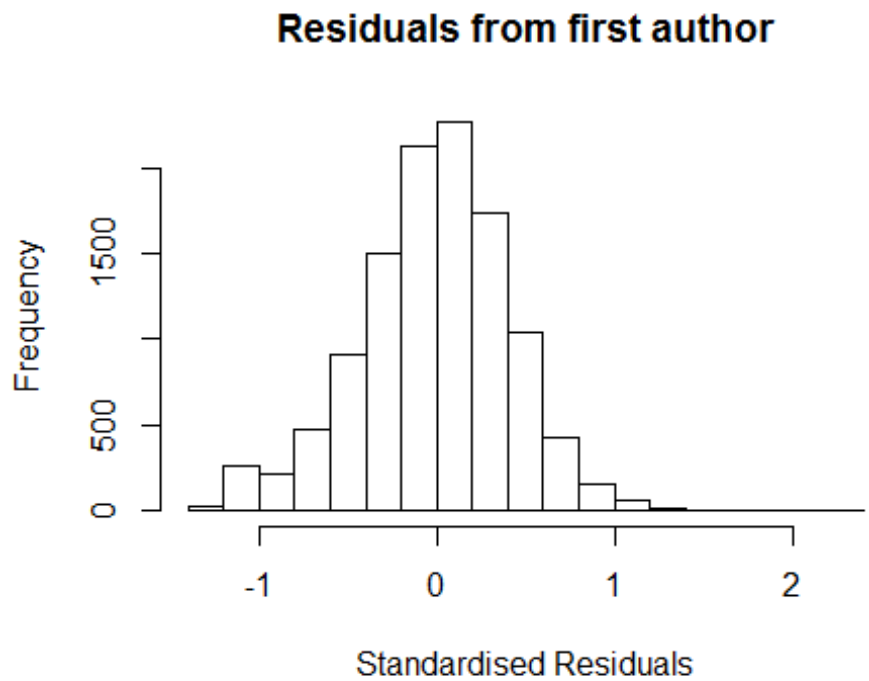

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
UniqueAuthors +
##     Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##     k.max = 1000))
##  \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.35972 -0.26220  0.00987  0.26109  2.38013
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.00353    0.02734   36.70  < 2e-16 ***
## FirstAuthorFemale1 -0.05497    0.00788   -6.98  3.2e-12 ***
## LastAuthorFemale1  -0.05770    0.00837   -6.89  5.8e-12 ***
## UniqueAuthors2      0.23051    0.02187   10.54  < 2e-16 ***
## UniqueAuthors3      0.23847    0.02176   10.96  < 2e-16 ***
## UniqueAuthors4      0.23911    0.02166   11.04  < 2e-16 ***
## UniqueAuthors5      0.34691    0.02021   17.17  < 2e-16 ***
## Year1997            0.06425    0.02772    2.32  0.02048 *
## Year1998           -0.03943    0.03235   -1.22  0.22286
## Year1999           -0.08387    0.02521   -3.33  0.00088 ***
```

```
## Year2000           -0.00943    0.03205   -0.29  0.76851
## Year2001           -0.08522    0.02719   -3.13  0.00173 **
## Year2002           -0.07556    0.02599   -2.91  0.00366 **
## Year2003           -0.10317    0.02554   -4.04  5.4e-05 ***
## Year2004           -0.12233    0.02535   -4.83  1.4e-06 ***
## Year2005           -0.13912    0.02473   -5.62  1.9e-08 ***
## Year2006           -0.09391    0.02445   -3.84  0.00012 ***
## Year2007           -0.11076    0.02427   -4.56  5.1e-06 ***
## Year2008           -0.08652    0.02499   -3.46  0.00054 ***
## Year2009           -0.11479    0.02532   -4.53  5.8e-06 ***
## Year2010           -0.10291    0.02436   -4.22  2.4e-05 ***
## Year2011           -0.11728    0.02373   -4.94  7.9e-07 ***
## Year2012           -0.13780    0.02426   -5.68  1.4e-08 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.388
## Multiple R-squared:  0.073,  Adjusted R-squared:  0.0712
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
##  2 observations c(6244,8507) are outliers with |weight| = 0 ( < 8.9e-06);
##  908 weights are ~= 1. The remaining 10307 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0146  0.8660  0.9500  0.8990  0.9860  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          8.92e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.031  1           1.015
## LastAuthorFemale  1.030  1           1.015
## Year              1.019 16           1.001
```



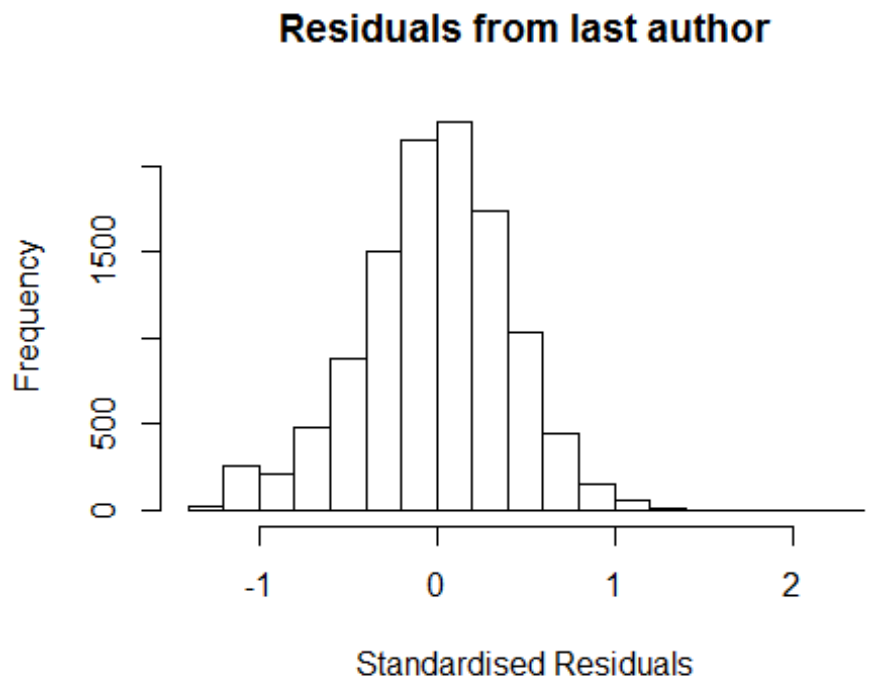

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.30539 -0.26365  0.00815  0.25991  2.34954
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)          1.25352    0.02050   61.14  < 2e-16 ***
## FirstAuthorFemale1 -0.03661    0.00797   -4.59  4.5e-06 ***
## LastAuthorFemale1  -0.06087    0.00852   -7.14  9.8e-13 ***
## Year1997            0.05187    0.02751    1.89  0.05939 .
## Year1998           -0.06265    0.03135   -2.00  0.04569 *
## Year1999           -0.10095    0.02516   -4.01  6.1e-05 ***
## Year2000           -0.02382    0.03118   -0.76  0.44484
## Year2001           -0.07506    0.02713   -2.77  0.00567 **
## Year2002           -0.06535    0.02615   -2.50  0.01246 *
## Year2003           -0.09465    0.02605   -3.63  0.00028 ***
## Year2004           -0.10239    0.02554   -4.01  6.2e-05 ***
## Year2005           -0.12691    0.02493   -5.09  3.6e-07 ***
```

```
## Year2006           -0.07290    0.02448   -2.98  0.00291 **
## Year2007           -0.09106    0.02431   -3.75  0.00018 ***
## Year2008           -0.06150    0.02522   -2.44  0.01477 *
## Year2009           -0.08469    0.02570   -3.30  0.00098 ***
## Year2010           -0.06811    0.02445   -2.79  0.00534 **
## Year2011           -0.08251    0.02391   -3.45  0.00056 ***
## Year2012           -0.10668    0.02469   -4.32  1.6e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.392
## Multiple R-squared:  0.0178, Adjusted R-squared:  0.0162
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  2 observations c(6244,8507) are outliers with |weight| = 0 ( < 8.9e-06);
##  970 weights are ~= 1. The remaining 10245 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0059  0.8640  0.9510  0.8950  0.9860  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          8.92e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##                   GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.01  1           1.005
## Year              1.01 16           1.000
```




```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.29224 -0.26432  0.00925  0.26118  2.36268
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.24070    0.02040   60.82  < 2e-16 ***
## FirstAuthorFemale1 -0.04750    0.00792   -6.00  2.1e-09 ***
## Year1997            0.05154    0.02763    1.87  0.06211 .
## Year1998           -0.06245    0.03131   -1.99  0.04614 *
## Year1999           -0.10068    0.02524   -3.99  6.7e-05 ***
## Year2000           -0.02676    0.03132   -0.85  0.39294
## Year2001           -0.07646    0.02720   -2.81  0.00495 **
## Year2002           -0.06661    0.02620   -2.54  0.01101 *
## Year2003           -0.09345    0.02605   -3.59  0.00034 ***
## Year2004           -0.10138    0.02556   -3.97  7.4e-05 ***
## Year2005           -0.12763    0.02499   -5.11  3.3e-07 ***
## Year2006           -0.07345    0.02454   -2.99  0.00277 **
```

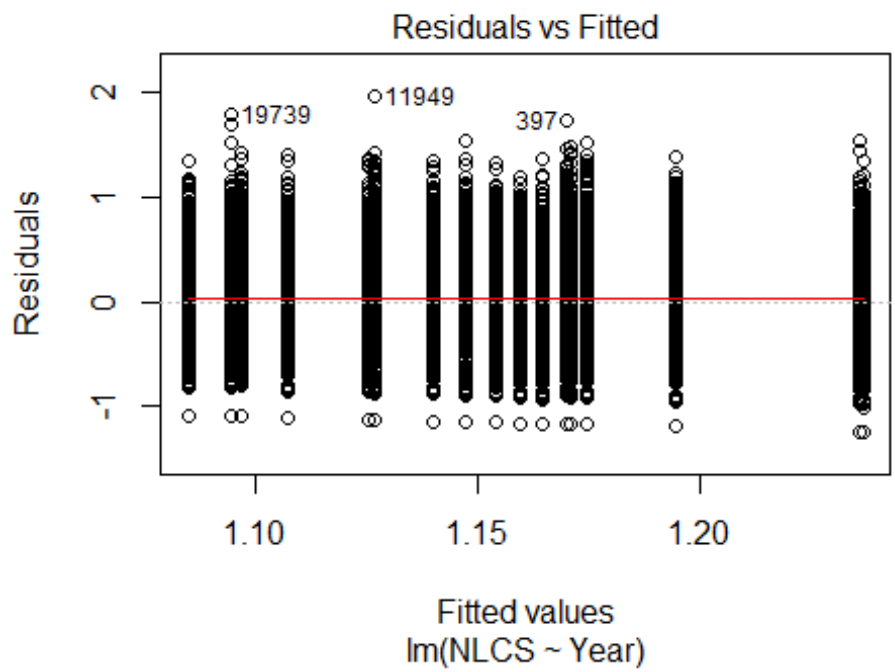
```
## Year2007           -0.09137    0.02435   -3.75  0.00018 ***
## Year2008           -0.06337    0.02531   -2.50  0.01229 *
## Year2009           -0.08747    0.02573   -3.40  0.00068 ***
## Year2010           -0.07138    0.02448   -2.92  0.00356 **
## Year2011           -0.08564    0.02395   -3.58  0.00035 ***
## Year2012           -0.10906    0.02474   -4.41  1.0e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.393
## Multiple R-squared:  0.0131, Adjusted R-squared:  0.0116
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  2 observations c(6244,8507) are outliers with |weight| = 0 ( < 8.9e-06);
##  990 weights are ~= 1. The remaining 10225 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0049  0.8640  0.9500  0.8950  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          8.92e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##                   GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.009  1           1.005
## Year             1.009 16           1.000
```



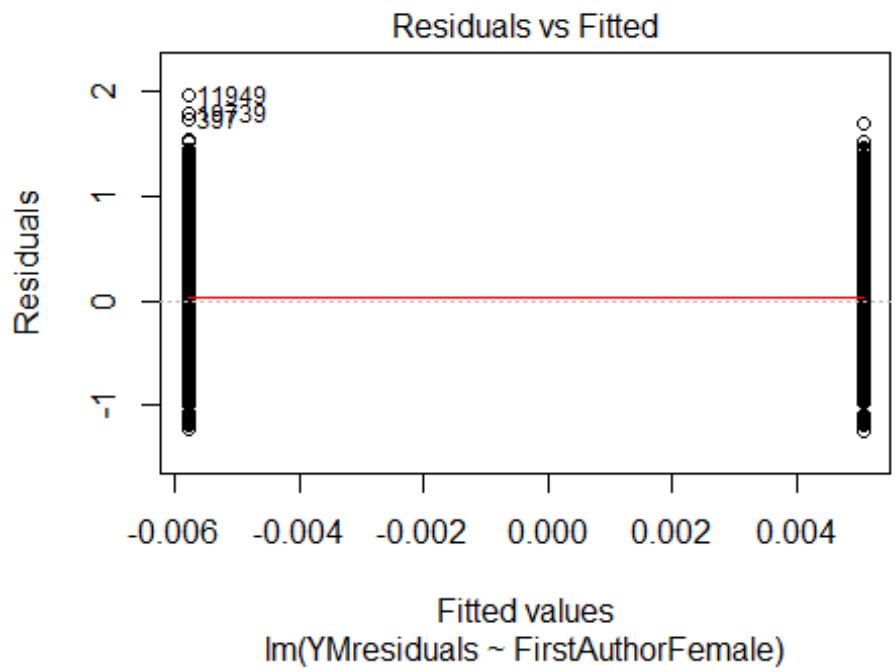

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.29298 -0.26253  0.00856  0.26135  2.36363
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.23920    0.02023   61.27  < 2e-16 ***
## LastAuthorFemale1 -0.06849    0.00847   -8.08  6.9e-16 ***
## Year1997            0.05378    0.02756    1.95  0.05107 .
## Year1998           -0.05950    0.03135   -1.90  0.05775 .
## Year1999           -0.09921    0.02522   -3.93  8.4e-05 ***
## Year2000           -0.02385    0.03114   -0.77  0.44388
## Year2001           -0.07472    0.02721   -2.75  0.00605 **
## Year2002           -0.06581    0.02619   -2.51  0.01200 *
## Year2003           -0.09361    0.02615   -3.58  0.00034 ***
## Year2004           -0.10250    0.02563   -4.00  6.4e-05 ***
## Year2005           -0.12657    0.02502   -5.06  4.3e-07 ***
## Year2006           -0.07255    0.02455   -2.96  0.00313 **
```

```
## Year2007          -0.09084    0.02437   -3.73  0.00020 ***
## Year2008          -0.06291    0.02526   -2.49  0.01278 *
## Year2009          -0.08710    0.02575   -3.38  0.00072 ***
## Year2010          -0.06867    0.02452   -2.80  0.00511 **
## Year2011          -0.08367    0.02400   -3.49  0.00049 ***
## Year2012          -0.10827    0.02476   -4.37  1.2e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.393
## Multiple R-squared:  0.0157, Adjusted R-squared:  0.0142
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  2 observations c(6244,8507) are outliers with |weight| = 0 ( < 8.9e-06);
##  986 weights are ~= 1. The remaining 10229 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0046  0.8640  0.9500  0.8950  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          8.92e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Sample size for the above analysis:  11217"
## [1] ""
## [1] ""
## [1] "####################################"
## [1] "Analysis of AJSC 2717"
## [1] "####################################"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  960  858  868  807 1048 1175 1058  849  969 1158 1222 1274 1396 1431 1456
## 2011 2012
## 1515 1593
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  771  642  647  580  653  678  908  709  802  957 1032 1046 1158 1194 1159
```

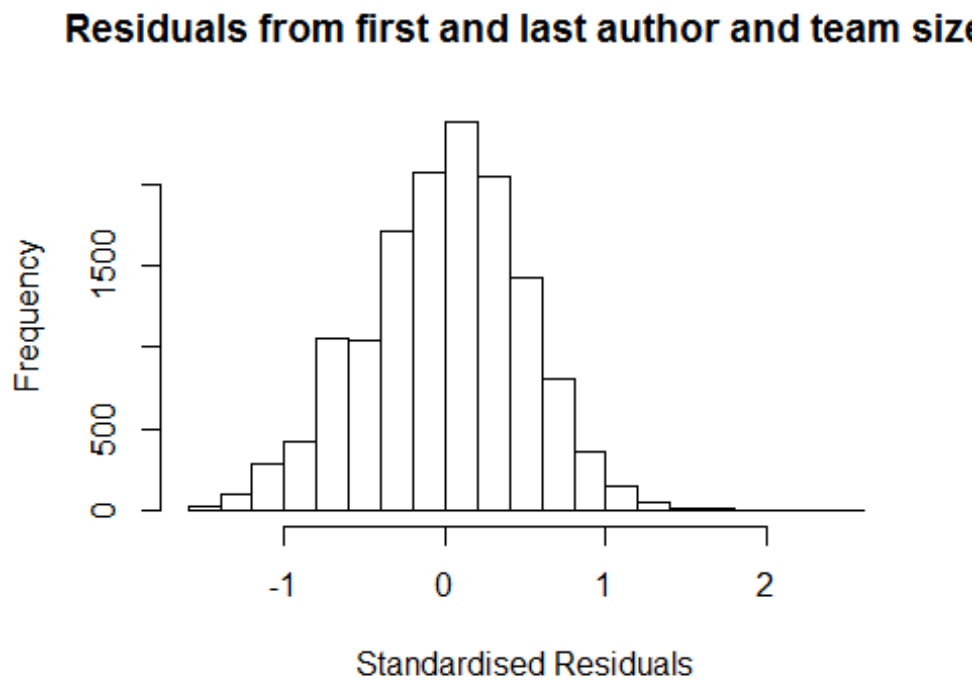
```
## 2011 2012
## 1197 1279
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  713  594  602  527  595  620  811  629  713  875  942  955 1038 1077 1022
## 2011 2012
## 1079 1166
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
##  Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 71, df = 16, p-value = 7e-09
```




```
##
##  Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 9.1, df = 1, p-value = 0.003
```

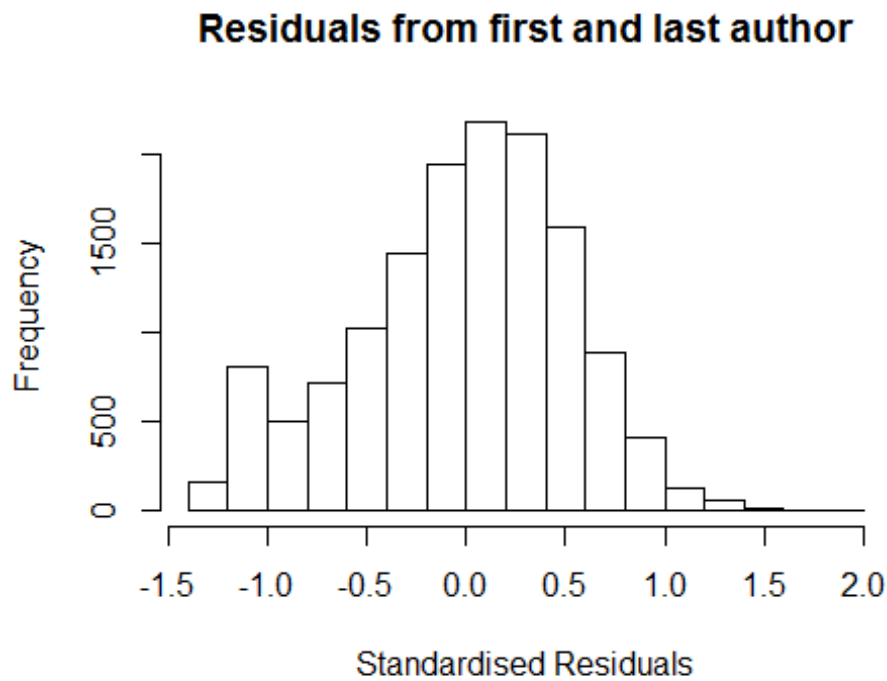



```
## [1] "Female first author team size 2018 geometric mean: 4.53472773201965"
## [1] "Male first author team size 2018 geometric mean: 4.64785803674514"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 120000, p-value = 0.4
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 4.29409478408711"
## [1] "Male last author team size 2018 geometric mean: 4.86544981999559"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 120000, p-value = 0.006
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.032  1           1.016
## LastAuthorFemale  1.072  1           1.035
## UniqueAuthors     1.132  4           1.016
## Year              1.094 16           1.003
```



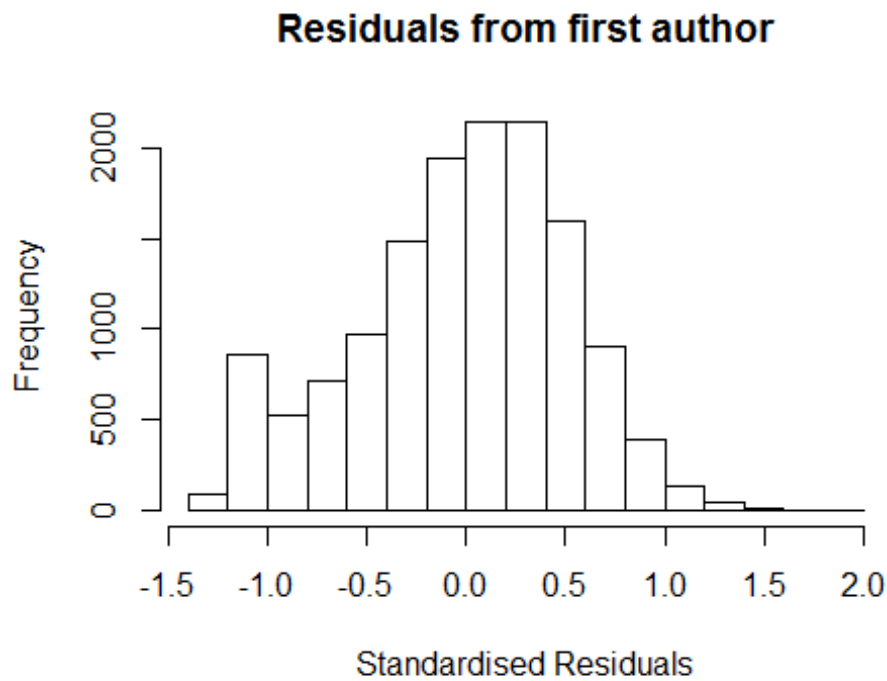

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
UniqueAuthors +
##     Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##     k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.5118 -0.3318  0.0223  0.3256  2.4043
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         0.81934    0.02565   31.95  < 2e-16 ***
## FirstAuthorFemale1  0.01532    0.00866    1.77   0.0768 .
## LastAuthorFemale1  -0.00998    0.00889   -1.12   0.2617
## UniqueAuthors2      0.33325    0.01706   19.54  < 2e-16 ***
## UniqueAuthors3      0.44279    0.01687   26.25  < 2e-16 ***
## UniqueAuthors4      0.54085    0.01708   31.66  < 2e-16 ***
## UniqueAuthors5      0.66288    0.01461   45.37  < 2e-16 ***
## Year1997           -0.02169    0.03226   -0.67   0.5014
## Year1998           -0.03025    0.03120   -0.97   0.3322
## Year1999            0.02962    0.03130    0.95   0.3440
```

```
## Year2000           -0.02659    0.03156   -0.84   0.3995
## Year2001           -0.05393    0.03097   -1.74   0.0816 .
## Year2002           -0.06309    0.02841   -2.22   0.0264 *
## Year2003           -0.10165    0.03120   -3.26   0.0011 **
## Year2004           -0.00777    0.02982   -0.26   0.7945
## Year2005           -0.10231    0.02830   -3.62   0.0003 ***
## Year2006           -0.13463    0.02842   -4.74  2.2e-06 ***
## Year2007           -0.14692    0.02742   -5.36  8.6e-08 ***
## Year2008           -0.19137    0.02713   -7.05  1.8e-12 ***
## Year2009           -0.12079    0.02732   -4.42  9.9e-06 ***
## Year2010           -0.16615    0.02714   -6.12  9.5e-10 ***
## Year2011           -0.20573    0.02710   -7.59  3.4e-14 ***
## Year2012           -0.19983    0.02681   -7.45  9.6e-14 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.483
## Multiple R-squared:  0.187,  Adjusted R-squared:  0.186
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
##  observation 7442 is an outlier with |weight| = 0 ( < 7.2e-06);
##  1145 weights are ~= 1. The remaining 12812 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0056  0.8610  0.9500  0.9050  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          7.16e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                    psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.014  1           1.007
## LastAuthorFemale  1.016  1           1.008
## Year              1.009 16           1.000
```



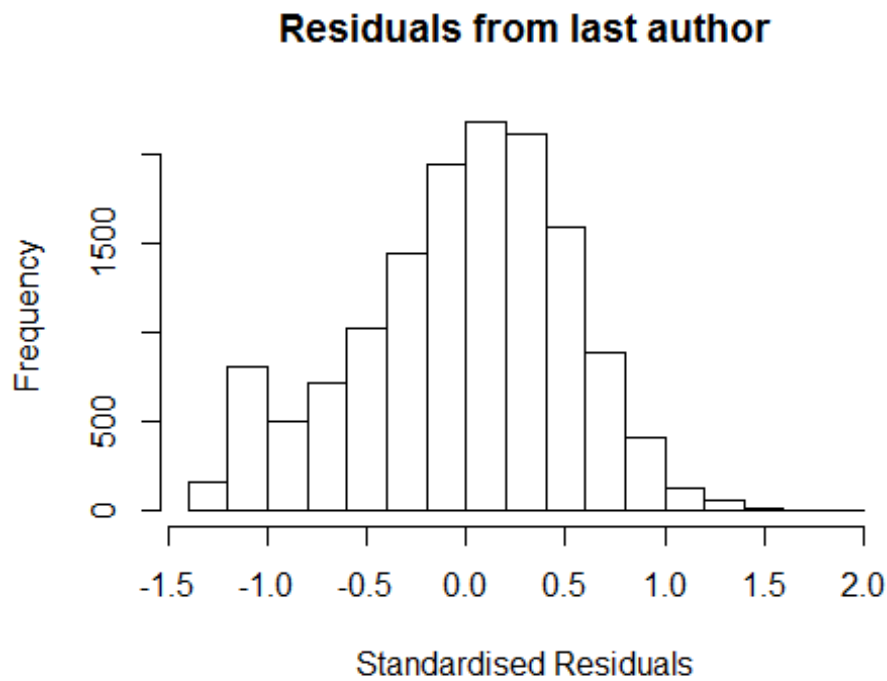

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.2996 -0.3471  0.0378  0.3552  1.9258
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)          1.20439    0.02469   48.78  < 2e-16 ***
## FirstAuthorFemale1  0.02273    0.00950    2.39   0.0168 *
## LastAuthorFemale1  -0.06921    0.00953   -7.26  3.9e-13 ***
## Year1997           -0.00718    0.03459   -0.21   0.8355
## Year1998           -0.01856    0.03304   -0.56   0.5744
## Year1999            0.07219    0.03302    2.19   0.0288 *
## Year2000            0.03524    0.03397    1.04   0.2996
## Year2001           -0.02190    0.03348   -0.65   0.5130
## Year2002            0.00314    0.03024    0.10   0.9173
## Year2003           -0.00605    0.03352   -0.18   0.8568
## Year2004            0.07252    0.03204    2.26   0.0236 *
## Year2005           -0.00630    0.03111   -0.20   0.8396
```

```
## Year2006           -0.04121    0.03103   -1.33   0.1842
## Year2007           -0.05891    0.02964   -1.99   0.0469 *
## Year2008           -0.08727    0.02953   -2.95   0.0031 **
## Year2009           -0.01136    0.02925   -0.39   0.6978
## Year2010           -0.05104    0.02901   -1.76   0.0786 .
## Year2011           -0.07304    0.02879   -2.54   0.0112 *
## Year2012           -0.07438    0.02829   -2.63   0.0086 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.512
## Multiple R-squared:  0.0103, Adjusted R-squared:  0.00905
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
##  1074 weights are ~= 1. The remaining 12884 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.127   0.861   0.949   0.899   0.985   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          7.16e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.004  1           1.002
## Year              1.004 16           1.000
```




```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.2617 -0.3513  0.0389  0.3594  1.9452
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)          1.18000    0.02444   48.28   <2e-16 ***
## FirstAuthorFemale1   0.00657    0.00955    0.69   0.4915
## Year1997            -0.00405    0.03459   -0.12   0.9068
## Year1998            -0.01692    0.03314   -0.51   0.6096
## Year1999             0.07512    0.03310    2.27   0.0233 *
## Year2000             0.03868    0.03405    1.14   0.2560
## Year2001            -0.02095    0.03362   -0.62   0.5331
## Year2002             0.00284    0.03037    0.09   0.9254
## Year2003            -0.00560    0.03349   -0.17   0.8673
## Year2004             0.07382    0.03206    2.30   0.0213 *
## Year2005            -0.00211    0.03119   -0.07   0.9461
## Year2006            -0.03625    0.03107   -1.17   0.2434
```

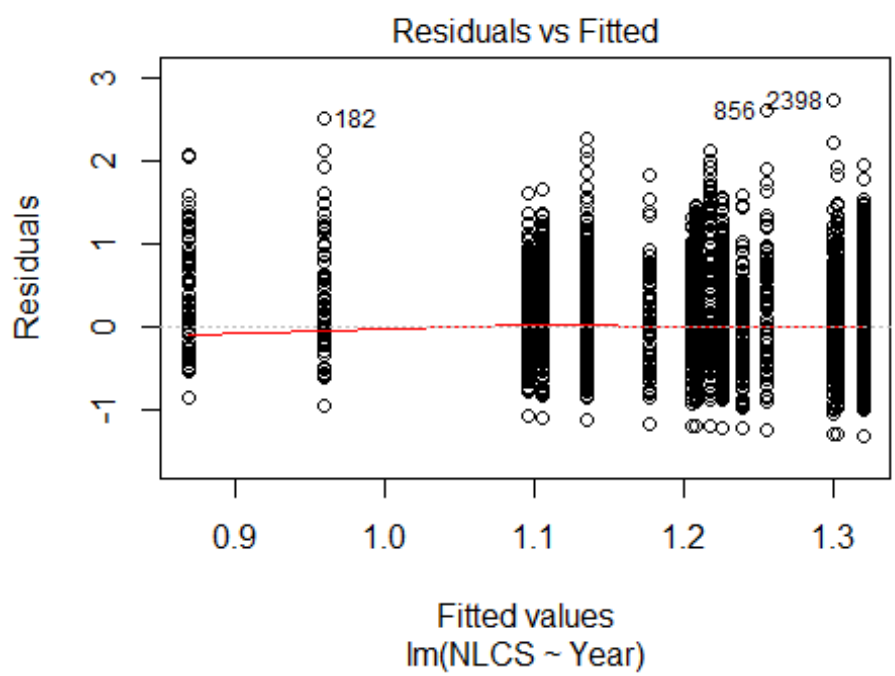
```
## Year2007           -0.05569    0.02971   -1.87   0.0609 .
## Year2008           -0.08450    0.02960   -2.85   0.0043 **
## Year2009           -0.00853    0.02928   -0.29   0.7709
## Year2010           -0.04878    0.02915   -1.67   0.0943 .
## Year2011           -0.06845    0.02889   -2.37   0.0178 *
## Year2012           -0.07137    0.02835   -2.52   0.0118 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.514
## Multiple R-squared:  0.00658,    Adjusted R-squared:  0.00536
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
##  1118 weights are ~= 1. The remaining 12840 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.120   0.860   0.948   0.899   0.984   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          7.16e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##                   GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.005  1           1.003
## Year             1.005 16           1.000
```



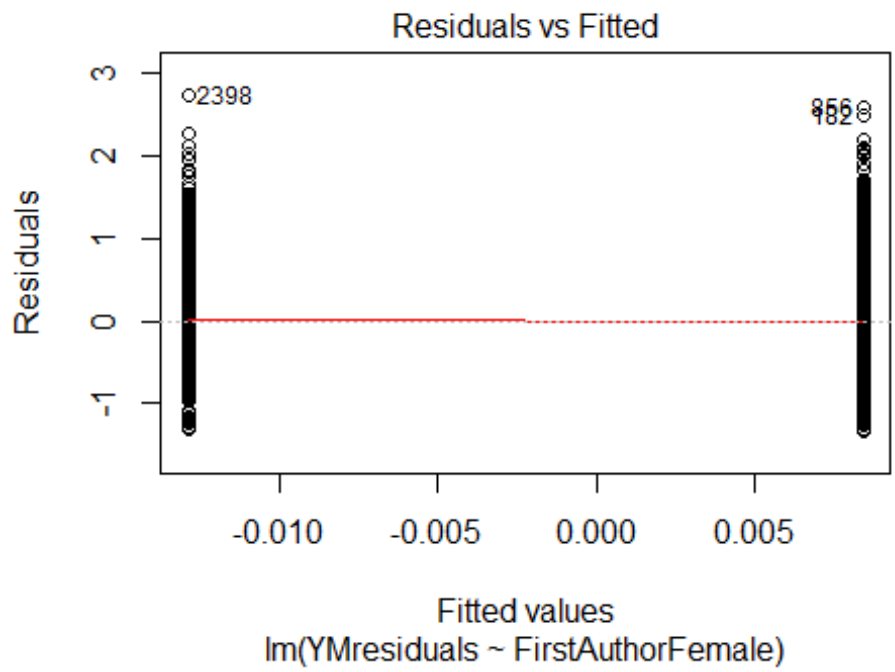

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.2862 -0.3473  0.0385  0.3550  1.9162
##
## Coefficients:
##                   Estimate Std. Error t value Pr(>|t|)
## (Intercept)        1.21316    0.02432   49.87  < 2e-16 ***
## LastAuthorFemale1 -0.06384    0.00956   -6.68  2.5e-11 ***
## Year1997          -0.00709    0.03457   -0.21   0.8375
## Year1998          -0.01842    0.03303   -0.56   0.5770
## Year1999           0.07255    0.03301    2.20   0.0280 *
## Year2000           0.03630    0.03395    1.07   0.2849
## Year2001          -0.02073    0.03343   -0.62   0.5352
## Year2002           0.00355    0.03025    0.12   0.9067
## Year2003          -0.00505    0.03351   -0.15   0.8803
## Year2004           0.07304    0.03200    2.28   0.0225 *
## Year2005          -0.00539    0.03111   -0.17   0.8624
## Year2006          -0.04036    0.03104   -1.30   0.1936
```

```
## Year2007          -0.05772    0.02962   -1.95   0.0514 .
## Year2008          -0.08528    0.02950   -2.89   0.0039 **
## Year2009          -0.00943    0.02922   -0.32   0.7469
## Year2010          -0.04865    0.02897   -1.68   0.0931 .
## Year2011          -0.07087    0.02876   -2.46   0.0138 *
## Year2012          -0.07283    0.02827   -2.58   0.0100 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.512
## Multiple R-squared:  0.01,   Adjusted R-squared:  0.00881
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
##  1086 weights are ~= 1. The remaining 12872 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.131   0.860   0.949   0.899   0.985   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          7.16e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Sample size for the above analysis:  13958"
## [1] ""
## [1] ""
## [1] "####################################"
## [1] "Analysis of AJSC 2718"
## [1] "####################################"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  185  321  166  180  280  184  215  213  268  272  423  492  567  571  615
## 2011 2012
##  793  787
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  123  228  120  128  160   86  141  160  204  194  288  358  450  437  469
## 2011 2012
```

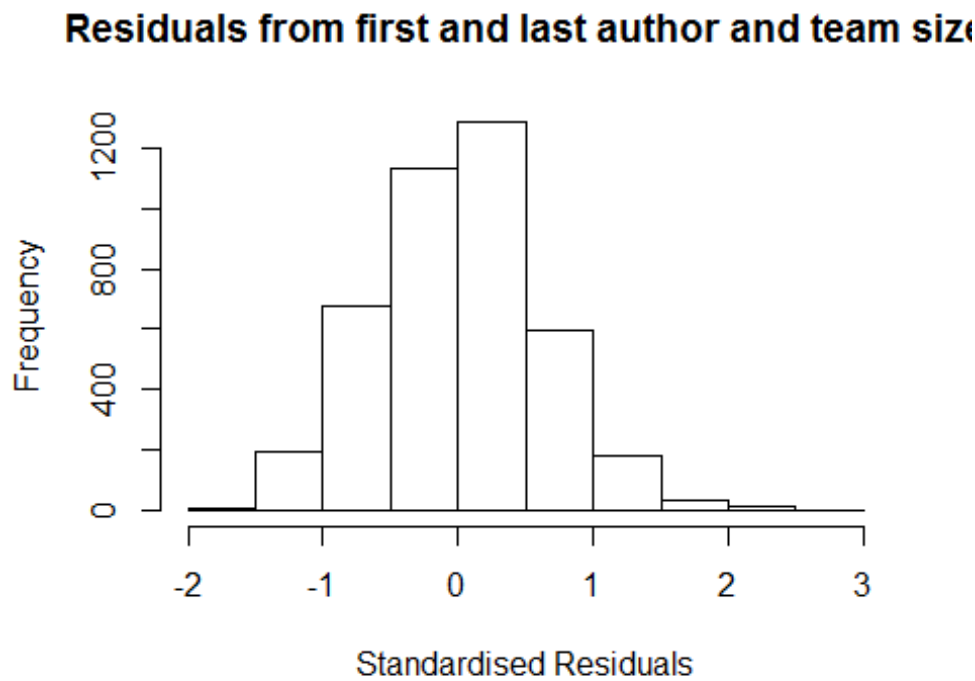
```
##  610  605
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  115  199  105  116  140   76  123  134  170  162  251  319  394  382  400
## 2011 2012
##  523  508
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
##  Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 180, df = 16, p-value <2e-16
```




```
##
##  Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 0.86, df = 1, p-value = 0.4
```

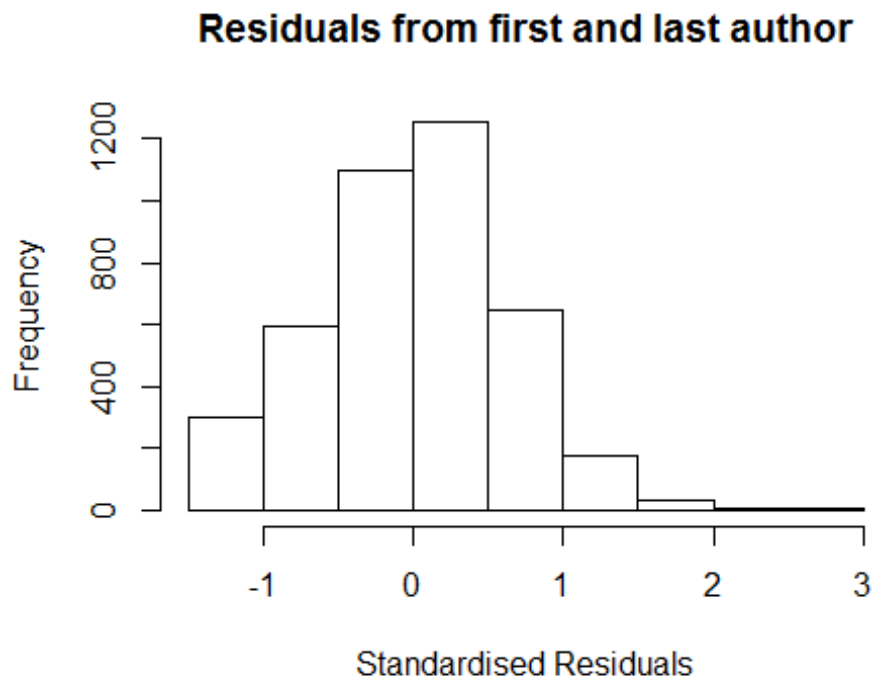



```
## [1] "Female first author team size 2018 geometric mean: 4.60157253597935"
## [1] "Male first author team size 2018 geometric mean: 4.74461904789526"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 94000, p-value = 1
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 4.27911773116693"
## [1] "Male last author team size 2018 geometric mean: 4.96719516051439"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 81000, p-value = 0.004
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.166  1           1.080
## LastAuthorFemale  1.193  1           1.092
## UniqueAuthors     1.215  4           1.025
## Year              1.221 16           1.006
```



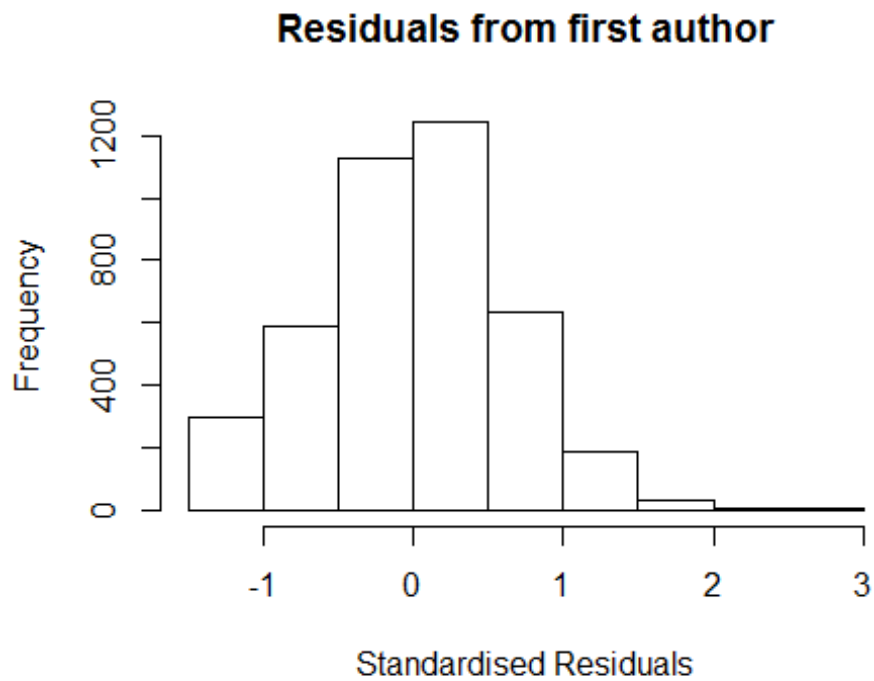

```
## [1] "List of  1 outliers with residuals above  2.5"
##         ScopusId  NLCS Year OneField Fields residuals
## 2398 21544449115 4.018 2005     2700      2     2.757
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
UniqueAuthors +
##     Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##     k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.5099 -0.3970  0.0187  0.3945  2.7567
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)           0.7158     0.0699   10.23  < 2e-16 ***
## FirstAuthorFemale1    0.0195     0.0211    0.92  0.35581
## LastAuthorFemale1    -0.0562     0.0217   -2.59  0.00966 **
## UniqueAuthors2        0.3132     0.0351    8.93  < 2e-16 ***
## UniqueAuthors3        0.2850     0.0352    8.11  6.7e-16 ***
## UniqueAuthors4        0.4085     0.0359   11.38  < 2e-16 ***
## UniqueAuthors5        0.4984     0.0312   15.98  < 2e-16 ***
## Year1997              0.1649     0.0943    1.75  0.08035 .
## Year1998              0.1766     0.0979    1.80  0.07142 .
## Year1999              0.2582     0.1053    2.45  0.01425 *
```

```
## Year2000            -0.1702     0.1007   -1.69  0.09104 .
## Year2001             0.2038     0.1131    1.80  0.07164 .
## Year2002             0.2084     0.0964    2.16  0.03072 *
## Year2003             0.2149     0.0821    2.62  0.00887 **
## Year2004             0.1584     0.0795    1.99  0.04641 *
## Year2005             0.2409     0.0832    2.89  0.00381 **
## Year2006             0.2957     0.0804    3.68  0.00024 ***
## Year2007             0.1871     0.0750    2.50  0.01261 *
## Year2008             0.0711     0.0736    0.97  0.33391
## Year2009             0.0583     0.0737    0.79  0.42855
## Year2010             0.0692     0.0725    0.96  0.33939
## Year2011             0.2577     0.0743    3.47  0.00052 ***
## Year2012             0.1755     0.0735    2.39  0.01701 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.596
## Multiple R-squared:  0.105,  Adjusted R-squared:   0.1
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
##  358 weights are ~= 1. The remaining 3759 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0006  0.8620  0.9510  0.9030  0.9860  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          2.43e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.153  1           1.074
## LastAuthorFemale  1.162  1           1.078
## Year              1.034 16           1.001
```



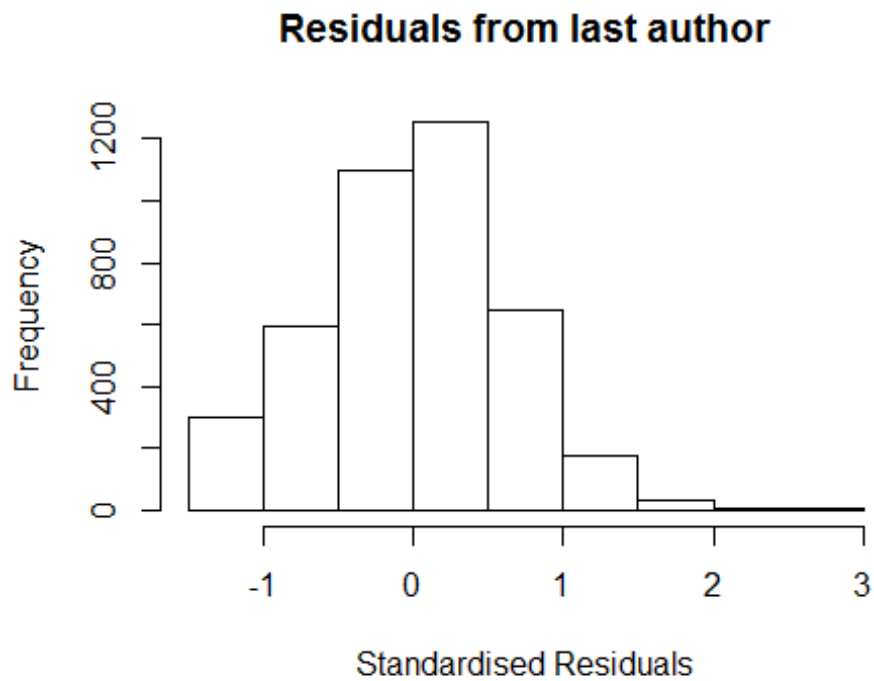

```
## [1] "List of  3 outliers with residuals above  2.5"
##         ScopusId  NLCS Year OneField Fields residuals
## 182   0030088712 3.461 1996     1704      5     2.533
## 856   0032813705 3.853 1999     2718      1     2.615
## 2398 21544449115 4.018 2005     2700      2     2.696
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.3576 -0.4138  0.0259  0.4176  2.6959
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)          0.9278     0.0777   11.94  < 2e-16 ***
## FirstAuthorFemale1   0.0117     0.0220    0.53  0.59442
## LastAuthorFemale1   -0.1077     0.0225   -4.78  1.8e-06 ***
## Year1997             0.2643     0.1003    2.63  0.00847 **
## Year1998             0.1984     0.1091    1.82  0.06911 .
## Year1999             0.3105     0.1140    2.72  0.00648 **
## Year2000            -0.0893     0.1040   -0.86  0.39036
## Year2001             0.2729     0.1183    2.31  0.02114 *
## Year2002             0.3309     0.1029    3.21  0.00132 **
## Year2003             0.3472     0.0902    3.85  0.00012 ***
```

```
## Year2004             0.2960     0.0872    3.40  0.00069 ***
## Year2005             0.3825     0.0904    4.23  2.4e-05 ***
## Year2006             0.4298     0.0888    4.84  1.4e-06 ***
## Year2007             0.3051     0.0840    3.63  0.00028 ***
## Year2008             0.2073     0.0823    2.52  0.01181 *
## Year2009             0.2293     0.0817    2.81  0.00501 **
## Year2010             0.2117     0.0811    2.61  0.00911 **
## Year2011             0.4163     0.0828    5.03  5.2e-07 ***
## Year2012             0.3586     0.0819    4.38  1.2e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.613
## Multiple R-squared:  0.0376, Adjusted R-squared:  0.0334
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
##  353 weights are ~= 1. The remaining 3764 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0141  0.8640  0.9500  0.9020  0.9860  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          2.43e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.014  1           1.007
## Year              1.014 16           1.000
```




```
## [1] "List of  3 outliers with residuals above  2.5"
##         ScopusId  NLCS Year OneField Fields residuals
## 182   0030088712 3.461 1996     1704      5     2.533
## 856   0032813705 3.853 1999     2718      1     2.615
## 2398 21544449115 4.018 2005     2700      2     2.696
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##    Min     1Q Median     3Q    Max
## -1.338 -0.415  0.023  0.422  2.756
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)          0.8976     0.0789   11.38  < 2e-16 ***
## FirstAuthorFemale1  -0.0319     0.0208   -1.53  0.12597
## Year1997             0.2691     0.1012    2.66  0.00785 **
## Year1998             0.2037     0.1112    1.83  0.06704 .
## Year1999             0.3252     0.1156    2.81  0.00492 **
## Year2000            -0.0767     0.1056   -0.73  0.46748
## Year2001             0.2790     0.1198    2.33  0.01990 *
## Year2002             0.3359     0.1048    3.20  0.00137 **
## Year2003             0.3603     0.0913    3.95  8.1e-05 ***
## Year2004             0.3068     0.0885    3.47  0.00053 ***
```

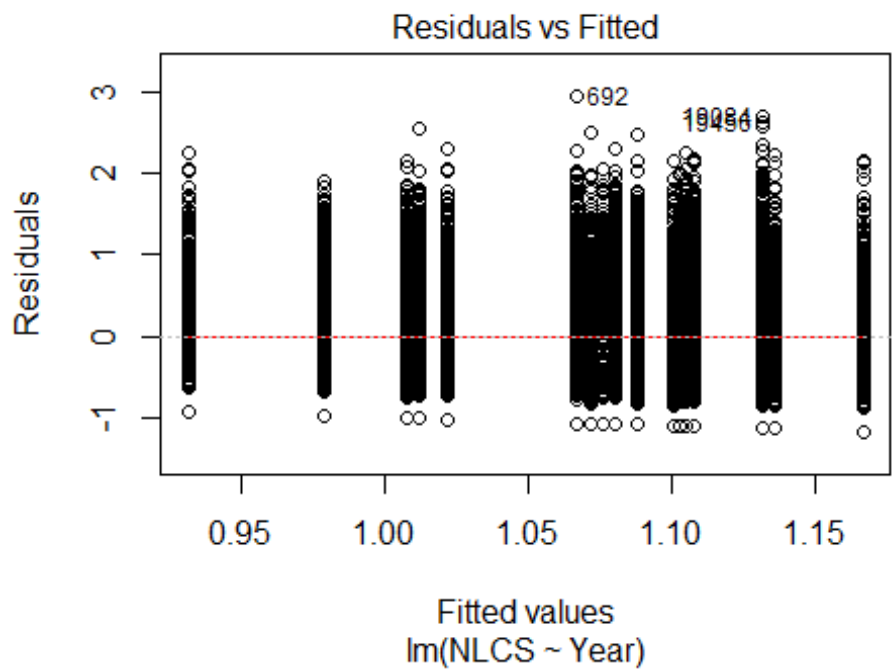
```
## Year2005             0.3965     0.0915    4.33  1.5e-05 ***
## Year2006             0.4409     0.0904    4.88  1.1e-06 ***
## Year2007             0.3181     0.0855    3.72  0.00020 ***
## Year2008             0.2174     0.0839    2.59  0.00958 **
## Year2009             0.2436     0.0832    2.93  0.00344 **
## Year2010             0.2224     0.0827    2.69  0.00720 **
## Year2011             0.4214     0.0844    4.99  6.2e-07 ***
## Year2012             0.3655     0.0836    4.37  1.3e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.615
## Multiple R-squared:  0.0323, Adjusted R-squared:  0.0283
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
##  346 weights are ~= 1. The remaining 3771 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0073  0.8600  0.9490  0.9020  0.9860  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          2.43e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##                   GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.021  1           1.011
## Year             1.021 16           1.001
```



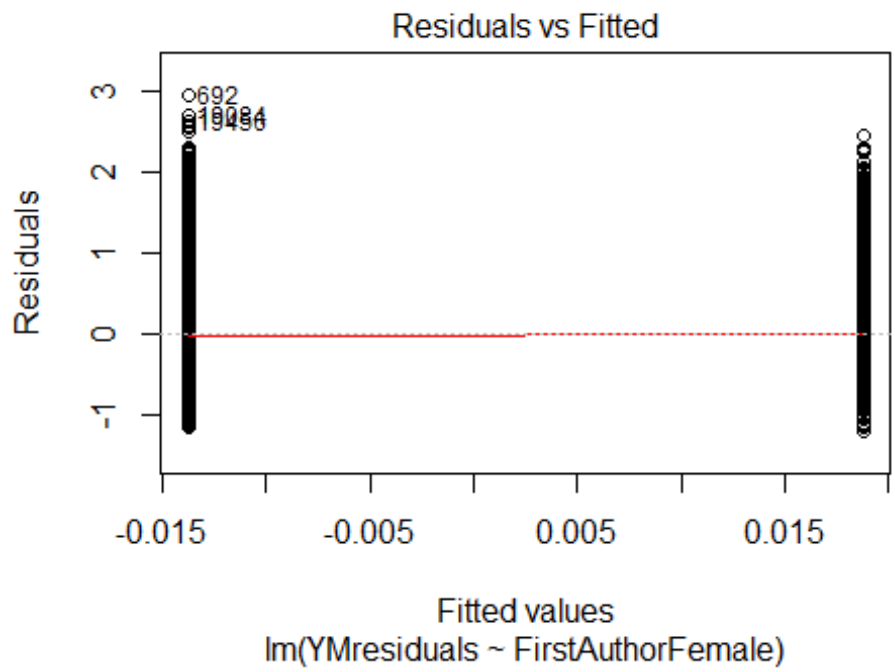

```
## [1] "List of  3 outliers with residuals above  2.5"
##         ScopusId  NLCS Year OneField Fields residuals
## 182   0030088712 3.461 1996     1704      5     2.533
## 856   0032813705 3.853 1999     2718      1     2.615
## 2398 21544449115 4.018 2005     2700      2     2.696
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.3599 -0.4134  0.0246  0.4153  2.7046
##
## Coefficients:
##                   Estimate Std. Error t value Pr(>|t|)
## (Intercept)         0.9311     0.0776   12.00  < 2e-16 ***
## LastAuthorFemale1  -0.1027     0.0212   -4.85  1.3e-06 ***
## Year1997            0.2628     0.1003    2.62  0.00881 **
## Year1998            0.1985     0.1091    1.82  0.06895 .
## Year1999            0.3092     0.1139    2.71  0.00669 **
## Year2000           -0.0899     0.1041   -0.86  0.38782
## Year2001            0.2724     0.1184    2.30  0.02146 *
## Year2002            0.3298     0.1030    3.20  0.00138 **
## Year2003            0.3469     0.0903    3.84  0.00012 ***
## Year2004            0.2958     0.0872    3.39  0.00070 ***
```

```
## Year2005            0.3823     0.0905    4.23  2.4e-05 ***
## Year2006            0.4288     0.0889    4.82  1.5e-06 ***
## Year2007            0.3048     0.0840    3.63  0.00029 ***
## Year2008            0.2073     0.0823    2.52  0.01186 *
## Year2009            0.2294     0.0817    2.81  0.00500 **
## Year2010            0.2113     0.0812    2.60  0.00928 **
## Year2011            0.4163     0.0829    5.02  5.3e-07 ***
## Year2012            0.3588     0.0820    4.38  1.2e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.613
## Multiple R-squared:  0.0375, Adjusted R-squared:  0.0336
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
##  359 weights are ~= 1. The remaining 3758 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0129  0.8640  0.9490  0.9020  0.9860  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          2.43e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                  psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Sample size for the above analysis:  4117"
## [1] ""
## [1] ""
## [1] "##################################"
## [1] "Analysis of AJSC 2719"
## [1] "##################################"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  897  825  868  926 1150 1375 1273 1195 1237 1453 1494 1640 1984 2254 2495
## 2011 2012
## 2706 2981
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
```

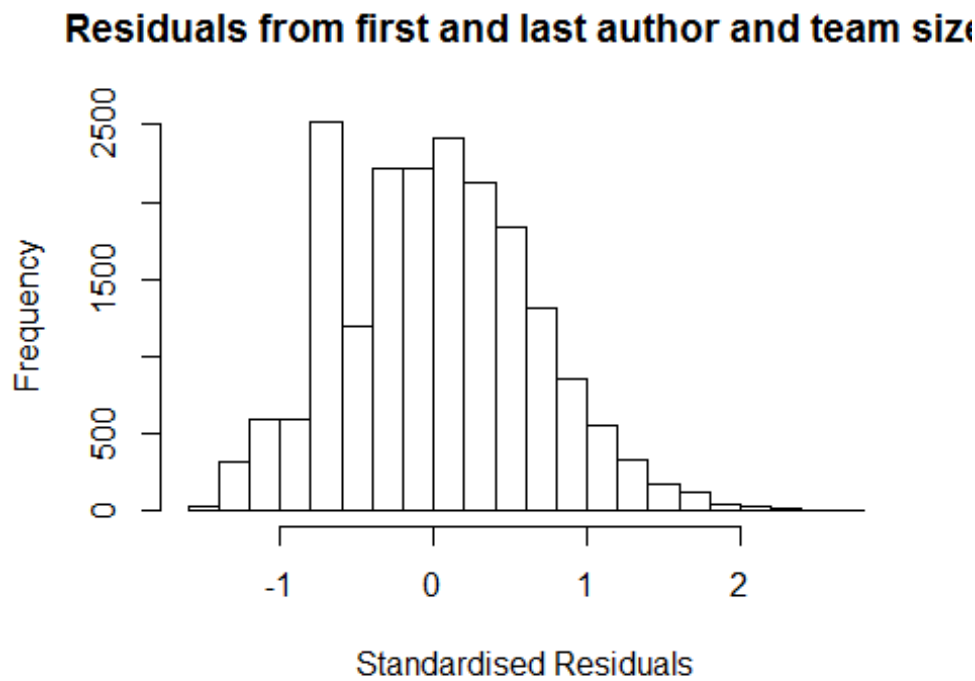
```
##  580  513  514  576  595  685 1089 1039 1028 1229 1287 1424 1682 1942 2092
## 2011 2012
## 2301 2536
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  554  478  474  535  544  648 1014  960  946 1129 1153 1298 1552 1768 1920
## 2011 2012
## 2110 2327
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
##  Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 100, df = 16, p-value = 1e-14
```




```
##
##  Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 23, df = 1, p-value = 2e-06
```




```
## [1] "Female first author team size 2018 geometric mean: 3.75066287070192"
## [1] "Male first author team size 2018 geometric mean: 3.0591067489645"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 1e+06, p-value = 5e-12
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 3.47981179803226"
## [1] "Male last author team size 2018 geometric mean: 3.34777868949035"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 920000, p-value = 0.3
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##                     GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.166  1           1.080
## LastAuthorFemale  1.150  1           1.072
## UniqueAuthors     1.057  4           1.007
## Year              1.046 16           1.001
```

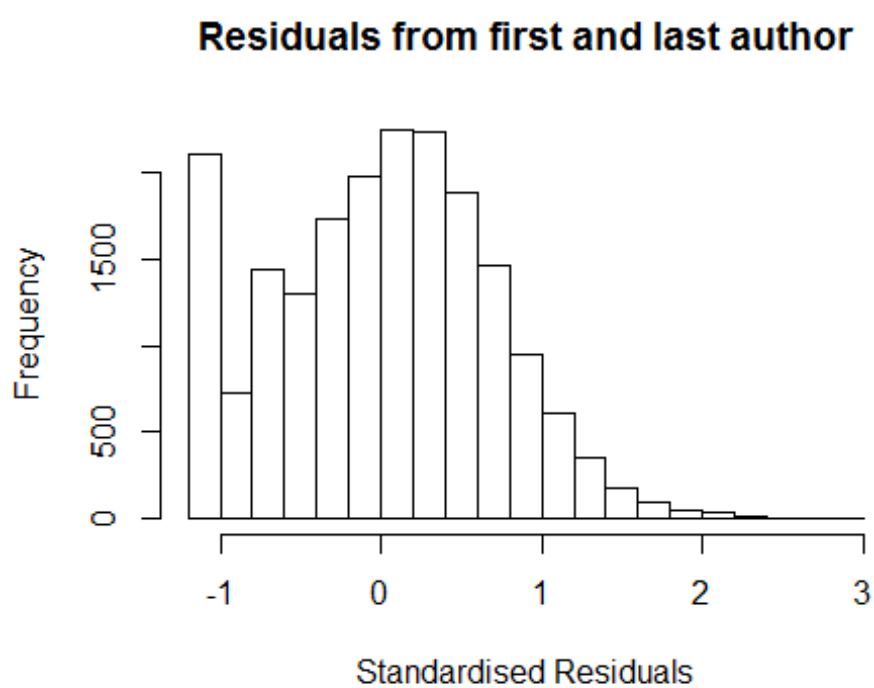



```
## [1] "List of  3 outliers with residuals above  2.5"
##          ScopusId  NLCS Year OneField Fields residuals
## 692    0029680562 4.006 1996     2719      2     2.733
## 2641   0032102084 3.179 1998     2719      2     2.531
## 19084 70349207410 3.825 2009     2719      1     2.509
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
UniqueAuthors +
##     Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##     k.max = 1000))
##  \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.50625 -0.45409  0.00363  0.43479  2.73327
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         0.76150    0.03450   22.07  < 2e-16 ***
## FirstAuthorFemale1 -0.00569    0.01013   -0.56   0.5744
## LastAuthorFemale1  -0.02911    0.01025   -2.84   0.0045 **
## UniqueAuthors2      0.37659    0.01515   24.86  < 2e-16 ***
## UniqueAuthors3      0.51123    0.01505   33.96  < 2e-16 ***
## UniqueAuthors4      0.57424    0.01537   37.35  < 2e-16 ***
## UniqueAuthors5      0.66562    0.01285   51.78  < 2e-16 ***
## Year1997            0.01139    0.04910    0.23   0.8166
```

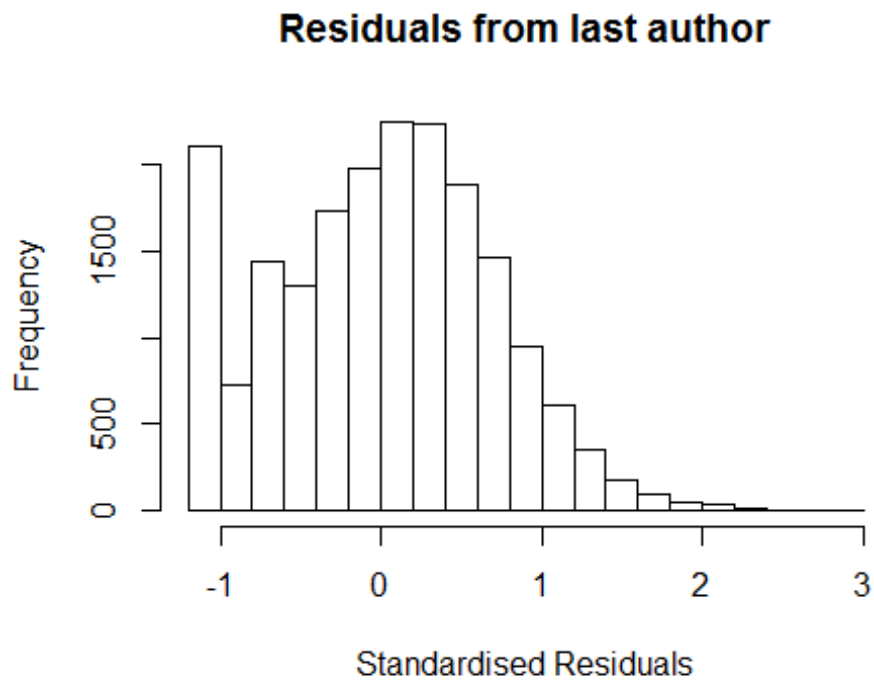
```
## Year1998           -0.11302    0.04708   -2.40   0.0164 *
## Year1999            0.06750    0.04701    1.44   0.1510
## Year2000            0.07914    0.04411    1.79   0.0728 .
## Year2001            0.06749    0.04391    1.54   0.1243
## Year2002           -0.09909    0.04010   -2.47   0.0135 *
## Year2003           -0.09071    0.03986   -2.28   0.0229 *
## Year2004           -0.10613    0.03956   -2.68   0.0073 **
## Year2005           -0.15472    0.03907   -3.96  7.5e-05 ***
## Year2006           -0.09813    0.03820   -2.57   0.0102 *
## Year2007           -0.06706    0.03784   -1.77   0.0763 .
## Year2008           -0.07548    0.03749   -2.01   0.0441 *
## Year2009           -0.01926    0.03725   -0.52   0.6052
## Year2010           -0.03566    0.03680   -0.97   0.3325
## Year2011           -0.05195    0.03647   -1.42   0.1543
## Year2012           -0.08487    0.03639   -2.33   0.0197 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.626
## Multiple R-squared:  0.156,  Adjusted R-squared:  0.155
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
##  1550 weights are ~= 1. The remaining 17860 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0176  0.8800  0.9450  0.9070  0.9840  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          5.15e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.115  1           1.056
## LastAuthorFemale  1.109  1           1.053
## Year              1.010 16           1.000
```

```
## [1] "List of  6 outliers with residuals above  2.5"
##          ScopusId  NLCS Year OneField Fields residuals
## 692    0029680562 4.006 1996     2719      2     2.990
## 9096   0037374498 3.555 2003     2719      2     2.570
## 13421 33947510199 3.556 2006     2719      5     2.515
## 19084 70349207410 3.825 2009     2719      1     2.727
## 19216 67449130309 3.706 2009     2719      5     2.647
## 19456 67649482684 3.748 2009     2719      5     2.650
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##    Min     1Q Median     3Q    Max
## -1.199 -0.491  0.040  0.475  2.990
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)          1.0165     0.0349   29.13  < 2e-16 ***
## FirstAuthorFemale1   0.0570     0.0108    5.28  1.3e-07 ***
## LastAuthorFemale1   -0.0398     0.0110   -3.61  0.00031 ***
## Year1997             0.0388     0.0515    0.75  0.45131
## Year1998            -0.1269     0.0494   -2.57  0.01020 *
## Year1999             0.0668     0.0483    1.38  0.16684
## Year2000             0.1252     0.0458    2.74  0.00623 **
## Year2001             0.1048     0.0450    2.33  0.01980 *
## Year2002            -0.0475     0.0416   -1.14  0.25272
## Year2003            -0.0319     0.0418   -0.76  0.44482
## Year2004            -0.0250     0.0420   -0.60  0.55052
## Year2005            -0.0632     0.0413   -1.53  0.12593
## Year2006             0.0250     0.0395    0.63  0.52739
## Year2007             0.0686     0.0390    1.76  0.07911 .
## Year2008             0.0542     0.0388    1.40  0.16246
## Year2009             0.0819     0.0387    2.11  0.03452 *
## Year2010             0.0606     0.0383    1.58  0.11355
## Year2011             0.0576     0.0378    1.52  0.12779
## Year2012             0.0335     0.0379    0.88  0.37785
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.709
## Multiple R-squared:  0.00737,    Adjusted R-squared:  0.00645
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  1554 weights are ~= 1. The remaining 17856 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0362  0.8590  0.9490  0.9150  0.9850  0.9990
```

```
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          5.15e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"

## Warning in lmrob.S(x, y, control = control, mf = mf): S refinements did
not
## converge (to refine.tol=1e-07) in 1000 (= k.max) steps
```



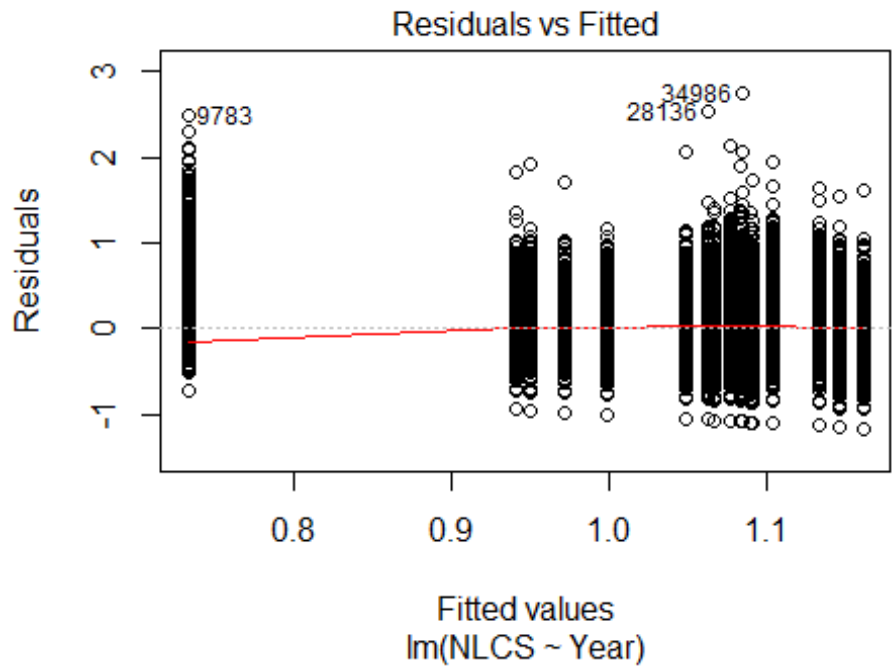

```
## [1] "Regression 4: Last author gender, Year as factors"
##                   GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.003  1           1.002
## Year             1.003 16           1.000
```



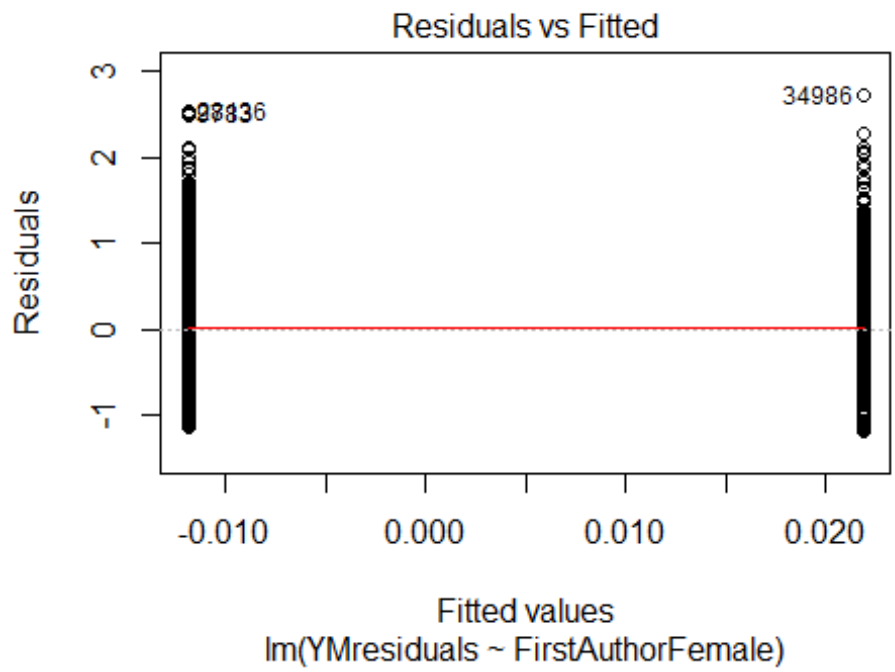

```
## [1] "List of  6 outliers with residuals above  2.5"
##          ScopusId  NLCS Year OneField Fields residuals
## 692    0029680562 4.006 1996     2719      2     2.990
## 9096   0037374498 3.555 2003     2719      2     2.570
## 13421 33947510199 3.556 2006     2719      5     2.515
## 19084 70349207410 3.825 2009     2719      1     2.727
## 19216 67449130309 3.706 2009     2719      5     2.647
## 19456 67649482684 3.748 2009     2719      5     2.650
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.1551 -0.4917  0.0383  0.4770  2.9792
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)          1.0268     0.0348   29.53   <2e-16 ***
## LastAuthorFemale1   -0.0170     0.0105   -1.62    0.104
## Year1997             0.0386     0.0513    0.75    0.452
## Year1998            -0.1239     0.0494   -2.51    0.012 *
## Year1999             0.0702     0.0483    1.46    0.146
## Year2000             0.1283     0.0457    2.81    0.005 **
## Year2001             0.1084     0.0449    2.41    0.016 *
```

```
## Year2002           -0.0429     0.0415   -1.03    0.301
## Year2003           -0.0278     0.0418   -0.67    0.505
## Year2004           -0.0208     0.0419   -0.50    0.619
## Year2005           -0.0566     0.0412   -1.37    0.170
## Year2006            0.0318     0.0395    0.81    0.420
## Year2007            0.0751     0.0390    1.93    0.054 .
## Year2008            0.0601     0.0387    1.55    0.121
## Year2009            0.0883     0.0387    2.28    0.022 *
## Year2010            0.0671     0.0382    1.76    0.079 .
## Year2011            0.0637     0.0377    1.69    0.091 .
## Year2012            0.0402     0.0379    1.06    0.288
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.71
## Multiple R-squared:  0.00602,    Adjusted R-squared:  0.00515
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  1563 weights are ~= 1. The remaining 17847 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0388  0.8580  0.9490  0.9150  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          5.15e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Sample size for the above analysis:  19410"
## [1] ""
## [1] ""
## [1] "####################################"
## [1] "Analysis of AJSC 2720"
## [1] "####################################"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 1744 1658 1519 1227 2798 1854 1797 1133 1207 1249 1467 1850 1947 1994 1982
## 2011 2012
```

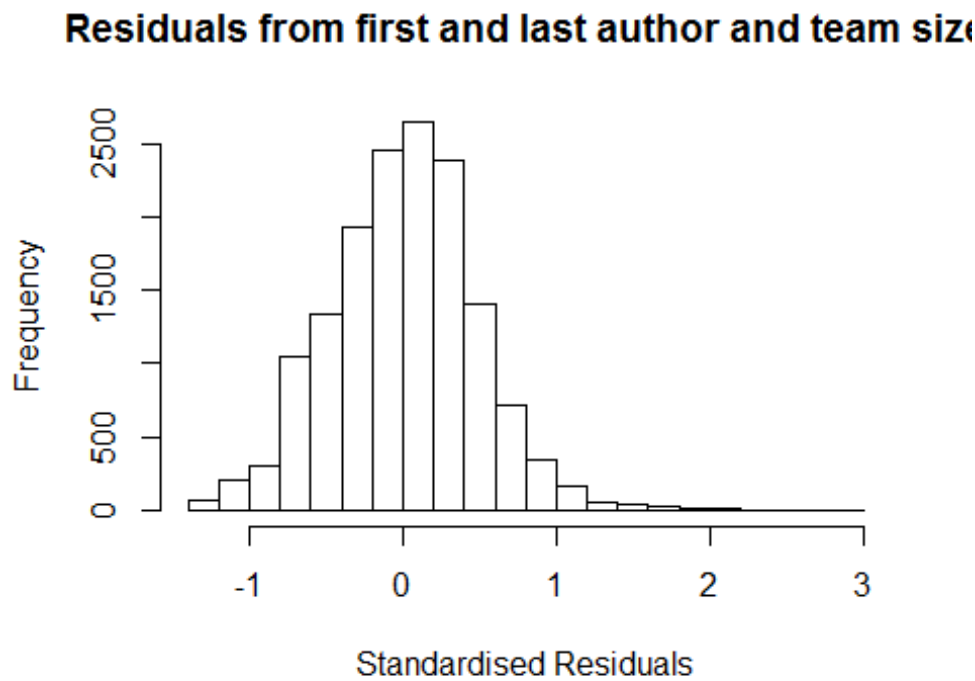
```
## 2193 2101
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  654  702  617  595 1340  946 1138  693  743  851  968 1247 1346 1380 1372
## 2011 2012
## 1498 1434
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  571  616  549  515 1169  829  998  602  604  704  829 1079 1134 1169 1197
## 2011 2012
## 1324 1252
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
##  Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 930, df = 16, p-value <2e-16
```




```
##
##  Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 38, df = 1, p-value = 6e-10
```



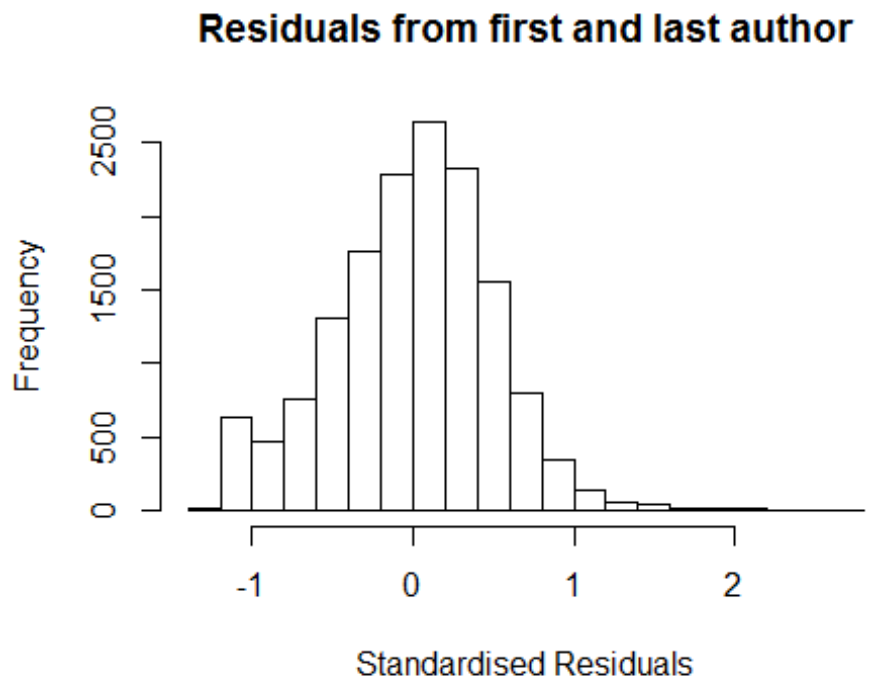

```
## [1] "Female first author team size 2018 geometric mean: 5.62710159162691"
## [1] "Male first author team size 2018 geometric mean: 5.77983300241576"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 180000, p-value = 0.4
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 5.15408292718835"
## [1] "Male last author team size 2018 geometric mean: 5.97075977946513"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 140000, p-value = 0.005
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.033  1           1.016
## LastAuthorFemale  1.023  1           1.011
## UniqueAuthors     1.110  4           1.013
## Year              1.119 16           1.004
```




```
## [1] "List of  3 outliers with residuals above  2.5"
##          ScopusId  NLCS Year OneField Fields residuals
## 9783   0033959520 3.226 2000     2720      2     2.652
## 28136 77955635233 3.588 2010     2720      2     2.599
## 34986 84855792427 3.831 2012     2720      2     2.837
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
UniqueAuthors +
##     Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##     k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.3593 -0.3026  0.0175  0.3074  2.8374
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         0.80625    0.02776   29.04  < 2e-16 ***
## FirstAuthorFemale1  0.01080    0.00801    1.35   0.1775
## LastAuthorFemale1  -0.02530    0.00938   -2.70   0.0070 **
## UniqueAuthors2      0.30152    0.01952   15.45  < 2e-16 ***
## UniqueAuthors3      0.33045    0.01855   17.81  < 2e-16 ***
## UniqueAuthors4      0.37548    0.01814   20.70  < 2e-16 ***
## UniqueAuthors5      0.54228    0.01575   34.43  < 2e-16 ***
## Year1997           -0.03442    0.03428   -1.00   0.3153
```

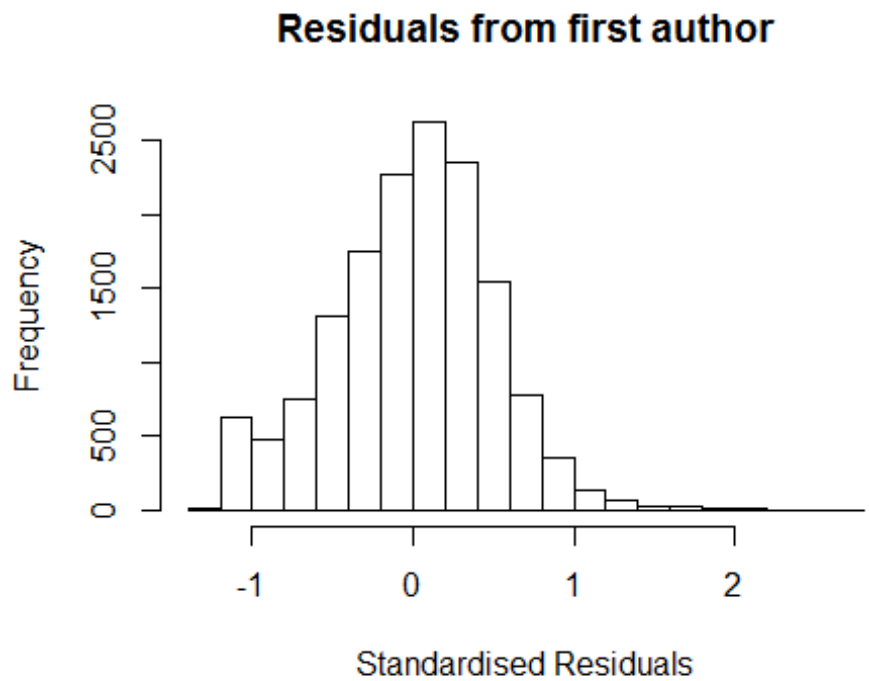
```
## Year1998           -0.08700    0.03239   -2.69   0.0072 **
## Year1999           -0.09555    0.02948   -3.24   0.0012 **
## Year2000           -0.58223    0.03955  -14.72  < 2e-16 ***
## Year2001           -0.03092    0.02720   -1.14   0.2558
## Year2002           -0.08049    0.02684   -3.00   0.0027 **
## Year2003           -0.23464    0.02871   -8.17  3.2e-16 ***
## Year2004           -0.25931    0.02879   -9.01  < 2e-16 ***
## Year2005           -0.22471    0.02880   -7.80  6.5e-15 ***
## Year2006           -0.18026    0.02784   -6.47  9.8e-11 ***
## Year2007           -0.16233    0.02691   -6.03  1.7e-09 ***
## Year2008           -0.14055    0.02688   -5.23  1.7e-07 ***
## Year2009           -0.13480    0.02680   -5.03  5.0e-07 ***
## Year2010           -0.16704    0.02762   -6.05  1.5e-09 ***
## Year2011           -0.13574    0.02665   -5.09  3.6e-07 ***
## Year2012           -0.15391    0.02728   -5.64  1.7e-08 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.435
## Multiple R-squared:  0.194,  Adjusted R-squared:  0.193
## Convergence in 18 IRWLS iterations
##
## Robustness weights:
##  13 observations
c(2284,2297,2913,2980,3029,3030,3060,3162,3167,3256,11699,13130,15092)
##   are outliers with |weight| = 0 ( < 6.6e-06);
##  1196 weights are ~= 1. The remaining 13932 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0001  0.8560  0.9470  0.8950  0.9840  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          6.60e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.016  1           1.008
```

```
## LastAuthorFemale  1.013  1           1.006
## Year              1.020 16           1.001
```



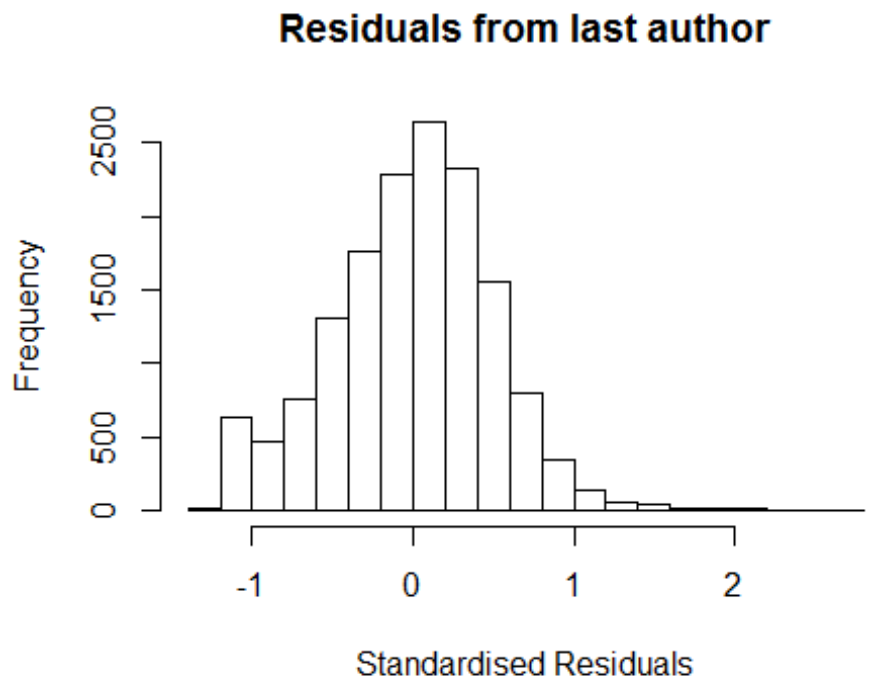

```
## [1] "List of  4 outliers with residuals above  2.5"
##          ScopusId  NLCS Year OneField Fields residuals
## 9033   0034254249 3.216 2000     2720      1     2.657
## 9783   0033959520 3.226 2000     2720      2     2.667
## 28136 77955635233 3.588 2010     2720      2     2.546
## 34986 84855792427 3.831 2012     2720      2     2.696
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.2168 -0.3240  0.0275  0.3242  2.6962
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)          1.17021    0.02555   45.79  < 2e-16 ***
## FirstAuthorFemale1   0.04661    0.00842    5.54  3.1e-08 ***
## LastAuthorFemale1   -0.03204    0.00988   -3.24  0.00119 **
## Year1997            -0.04270    0.03714   -1.15  0.25028
## Year1998            -0.06445    0.03404   -1.89  0.05831 .
## Year1999            -0.08232    0.03073   -2.68  0.00739 **
```

```
## Year2000           -0.57887    0.04040  -14.33  < 2e-16 ***
## Year2001           -0.00311    0.02878   -0.11  0.91400
## Year2002           -0.02413    0.02841   -0.85  0.39562
## Year2003           -0.20532    0.03053   -6.73  1.8e-11 ***
## Year2004           -0.24482    0.03154   -7.76  8.8e-15 ***
## Year2005           -0.23266    0.03088   -7.53  5.2e-14 ***
## Year2006           -0.17938    0.02983   -6.01  1.9e-09 ***
## Year2007           -0.11605    0.02918   -3.98  7.0e-05 ***
## Year2008           -0.10476    0.02887   -3.63  0.00029 ***
## Year2009           -0.07496    0.02884   -2.60  0.00934 **
## Year2010           -0.09618    0.02962   -3.25  0.00117 **
## Year2011           -0.07578    0.02849   -2.66  0.00781 **
## Year2012           -0.08201    0.02886   -2.84  0.00450 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.448
## Multiple R-squared:  0.0774, Adjusted R-squared:  0.0763
## Convergence in 18 IRWLS iterations
##
## Robustness weights:
##  8 observations c(2913,2980,3029,3030,3167,3256,11699,15092)
##   are outliers with |weight| = 0 ( < 6.6e-06);
##  1227 weights are ~= 1. The remaining 13906 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0009  0.8480  0.9430  0.8890  0.9840  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          6.60e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.011  1           1.006
## Year              1.011 16           1.000
```




```
## [1] "List of  4 outliers with residuals above  2.5"
##          ScopusId  NLCS Year OneField Fields residuals
## 9033   0034254249 3.216 2000     2720      1     2.657
## 9783   0033959520 3.226 2000     2720      2     2.667
## 28136 77955635233 3.588 2010     2720      2     2.546
## 34986 84855792427 3.831 2012     2720      2     2.696
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.2079 -0.3256  0.0281  0.3259  2.7067
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.16484    0.02549   45.70  < 2e-16 ***
## FirstAuthorFemale1  0.04307    0.00842    5.11  3.2e-07 ***
## Year1997           -0.04156    0.03721   -1.12  0.26406
## Year1998           -0.06414    0.03405   -1.88  0.05961 .
## Year1999           -0.08261    0.03074   -2.69  0.00722 **
## Year2000           -0.58159    0.04040  -14.40  < 2e-16 ***
## Year2001           -0.00313    0.02880   -0.11  0.91359
## Year2002           -0.02448    0.02843   -0.86  0.38923
## Year2003           -0.20521    0.03052   -6.72  1.8e-11 ***
```

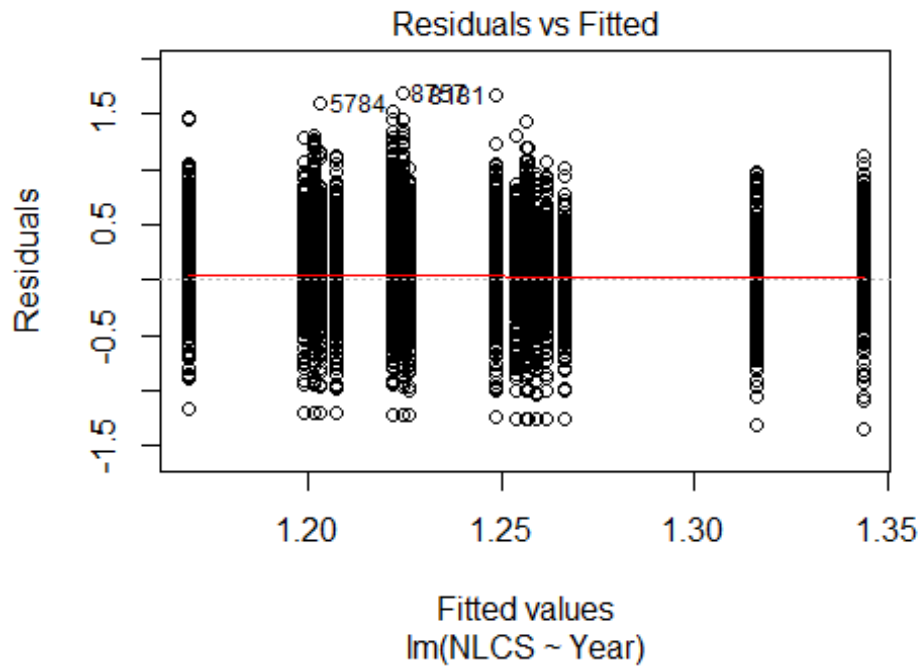
```
## Year2004           -0.24478    0.03154   -7.76  9.0e-15 ***
## Year2005           -0.23405    0.03091   -7.57  3.9e-14 ***
## Year2006           -0.17912    0.02987   -6.00  2.1e-09 ***
## Year2007           -0.11597    0.02918   -3.97  7.1e-05 ***
## Year2008           -0.10522    0.02890   -3.64  0.00027 ***
## Year2009           -0.07539    0.02886   -2.61  0.00901 **
## Year2010           -0.09735    0.02965   -3.28  0.00103 **
## Year2011           -0.07718    0.02851   -2.71  0.00680 **
## Year2012           -0.08361    0.02889   -2.89  0.00381 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.448
## Multiple R-squared:  0.0769, Adjusted R-squared:  0.0758
## Convergence in 18 IRWLS iterations
##
## Robustness weights:
##  8 observations c(2913,2980,3029,3030,3167,3256,11699,15092)
##   are outliers with |weight| = 0 ( < 6.6e-06);
##  1197 weights are ~= 1. The remaining 13936 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0006  0.8520  0.9430  0.8890  0.9840  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          6.60e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##                   GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.008  1           1.004
## Year             1.008 16           1.000
```



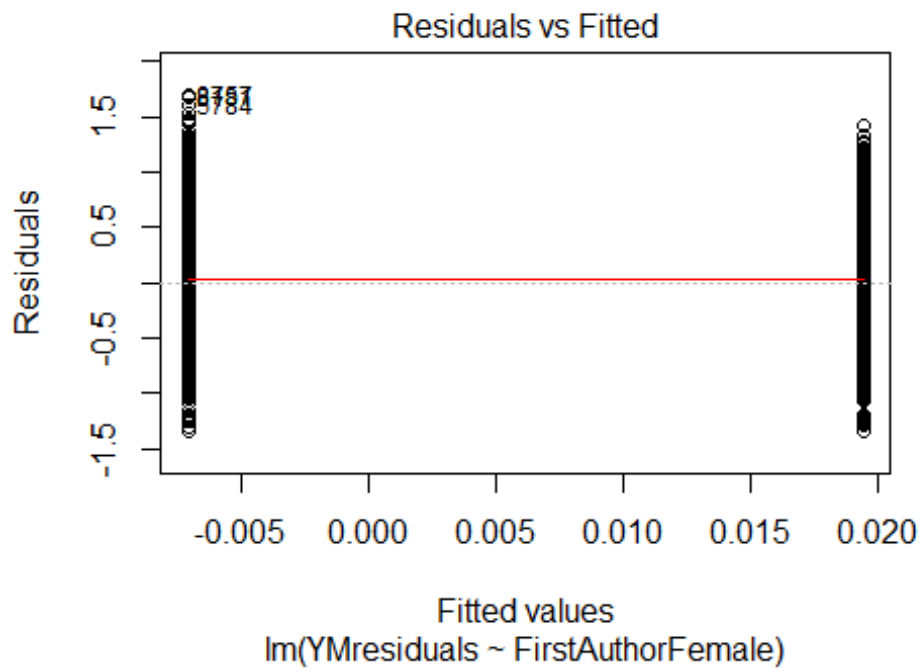

```
## [1] "List of  4 outliers with residuals above  2.5"
##          ScopusId  NLCS Year OneField Fields residuals
## 9033   0034254249 3.216 2000     2720      1     2.657
## 9783   0033959520 3.226 2000     2720      2     2.667
## 28136 77955635233 3.588 2010     2720      2     2.546
## 34986 84855792427 3.831 2012     2720      2     2.696
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##    Min     1Q Median     3Q    Max
## -1.183 -0.326  0.029  0.323  2.727
##
## Coefficients:
##                   Estimate Std. Error t value Pr(>|t|)
## (Intercept)        1.18307    0.02540   46.58  < 2e-16 ***
## LastAuthorFemale1 -0.02517    0.00986   -2.55  0.01070 *
## Year1997          -0.04146    0.03715   -1.12  0.26439
## Year1998          -0.06510    0.03407   -1.91  0.05602 .
## Year1999          -0.08109    0.03076   -2.64  0.00840 **
## Year2000          -0.57751    0.04045  -14.28  < 2e-16 ***
## Year2001          -0.00142    0.02881   -0.05  0.96063
## Year2002          -0.02258    0.02845   -0.79  0.42745
## Year2003          -0.20301    0.03052   -6.65  3.0e-11 ***
```

```
## Year2004          -0.24418    0.03152   -7.75  1.0e-14 ***
## Year2005          -0.23112    0.03094   -7.47  8.4e-14 ***
## Year2006          -0.17675    0.02985   -5.92  3.3e-09 ***
## Year2007          -0.11279    0.02922   -3.86  0.00011 ***
## Year2008          -0.10028    0.02890   -3.47  0.00052 ***
## Year2009          -0.07040    0.02886   -2.44  0.01472 *
## Year2010          -0.09159    0.02960   -3.09  0.00198 **
## Year2011          -0.07111    0.02850   -2.49  0.01261 *
## Year2012          -0.07861    0.02888   -2.72  0.00650 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.448
## Multiple R-squared:  0.0756, Adjusted R-squared:  0.0746
## Convergence in 18 IRWLS iterations
##
## Robustness weights:
##  8 observations c(2913,2980,3029,3030,3167,3256,11699,15092)
##   are outliers with |weight| = 0 ( < 6.6e-06);
##  1190 weights are ~= 1. The remaining 13943 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0001  0.8450  0.9440  0.8890  0.9840  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          6.60e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Sample size for the above analysis:  15141"
## [1] ""
## [1] ""
## [1] "####################################"
## [1] "Analysis of AJSC 2721"
## [1] "####################################"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  510  483  553  582  535  677  565  466  393  454  527  607  695  891  770
## 2011 2012
```

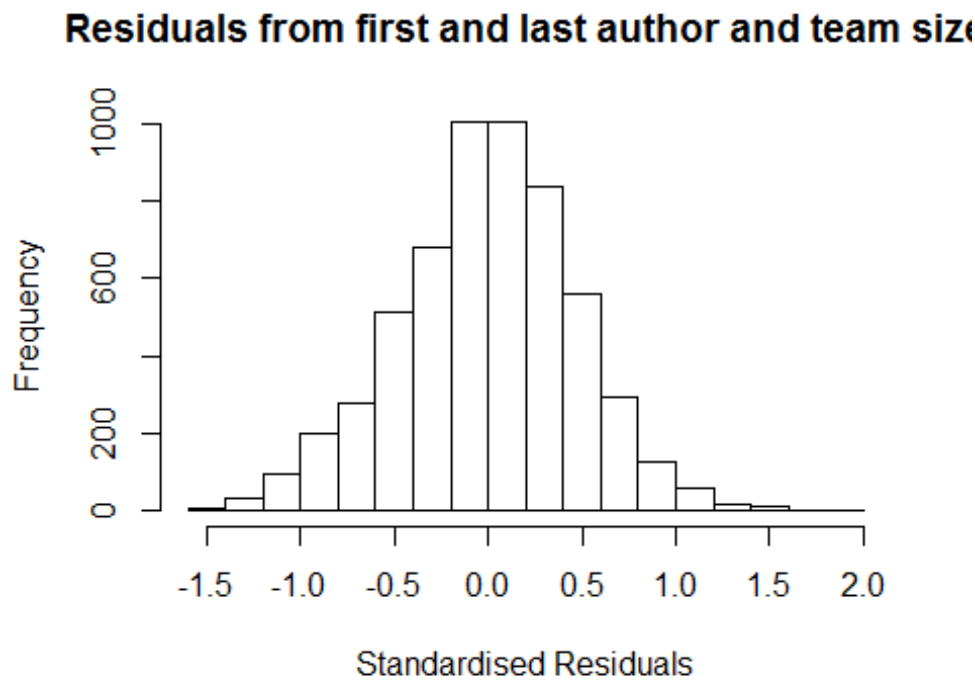
```
##  843  832
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  232  238  272  297  324  371  369  338  257  314  364  422  491  633  530
## 2011 2012
##  611  572
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  198  204  244  258  283  326  320  302  221  258  316  354  414  529  462
## 2011 2012
##  510  501
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
##  Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 150, df = 16, p-value <2e-16
```




```
##
##  Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 0.43, df = 1, p-value = 0.5
```

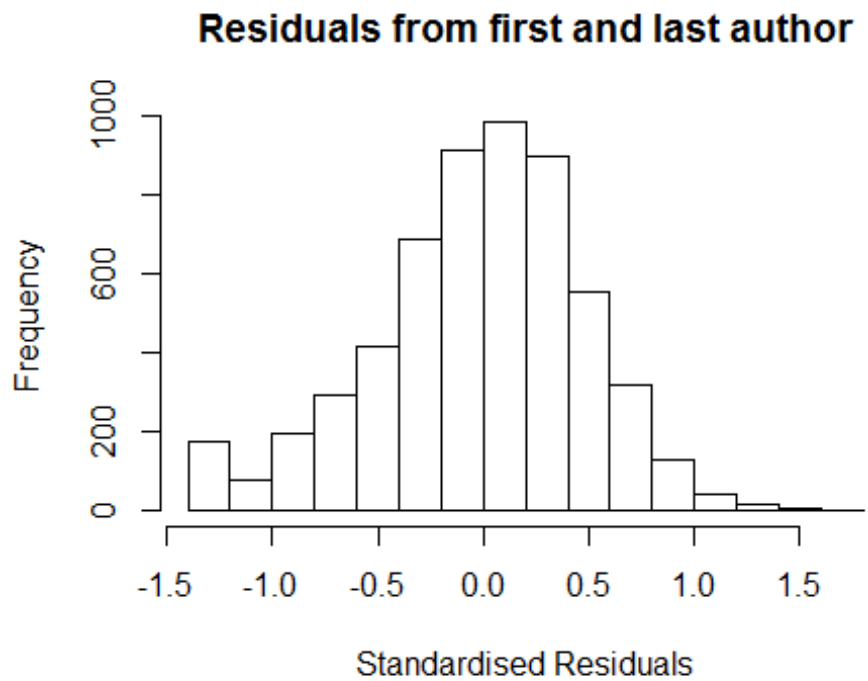



```
## [1] "Female first author team size 2018 geometric mean: 5.19140803004158"
## [1] "Male first author team size 2018 geometric mean: 5.31662227306746"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 26000, p-value = 0.7
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 4.4709439713507"
## [1] "Male last author team size 2018 geometric mean: 5.51077923223257"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 17000, p-value = 0.02
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.037  1           1.019
## LastAuthorFemale  1.019  1           1.010
## UniqueAuthors     1.154  4           1.018
## Year              1.170 16           1.005
```



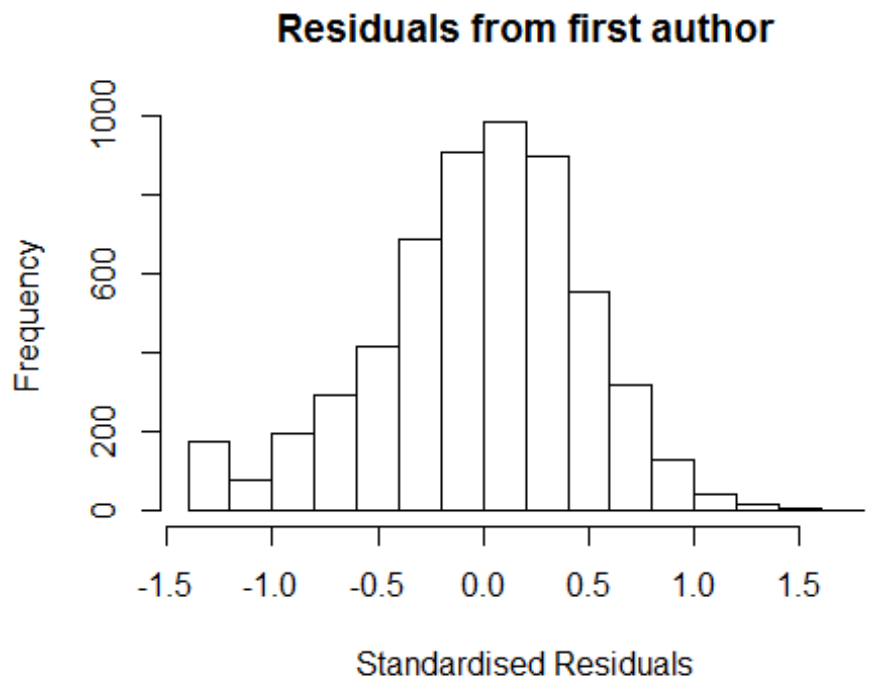

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
UniqueAuthors +
##     Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##     k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.4490 -0.2902  0.0086  0.3042  1.8227
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)         0.886430   0.047171   18.79  < 2e-16 ***
## FirstAuthorFemale1  0.000207   0.014197    0.01   0.9884
## LastAuthorFemale1   0.005708   0.017271    0.33   0.7410
## UniqueAuthors2      0.147588   0.037223    3.96  7.4e-05 ***
## UniqueAuthors3      0.182550   0.036136    5.05  4.5e-07 ***
## UniqueAuthors4      0.363874   0.034164   10.65  < 2e-16 ***
## UniqueAuthors5      0.515638   0.030856   16.71  < 2e-16 ***
## Year1997            0.011358   0.045053    0.25   0.8010
## Year1998            0.133296   0.048303    2.76   0.0058 **
## Year1999            0.041064   0.044934    0.91   0.3608
```

```
## Year2000            0.070566   0.044195    1.60   0.1104
## Year2001            0.017120   0.043342    0.40   0.6929
## Year2002           -0.007197   0.045629   -0.16   0.8747
## Year2003            0.023323   0.042918    0.54   0.5868
## Year2004            0.006478   0.047995    0.13   0.8926
## Year2005           -0.073602   0.045666   -1.61   0.1071
## Year2006           -0.021095   0.044502   -0.47   0.6355
## Year2007           -0.009277   0.043995   -0.21   0.8330
## Year2008           -0.022118   0.042045   -0.53   0.5989
## Year2009           -0.028370   0.044149   -0.64   0.5205
## Year2010           -0.070949   0.043459   -1.63   0.1026
## Year2011           -0.027755   0.043266   -0.64   0.5212
## Year2012           -0.078510   0.043138   -1.82   0.0688 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.446
## Multiple R-squared:  0.145,  Adjusted R-squared:  0.142
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
##  468 weights are ~= 1. The remaining 5232 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0572  0.8640  0.9500  0.9010  0.9860  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.75e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.026  1           1.013
## LastAuthorFemale  1.019  1           1.009
## Year              1.043 16           1.001
```



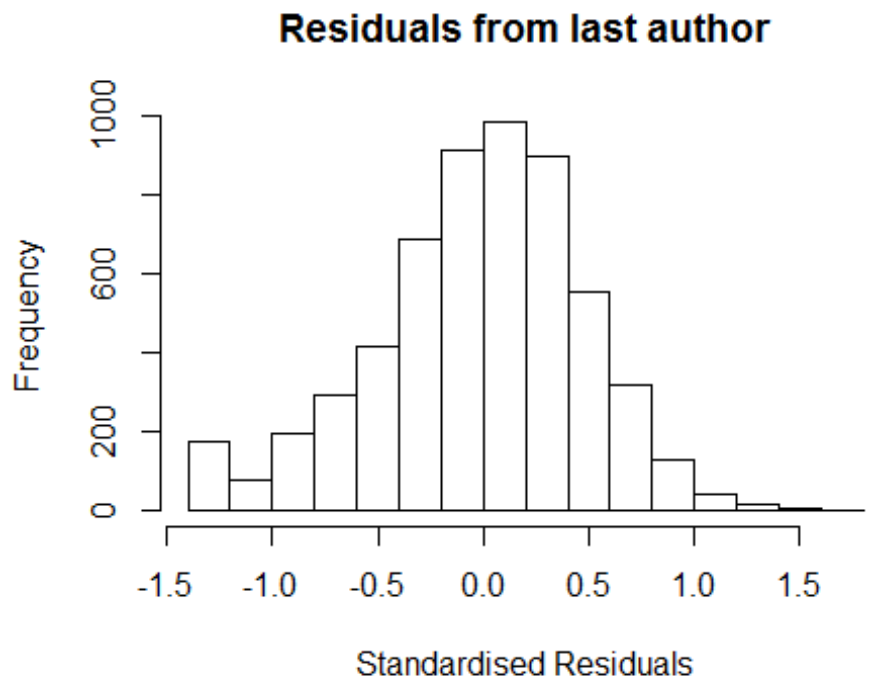

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.3522 -0.3118  0.0201  0.3160  1.6653
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)          1.21389    0.03896   31.15   <2e-16 ***
## FirstAuthorFemale1  0.02664    0.01487    1.79   0.0732 .
## LastAuthorFemale1   0.00195    0.01834    0.11   0.9155
## Year1997            0.05300    0.04687    1.13   0.2582
## Year1998            0.13830    0.04959    2.79   0.0053 **
## Year1999            0.04993    0.04677    1.07   0.2857
## Year2000            0.10479    0.04626    2.27   0.0235 *
## Year2001            0.02627    0.04584    0.57   0.5666
## Year2002            0.00387    0.04831    0.08   0.9361
## Year2003            0.04054    0.04554    0.89   0.3734
## Year2004            0.08268    0.05084    1.63   0.1040
## Year2005            0.01264    0.04893    0.26   0.7962
```

```
## Year2006            0.06276    0.04672    1.34   0.1792
## Year2007            0.05094    0.04673    1.09   0.2757
## Year2008            0.05037    0.04551    1.11   0.2684
## Year2009            0.02981    0.04698    0.63   0.5258
## Year2010           -0.02344    0.04728   -0.50   0.6200
## Year2011            0.01022    0.04684    0.22   0.8273
## Year2012           -0.04272    0.04615   -0.93   0.3547
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.462
## Multiple R-squared:  0.00815,    Adjusted R-squared:  0.00501
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
##  473 weights are ~= 1. The remaining 5227 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.167   0.860   0.951   0.894   0.985   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.75e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.026  1           1.013
## Year              1.026 16           1.001
```




```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##    Min     1Q Median     3Q    Max
## -1.352 -0.312  0.020  0.316  1.665
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)          1.21407    0.03899   31.14   <2e-16 ***
## FirstAuthorFemale1  0.02680    0.01490    1.80   0.0721 .
## Year1997            0.05300    0.04687    1.13   0.2583
## Year1998            0.13842    0.04954    2.79   0.0052 **
## Year1999            0.05001    0.04674    1.07   0.2847
## Year2000            0.10484    0.04624    2.27   0.0234 *
## Year2001            0.02631    0.04582    0.57   0.5659
## Year2002            0.00389    0.04831    0.08   0.9358
## Year2003            0.04059    0.04553    0.89   0.3728
## Year2004            0.08268    0.05085    1.63   0.1040
## Year2005            0.01274    0.04890    0.26   0.7944
## Year2006            0.06282    0.04671    1.34   0.1787
```

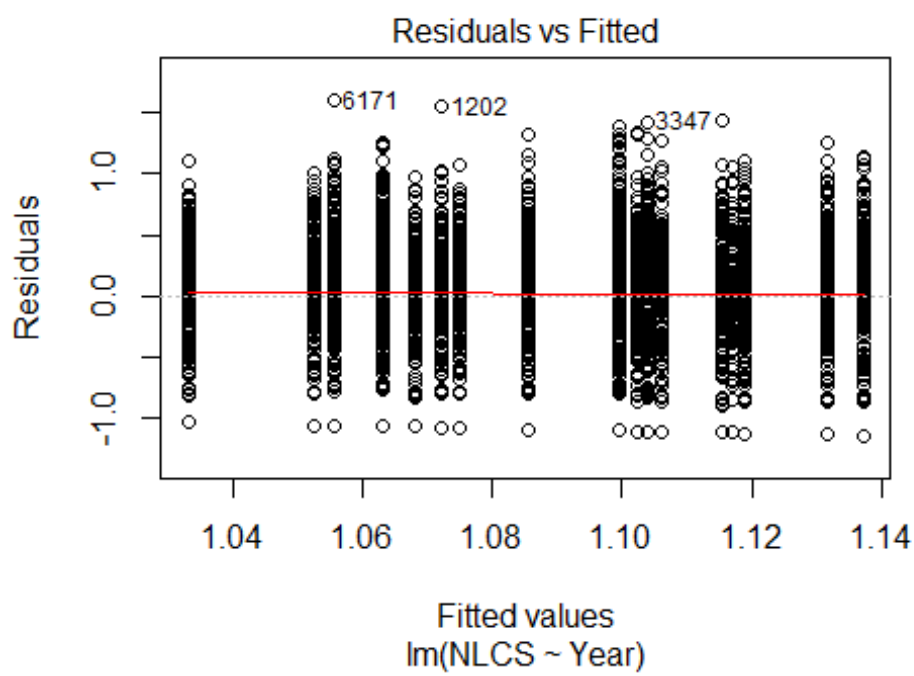
```
## Year2007            0.05104    0.04670    1.09   0.2744
## Year2008            0.05048    0.04545    1.11   0.2667
## Year2009            0.02997    0.04691    0.64   0.5229
## Year2010           -0.02332    0.04722   -0.49   0.6214
## Year2011            0.01031    0.04679    0.22   0.8256
## Year2012           -0.04258    0.04607   -0.92   0.3554
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.462
## Multiple R-squared:  0.00814,    Adjusted R-squared:  0.00518
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
##  471 weights are ~= 1. The remaining 5229 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.167   0.860   0.951   0.894   0.985   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.75e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##                   GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.018  1           1.009
## Year             1.018 16           1.001
```



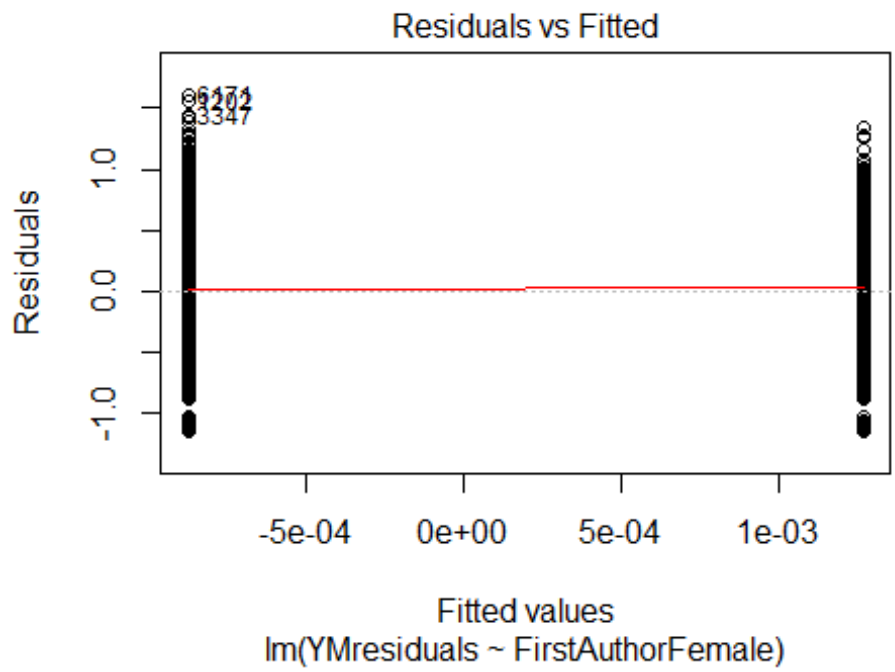

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##    Min     1Q Median     3Q    Max
## -1.356 -0.314  0.022  0.313  1.657
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.21931    0.03875   31.47   <2e-16 ***
## LastAuthorFemale1   0.00517    0.01833    0.28   0.7779
## Year1997            0.05378    0.04680    1.15   0.2505
## Year1998            0.13709    0.04947    2.77   0.0056 **
## Year1999            0.04925    0.04664    1.06   0.2911
## Year2000            0.10444    0.04613    2.26   0.0236 *
## Year2001            0.02602    0.04578    0.57   0.5698
## Year2002            0.00291    0.04823    0.06   0.9519
## Year2003            0.04207    0.04540    0.93   0.3541
## Year2004            0.08352    0.05081    1.64   0.1003
## Year2005            0.01258    0.04885    0.26   0.7968
## Year2006            0.06352    0.04661    1.36   0.1730
```

```
## Year2007           0.05336    0.04656    1.15   0.2518
## Year2008           0.05266    0.04546    1.16   0.2467
## Year2009           0.03254    0.04685    0.69   0.4874
## Year2010          -0.02107    0.04718   -0.45   0.6553
## Year2011           0.01251    0.04677    0.27   0.7891
## Year2012          -0.03906    0.04604   -0.85   0.3962
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.462
## Multiple R-squared:  0.00759,    Adjusted R-squared:  0.00462
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
##  467 weights are ~= 1. The remaining 5233 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.171   0.860   0.951   0.894   0.985   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.75e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Sample size for the above analysis:  5700"
## [1] ""
## [1] ""
## [1] "####################################"
## [1] "Analysis of AJSC 2722"
## [1] "####################################"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  425  486  372  406  408  396  389  354  389  358  427  486  525  650  783
## 2011 2012
##  705  631
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  279  336  237  236  206  143  233  220  267  226  287  329  351  430  540
## 2011 2012
```

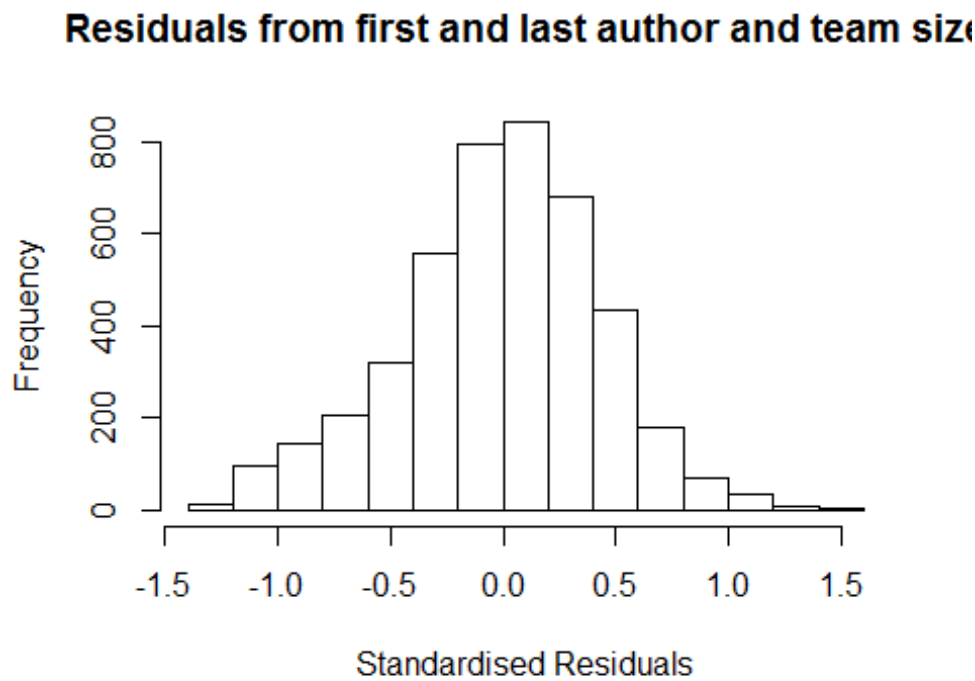
```
##  491  455
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  232  279  198  197  170  127  203  176  222  184  244  283  295  354  443
## 2011 2012
##  423  363
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
##  Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 49, df = 16, p-value = 4e-05
```




```
##
##  Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 8.3, df = 1, p-value = 0.004
```

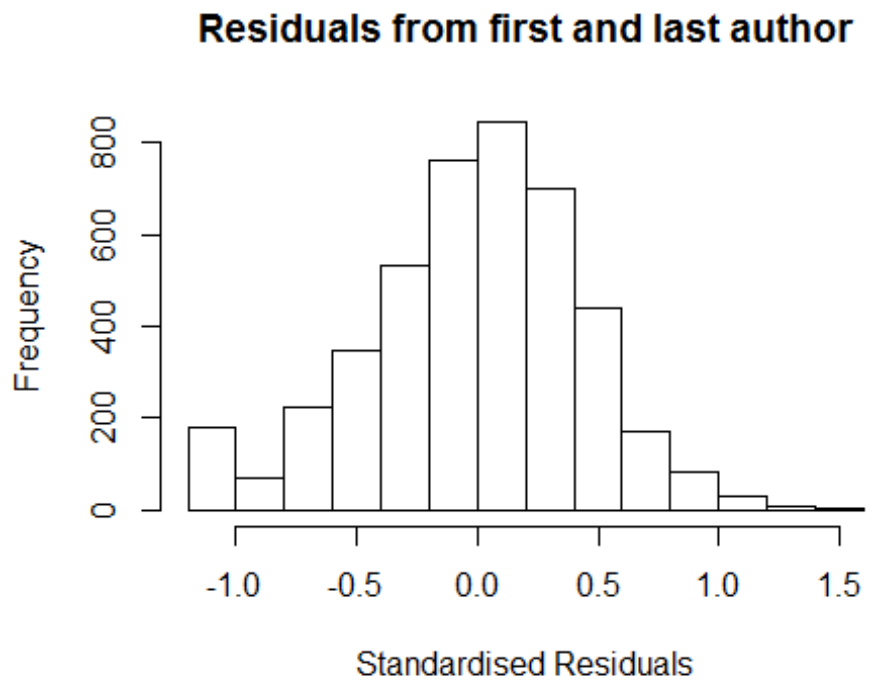



```
## [1] "Female first author team size 2018 geometric mean: 4.38122511192199"
## [1] "Male first author team size 2018 geometric mean: 4.04862140841308"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 26000, p-value = 0.2
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 4.15122666547634"
## [1] "Male last author team size 2018 geometric mean: 4.20949451317323"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 23000, p-value = 0.8
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.031  1           1.015
## LastAuthorFemale  1.039  1           1.019
## UniqueAuthors     1.148  4           1.017
## Year              1.168 16           1.005
```



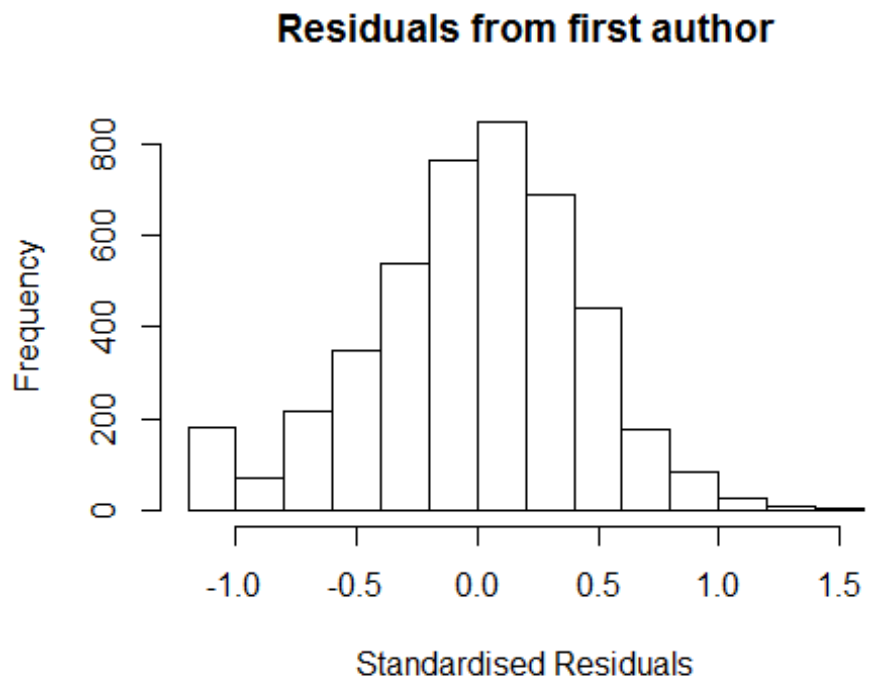

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
UniqueAuthors +
##     Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##     k.max = 1000))
##  \--> method = "MM"
## Residuals:
##    Min     1Q Median     3Q    Max
## -1.226 -0.278  0.014  0.285  1.589
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         0.99628    0.04263   23.37  < 2e-16 ***
## FirstAuthorFemale1  0.01193    0.01378    0.87   0.3866
## LastAuthorFemale1  -0.01620    0.01495   -1.08   0.2787
## UniqueAuthors2      0.09755    0.03434    2.84   0.0045 **
## UniqueAuthors3      0.14790    0.03432    4.31  1.7e-05 ***
## UniqueAuthors4      0.18008    0.03375    5.34  1.0e-07 ***
## UniqueAuthors5      0.28698    0.03263    8.80  < 2e-16 ***
## Year1997           -0.00926    0.04163   -0.22   0.8240
## Year1998           -0.07291    0.04227   -1.72   0.0847 .
## Year1999           -0.02233    0.04287   -0.52   0.6024
```

```
## Year2000           -0.02485    0.04570   -0.54   0.5866
## Year2001           -0.03208    0.04622   -0.69   0.4877
## Year2002           -0.12007    0.04446   -2.70   0.0070 **
## Year2003           -0.06367    0.04490   -1.42   0.1563
## Year2004           -0.13396    0.04218   -3.18   0.0015 **
## Year2005           -0.10556    0.04073   -2.59   0.0096 **
## Year2006           -0.09423    0.04315   -2.18   0.0290 *
## Year2007           -0.08457    0.04009   -2.11   0.0350 *
## Year2008           -0.05732    0.04130   -1.39   0.1653
## Year2009           -0.10801    0.03933   -2.75   0.0061 **
## Year2010           -0.09818    0.03871   -2.54   0.0112 *
## Year2011           -0.05887    0.03843   -1.53   0.1256
## Year2012           -0.08623    0.04146   -2.08   0.0376 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.417
## Multiple R-squared:  0.0456, Adjusted R-squared:  0.0408
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
##  406 weights are ~= 1. The remaining 3987 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.115   0.860   0.949   0.896   0.986   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          2.28e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.025  1           1.012
## LastAuthorFemale  1.031  1           1.015
## Year              1.047 16           1.001
```



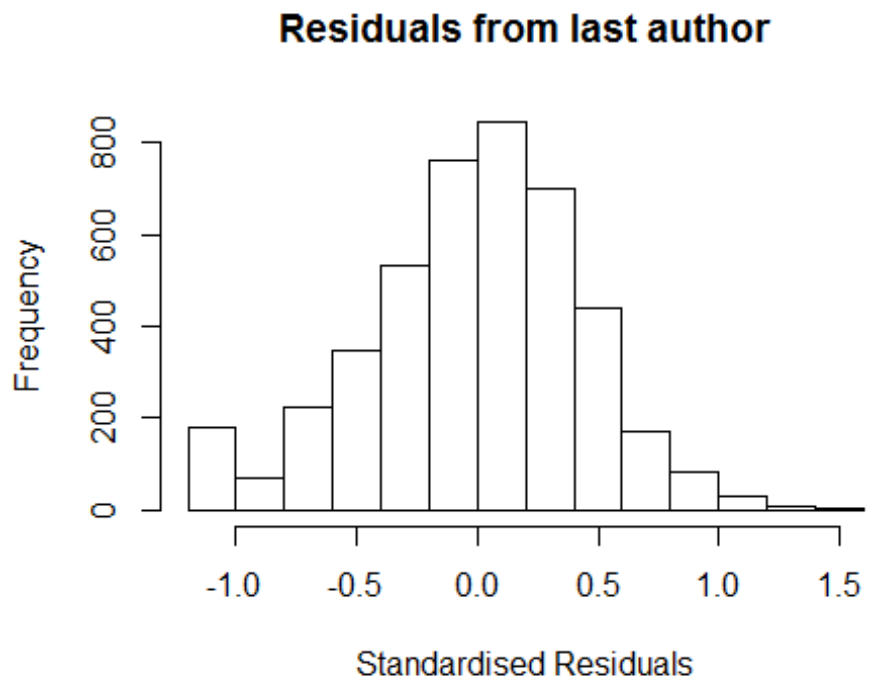

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.1682 -0.2818  0.0206  0.2951  1.5846
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)          1.14738    0.03155   36.37   <2e-16 ***
## FirstAuthorFemale1  0.02078    0.01395    1.49    0.136
## LastAuthorFemale1  -0.02271    0.01508   -1.51    0.132
## Year1997           -0.00873    0.04138   -0.21    0.833
## Year1998           -0.07330    0.04166   -1.76    0.079 .
## Year1999           -0.00861    0.04230   -0.20    0.839
## Year2000           -0.02186    0.04635   -0.47    0.637
## Year2001           -0.02281    0.04485   -0.51    0.611
## Year2002           -0.09373    0.04430   -2.12    0.034 *
## Year2003           -0.02716    0.04513   -0.60    0.547
## Year2004           -0.09776    0.04231   -2.31    0.021 *
## Year2005           -0.05892    0.04105   -1.44    0.151
```

```
## Year2006           -0.04862    0.04240   -1.15    0.252
## Year2007           -0.04494    0.03934   -1.14    0.253
## Year2008           -0.03160    0.04193   -0.75    0.451
## Year2009           -0.07497    0.03964   -1.89    0.059 .
## Year2010           -0.06097    0.03862   -1.58    0.115
## Year2011           -0.02543    0.03808   -0.67    0.504
## Year2012           -0.04302    0.04110   -1.05    0.295
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.424
## Multiple R-squared:  0.0047, Adjusted R-squared:  0.000606
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  371 weights are ~= 1. The remaining 4022 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.132   0.865   0.948   0.896   0.986   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          2.28e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.021  1           1.011
## Year              1.021 16           1.001
```




```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.1601 -0.2842  0.0191  0.2936  1.5914
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)          1.14178    0.03133   36.44   <2e-16 ***
## FirstAuthorFemale1  0.01831    0.01394    1.31    0.189
## Year1997           -0.00886    0.04145   -0.21    0.831
## Year1998           -0.07269    0.04178   -1.74    0.082 .
## Year1999           -0.00959    0.04227   -0.23    0.821
## Year2000           -0.02203    0.04638   -0.47    0.635
## Year2001           -0.02182    0.04470   -0.49    0.625
## Year2002           -0.09339    0.04430   -2.11    0.035 *
## Year2003           -0.02506    0.04515   -0.55    0.579
## Year2004           -0.09779    0.04238   -2.31    0.021 *
## Year2005           -0.05748    0.04105   -1.40    0.162
## Year2006           -0.04664    0.04238   -1.10    0.271
```

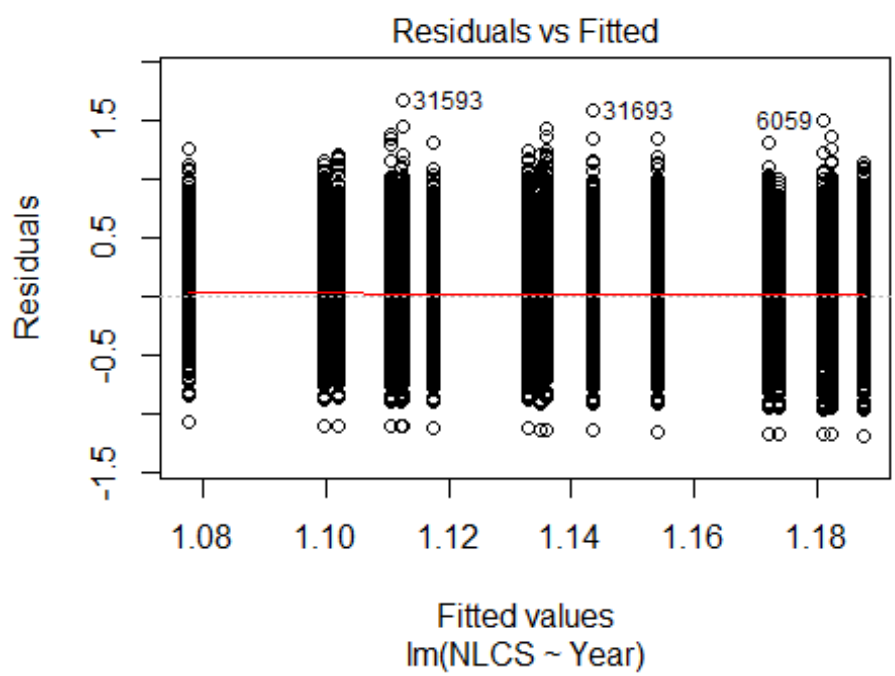
```
## Year2007           -0.04454    0.03941   -1.13    0.258
## Year2008           -0.03102    0.04196   -0.74    0.460
## Year2009           -0.07617    0.03965   -1.92    0.055 .
## Year2010           -0.06088    0.03866   -1.57    0.115
## Year2011           -0.02597    0.03808   -0.68    0.495
## Year2012           -0.04391    0.04110   -1.07    0.285
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.424
## Multiple R-squared:  0.00419,    Adjusted R-squared:  0.000326
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  365 weights are ~= 1. The remaining 4028 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.129   0.866   0.949   0.896   0.985   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          2.28e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##                   GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.027  1           1.013
## Year             1.027 16           1.001
```



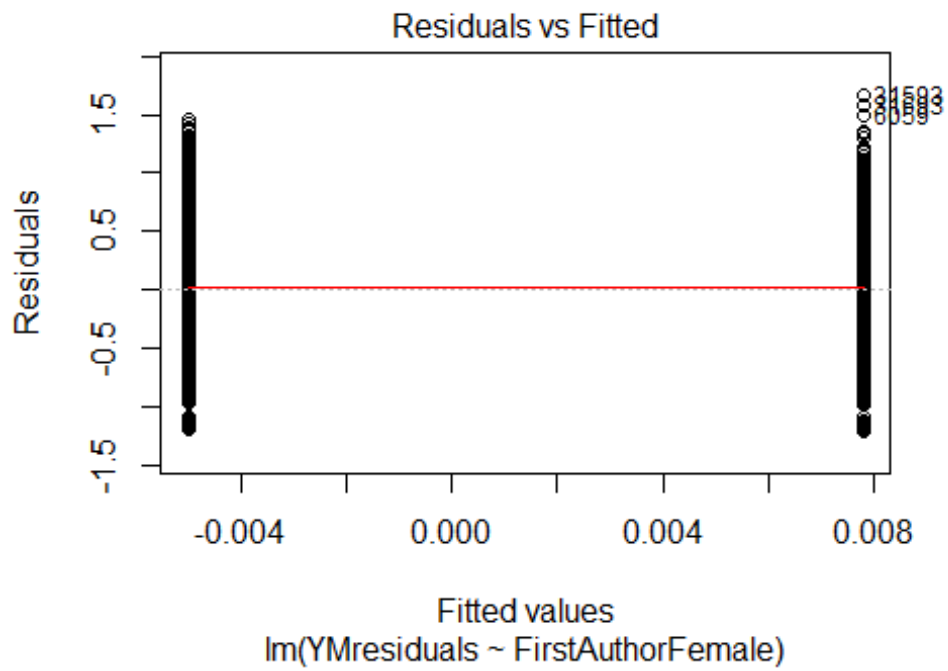

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.1555 -0.2821  0.0206  0.2931  1.5759
##
## Coefficients:
##                   Estimate Std. Error t value Pr(>|t|)
## (Intercept)        1.15553    0.03084   37.47   <2e-16 ***
## LastAuthorFemale1 -0.02002    0.01509   -1.33    0.185
## Year1997          -0.00992    0.04128   -0.24    0.810
## Year1998          -0.07494    0.04163   -1.80    0.072 .
## Year1999          -0.00870    0.04227   -0.21    0.837
## Year2000          -0.02305    0.04635   -0.50    0.619
## Year2001          -0.02137    0.04495   -0.48    0.635
## Year2002          -0.09402    0.04430   -2.12    0.034 *
## Year2003          -0.02826    0.04510   -0.63    0.531
## Year2004          -0.09814    0.04229   -2.32    0.020 *
## Year2005          -0.05931    0.04100   -1.45    0.148
## Year2006          -0.05061    0.04231   -1.20    0.232
```

```
## Year2007          -0.04613    0.03928   -1.17    0.240
## Year2008          -0.03050    0.04189   -0.73    0.467
## Year2009          -0.07441    0.03958   -1.88    0.060 .
## Year2010          -0.06097    0.03855   -1.58    0.114
## Year2011          -0.02505    0.03802   -0.66    0.510
## Year2012          -0.04239    0.04102   -1.03    0.301
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.424
## Multiple R-squared:  0.00421,    Adjusted R-squared:  0.000338
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  372 weights are ~= 1. The remaining 4021 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.138   0.865   0.949   0.896   0.985   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          2.28e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Sample size for the above analysis:  4393"
## [1] ""
## [1] ""
## [1] "####################################"
## [1] "Analysis of AJSC 2723"
## [1] "####################################"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 2031 2093 2031 2019 2056 2094 2075 1651 1791 1887 1946 1904 2133 2240 2279
## 2011 2012
## 2519 2550
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 1280 1318 1324 1363 1191 1054 1484 1180 1276 1385 1285 1311 1444 1497 1536
## 2011 2012
```

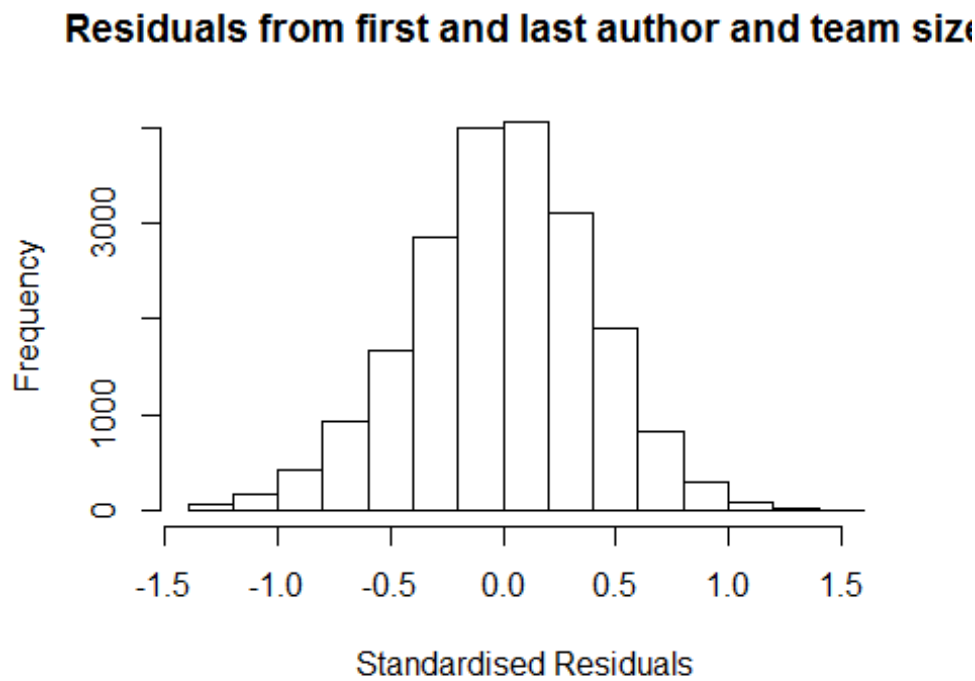
```
## 1739 1762
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 1123 1165 1170 1177 1051  900 1304 1034 1112 1208 1120 1125 1254 1287 1314
## 2011 2012
## 1518 1512
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
##  Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 64, df = 16, p-value = 1e-07
```




```
##
##  Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 83, df = 1, p-value <2e-16
```

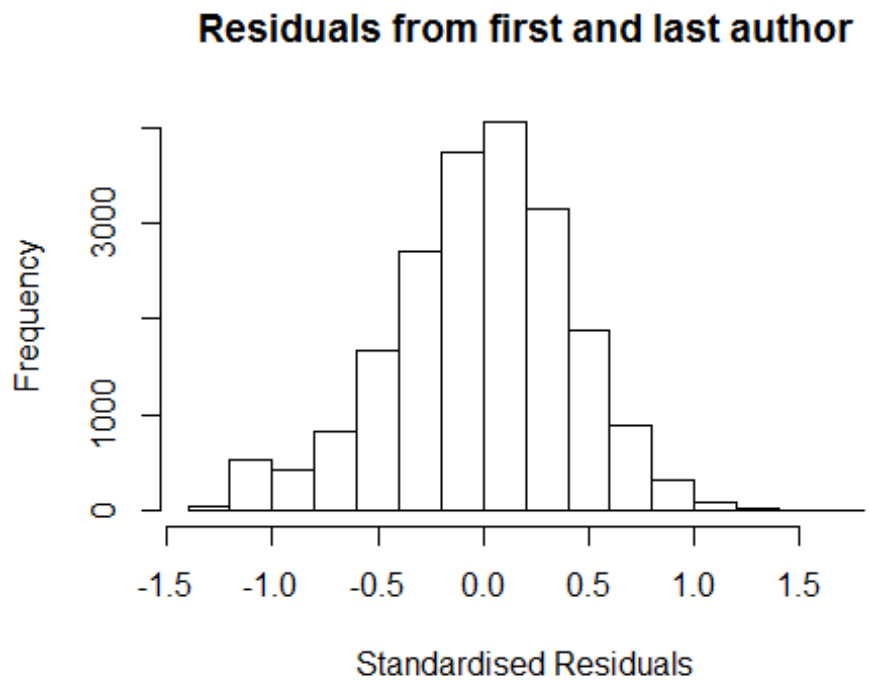



```
## [1] "Female first author team size 2018 geometric mean: 5.80149587099122"
## [1] "Male first author team size 2018 geometric mean: 5.705210991105"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 4e+05, p-value = 0.8
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 5.77000442212655"
## [1] "Male last author team size 2018 geometric mean: 5.74328404168252"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 350000, p-value = 0.8
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##                     GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.026  1           1.013
## LastAuthorFemale  1.020  1           1.010
## UniqueAuthors     1.049  4           1.006
## Year              1.058 16           1.002
```



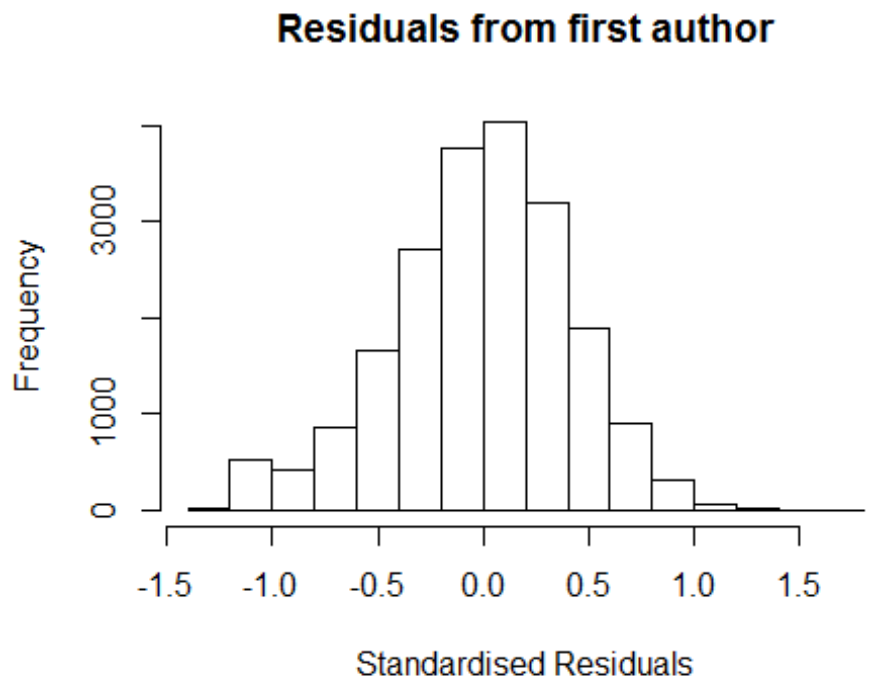

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
UniqueAuthors +
##     Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##     k.max = 1000))
##  \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.33295 -0.26136  0.00453  0.26429  1.56857
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         0.88141    0.02042   43.16  < 2e-16 ***
## FirstAuthorFemale1 -0.00834    0.00586   -1.42    0.155
## LastAuthorFemale1  -0.01693    0.00671   -2.52    0.012 *
## UniqueAuthors2      0.19186    0.01889   10.16  < 2e-16 ***
## UniqueAuthors3      0.24044    0.01772   13.57  < 2e-16 ***
## UniqueAuthors4      0.32069    0.01749   18.34  < 2e-16 ***
## UniqueAuthors5      0.44904    0.01638   27.41  < 2e-16 ***
## Year1997            0.00167    0.01854    0.09    0.928
## Year1998            0.00251    0.01814    0.14    0.890
## Year1999           -0.01005    0.01772   -0.57    0.571
```

```
## Year2000            0.00243    0.01874    0.13    0.897
## Year2001           -0.06376    0.01937   -3.29    0.001 ***
## Year2002           -0.09378    0.01730   -5.42  6.0e-08 ***
## Year2003           -0.09891    0.01823   -5.43  5.9e-08 ***
## Year2004           -0.11071    0.01784   -6.20  5.6e-10 ***
## Year2005           -0.11187    0.01801   -6.21  5.4e-10 ***
## Year2006           -0.13219    0.01838   -7.19  6.6e-13 ***
## Year2007           -0.10991    0.01830   -6.00  2.0e-09 ***
## Year2008           -0.10868    0.01770   -6.14  8.5e-10 ***
## Year2009           -0.08186    0.01750   -4.68  2.9e-06 ***
## Year2010           -0.07678    0.01795   -4.28  1.9e-05 ***
## Year2011           -0.09465    0.01741   -5.44  5.5e-08 ***
## Year2012           -0.07970    0.01800   -4.43  9.6e-06 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.391
## Multiple R-squared:  0.108,  Adjusted R-squared:  0.107
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
##  1696 weights are ~= 1. The remaining 18678 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0712  0.8680  0.9510  0.9010  0.9860  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          4.91e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.017  1           1.009
## LastAuthorFemale  1.015  1           1.007
## Year              1.015 16           1.000
```



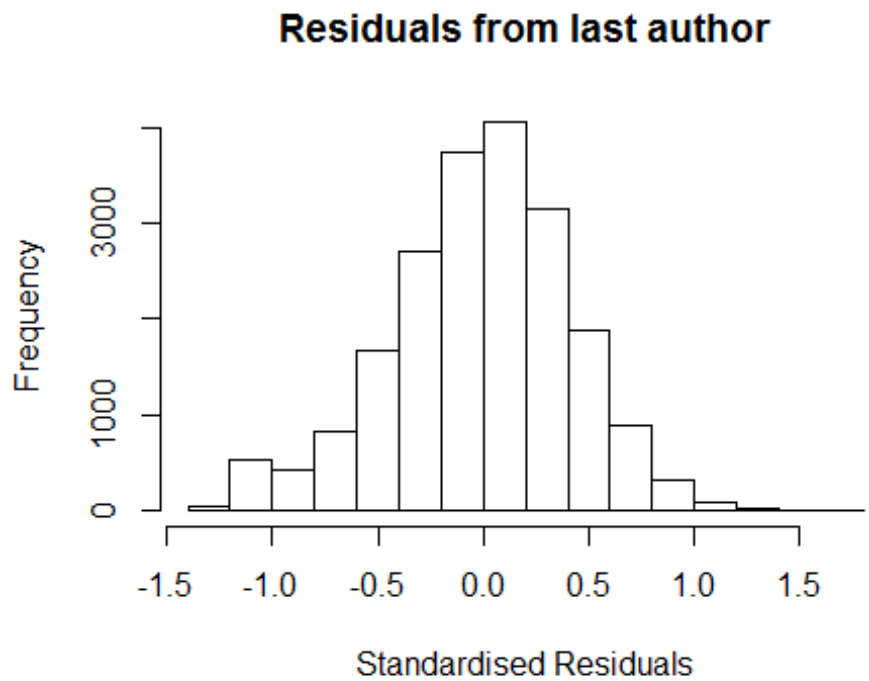

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.2160 -0.2741  0.0102  0.2696  1.6434
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)          1.19772    0.01439   83.25  < 2e-16 ***
## FirstAuthorFemale1  0.01222    0.00607    2.01  0.04422 *
## LastAuthorFemale1  -0.02172    0.00696   -3.12  0.00180 **
## Year1997           -0.00320    0.01915   -0.17  0.86713
## Year1998           -0.00409    0.01856   -0.22  0.82541
## Year1999           -0.00643    0.01806   -0.36  0.72197
## Year2000            0.00604    0.01945    0.31  0.75597
## Year2001           -0.02970    0.02007   -1.48  0.13889
## Year2002           -0.07682    0.01800   -4.27  2.0e-05 ***
## Year2003           -0.05763    0.01900   -3.03  0.00243 **
## Year2004           -0.06805    0.01869   -3.64  0.00027 ***
## Year2005           -0.07111    0.01894   -3.75  0.00017 ***
```

```
## Year2006           -0.09636    0.01927   -5.00  5.7e-07 ***
## Year2007           -0.08365    0.01906   -4.39  1.1e-05 ***
## Year2008           -0.07137    0.01868   -3.82  0.00013 ***
## Year2009           -0.03736    0.01807   -2.07  0.03870 *
## Year2010           -0.02230    0.01839   -1.21  0.22545
## Year2011           -0.05609    0.01801   -3.11  0.00184 **
## Year2012           -0.04589    0.01852   -2.48  0.01325 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.403
## Multiple R-squared:  0.00633,    Adjusted R-squared:  0.00545
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  1716 weights are ~= 1. The remaining 18658 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0595  0.8680  0.9510  0.8970  0.9860  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          4.91e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##                   GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.01  1           1.005
## Year              1.01 16           1.000
```




```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.2093 -0.2751  0.0101  0.2688  1.6508
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)          1.19414    0.01433   83.31  < 2e-16 ***
## FirstAuthorFemale1   0.00997    0.00605    1.65  0.09960 .
## Year1997            -0.00328    0.01916   -0.17  0.86425
## Year1998            -0.00455    0.01857   -0.25  0.80641
## Year1999            -0.00692    0.01807   -0.38  0.70162
## Year2000             0.00521    0.01945    0.27  0.78892
## Year2001            -0.03013    0.02006   -1.50  0.13321
## Year2002            -0.07712    0.01801   -4.28  1.9e-05 ***
## Year2003            -0.05768    0.01900   -3.04  0.00241 **
## Year2004            -0.06880    0.01870   -3.68  0.00023 ***
## Year2005            -0.07220    0.01895   -3.81  0.00014 ***
## Year2006            -0.09722    0.01927   -5.05  4.6e-07 ***
```

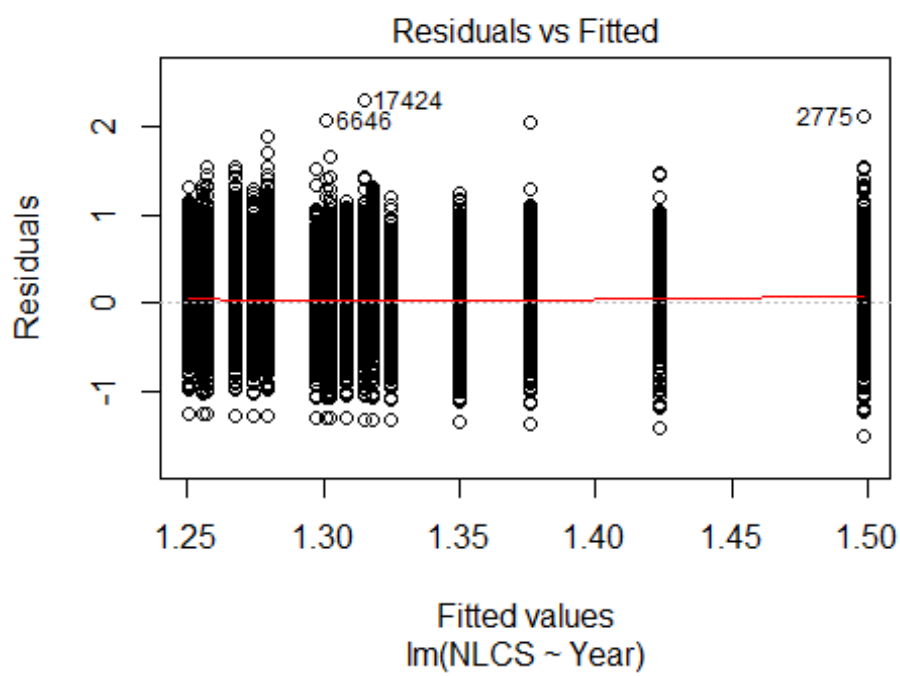
```
## Year2007           -0.08503    0.01907   -4.46  8.3e-06 ***
## Year2008           -0.07291    0.01868   -3.90  9.5e-05 ***
## Year2009           -0.03871    0.01808   -2.14  0.03227 *
## Year2010           -0.02394    0.01839   -1.30  0.19314
## Year2011           -0.05768    0.01801   -3.20  0.00136 **
## Year2012           -0.04703    0.01853   -2.54  0.01113 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.404
## Multiple R-squared:  0.00585,    Adjusted R-squared:  0.00502
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  1701 weights are ~= 1. The remaining 18673 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0564  0.8680  0.9510  0.8970  0.9860  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          4.91e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##                   GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.007  1           1.004
## Year             1.007 16           1.000
```



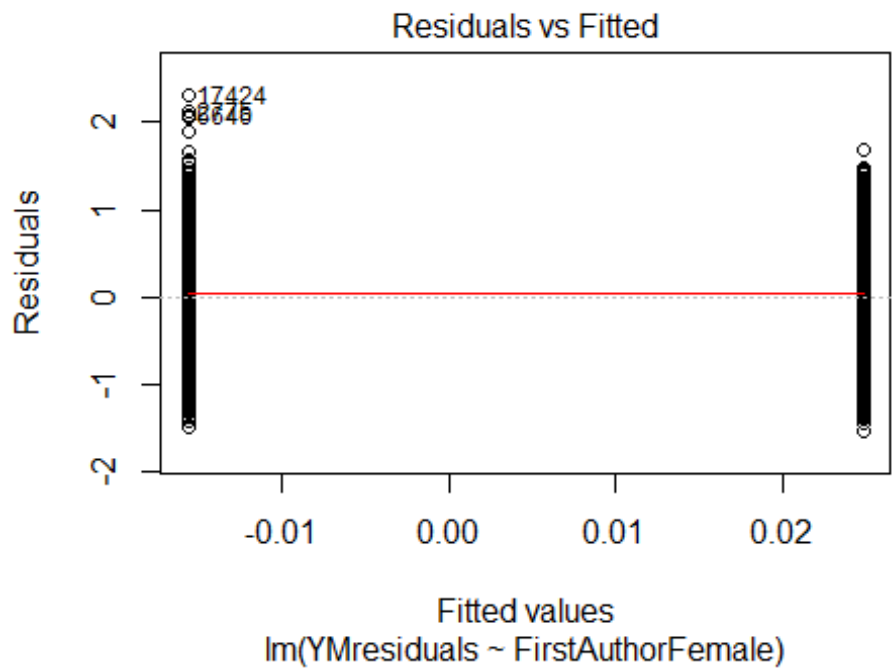

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.2079 -0.2737  0.0102  0.2696  1.6509
##
## Coefficients:
##                   Estimate Std. Error t value Pr(>|t|)
## (Intercept)        1.20175    0.01420   84.64  < 2e-16 ***
## LastAuthorFemale1 -0.02008    0.00693   -2.90  0.00379 **
## Year1997          -0.00339    0.01915   -0.18  0.85969
## Year1998          -0.00408    0.01856   -0.22  0.82588
## Year1999          -0.00639    0.01807   -0.35  0.72363
## Year2000           0.00619    0.01944    0.32  0.75026
## Year2001          -0.02982    0.02007   -1.49  0.13725
## Year2002          -0.07687    0.01801   -4.27  2.0e-05 ***
## Year2003          -0.05706    0.01900   -3.00  0.00267 **
## Year2004          -0.06736    0.01869   -3.60  0.00031 ***
## Year2005          -0.07042    0.01895   -3.72  0.00020 ***
## Year2006          -0.09573    0.01927   -4.97  6.8e-07 ***
```

```
## Year2007          -0.08262    0.01906   -4.33  1.5e-05 ***
## Year2008          -0.07063    0.01868   -3.78  0.00016 ***
## Year2009          -0.03651    0.01807   -2.02  0.04341 *
## Year2010          -0.02132    0.01839   -1.16  0.24635
## Year2011          -0.05497    0.01800   -3.05  0.00227 **
## Year2012          -0.04475    0.01852   -2.42  0.01569 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.403
## Multiple R-squared:  0.00612,    Adjusted R-squared:  0.00529
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  1702 weights are ~= 1. The remaining 18672 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0562  0.8680  0.9510  0.8970  0.9860  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          4.91e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Sample size for the above analysis:  20374"
## [1] ""
## [1] ""
## [1] "####################################"
## [1] "Analysis of AJSC 2724"
## [1] "####################################"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 1031 1481  909  976  921 1089 1043  867  980 1096 1243 1358 1472 1232 1416
## 2011 2012
## 1600 1411
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  759  700  663  719  585  653  773  681  769  819  934 1025 1119  930 1037
## 2011 2012
```

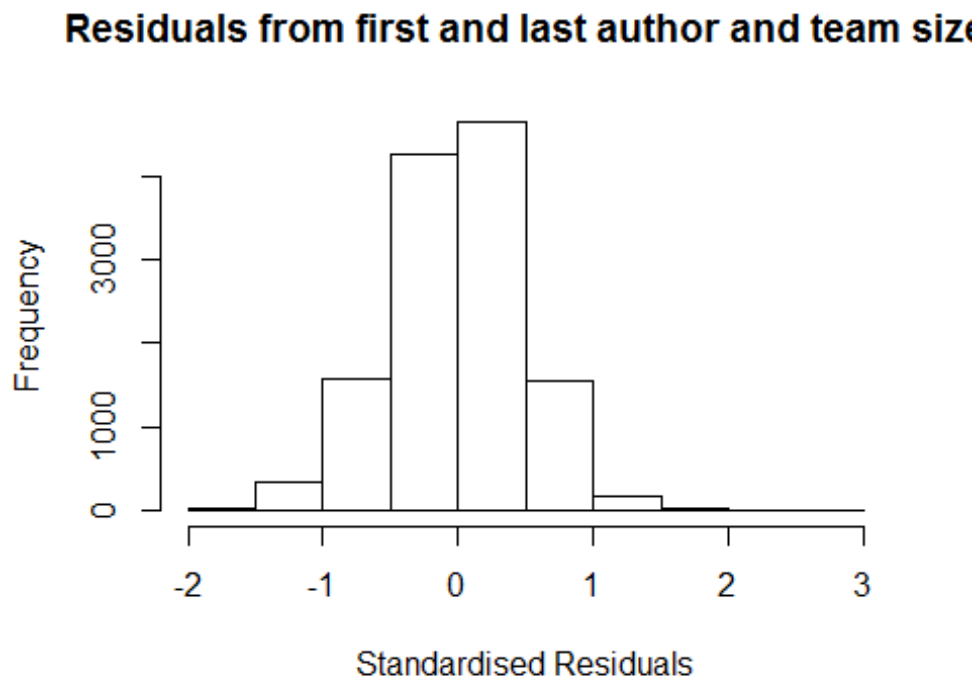
```
## 1156 1000
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  674  611  581  626  513  575  694  601  680  697  819  913  992  804  912
## 2011 2012
## 1019  844
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
##  Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 250, df = 16, p-value <2e-16
```




```
##
##  Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 38, df = 1, p-value = 8e-10
```

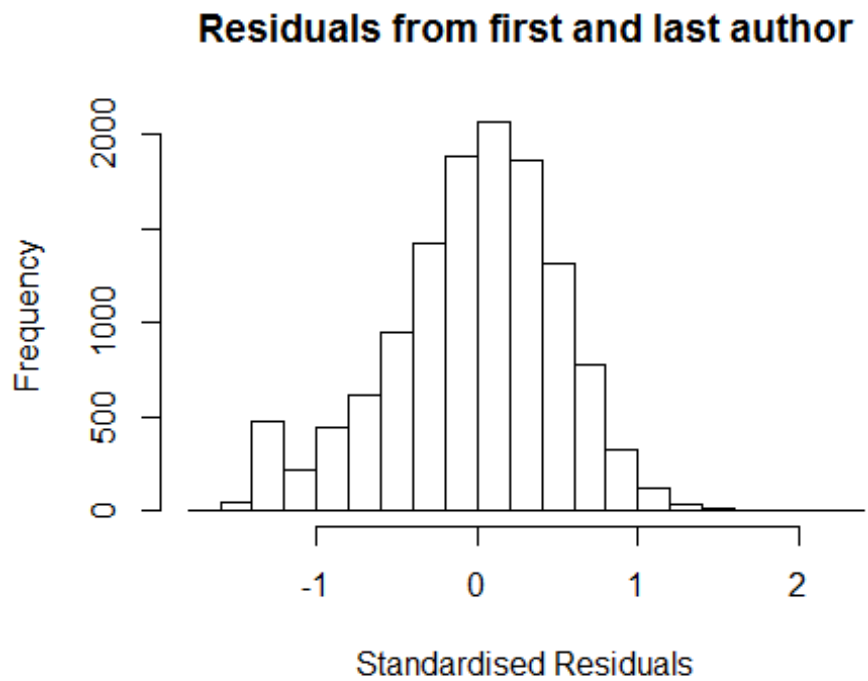



```
## [1] "Female first author team size 2018 geometric mean: 5.20938121475472"
## [1] "Male first author team size 2018 geometric mean: 4.92074244016676"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 120000, p-value = 0.2
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 4.66572841613816"
## [1] "Male last author team size 2018 geometric mean: 5.30677998207053"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 92000, p-value = 0.003
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##                     GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.045  1           1.022
## LastAuthorFemale  1.047  1           1.023
## UniqueAuthors     1.054  4           1.007
## Year              1.081 16           1.002
```



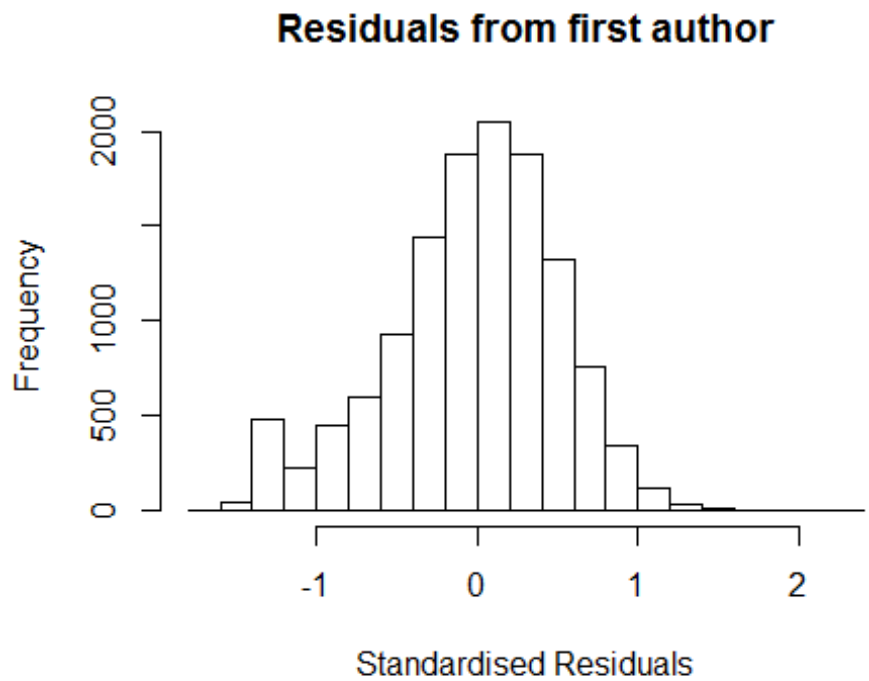

```
## [1] "List of  1 outliers with residuals above  2.5"
##        ScopusId  NLCS Year OneField Fields residuals
## 2775 1842334456 3.603 1997     2712      3     2.634
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
UniqueAuthors +
##     Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##     k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.7282 -0.3210  0.0106  0.3178  2.6338
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)         0.820776   0.030011   27.35  < 2e-16 ***
## FirstAuthorFemale1 -0.002058   0.008995   -0.23   0.8191
## LastAuthorFemale1  -0.011122   0.010164   -1.09   0.2739
## UniqueAuthors2      0.385148   0.025049   15.38  < 2e-16 ***
## UniqueAuthors3      0.542734   0.023147   23.45  < 2e-16 ***
## UniqueAuthors4      0.644520   0.022313   28.89  < 2e-16 ***
## UniqueAuthors5      0.759032   0.019878   38.18  < 2e-16 ***
## Year1997            0.148418   0.033599    4.42  1.0e-05 ***
## Year1998            0.020323   0.030939    0.66   0.5113
## Year1999           -0.000783   0.030619   -0.03   0.9796
```

```
## Year2000           -0.041623   0.031286   -1.33   0.1834
## Year2001           -0.095166   0.030503   -3.12   0.0018 **
## Year2002           -0.068741   0.029826   -2.30   0.0212 *
## Year2003           -0.123466   0.029760   -4.15  3.4e-05 ***
## Year2004           -0.135022   0.029167   -4.63  3.7e-06 ***
## Year2005           -0.161684   0.028958   -5.58  2.4e-08 ***
## Year2006           -0.113902   0.028853   -3.95  7.9e-05 ***
## Year2007           -0.118305   0.028608   -4.14  3.6e-05 ***
## Year2008           -0.175510   0.028051   -6.26  4.1e-10 ***
## Year2009           -0.123449   0.029560   -4.18  3.0e-05 ***
## Year2010           -0.162611   0.029833   -5.45  5.1e-08 ***
## Year2011           -0.130102   0.029626   -4.39  1.1e-05 ***
## Year2012           -0.135742   0.030491   -4.45  8.6e-06 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.466
## Multiple R-squared:  0.209,  Adjusted R-squared:  0.208
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
##  observation 1269 is an outlier with |weight| = 0 ( < 8e-06);
##  1045 weights are ~= 1. The remaining 11509 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0008  0.8650  0.9490  0.9000  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          7.96e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.023  1           1.011
## LastAuthorFemale  1.025  1           1.012
## Year              1.039 16           1.001
```



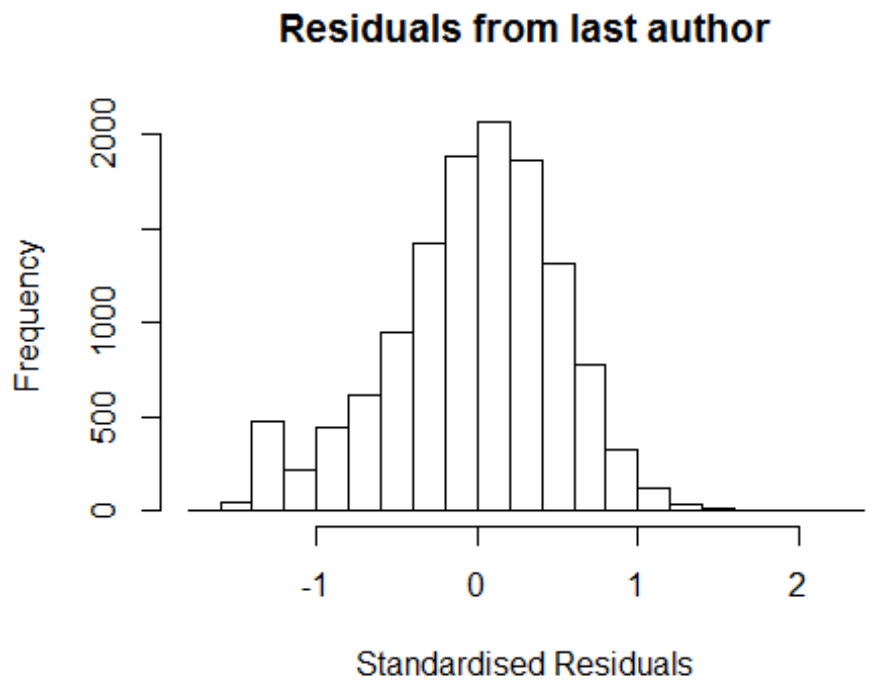

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##    Min     1Q Median     3Q    Max
## -1.614 -0.334  0.024  0.336  2.280
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)          1.36199    0.02645   51.49  < 2e-16 ***
## FirstAuthorFemale1   0.04702    0.00969    4.85  1.2e-06 ***
## LastAuthorFemale1   -0.03455    0.01113   -3.10  0.00192 **
## Year1997             0.20476    0.03653    5.60  2.1e-08 ***
## Year1998             0.07403    0.03339    2.22  0.02665 *
## Year1999             0.01752    0.03330    0.53  0.59890
## Year2000            -0.01408    0.03348   -0.42  0.67404
## Year2001            -0.06146    0.03345   -1.84  0.06617 .
## Year2002             0.00415    0.03223    0.13  0.89744
## Year2003            -0.04931    0.03216   -1.53  0.12526
## Year2004            -0.06266    0.03179   -1.97  0.04875 *
## Year2005            -0.08579    0.03245   -2.64  0.00822 **
```

```
## Year2006           -0.03959    0.03159   -1.25  0.21006
## Year2007           -0.03644    0.03115   -1.17  0.24204
## Year2008           -0.10403    0.03084   -3.37  0.00074 ***
## Year2009           -0.03439    0.03176   -1.08  0.27893
## Year2010           -0.07688    0.03313   -2.32  0.02034 *
## Year2011           -0.08327    0.03261   -2.55  0.01069 *
## Year2012           -0.07025    0.03370   -2.08  0.03710 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.494
## Multiple R-squared:  0.0183, Adjusted R-squared:  0.0169
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
##  1075 weights are ~= 1. The remaining 11480 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0008  0.8600  0.9490  0.8930  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          7.96e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##                   GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.02  1           1.010
## Year              1.02 16           1.001
```




```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.6035 -0.3344  0.0239  0.3356  2.2879
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)          1.35648    0.02637   51.44  < 2e-16 ***
## FirstAuthorFemale1   0.04217    0.00973    4.33  1.5e-05 ***
## Year1997             0.20489    0.03646    5.62  2.0e-08 ***
## Year1998             0.07408    0.03338    2.22  0.02649 *
## Year1999             0.01812    0.03330    0.54  0.58639
## Year2000            -0.01312    0.03344   -0.39  0.69482
## Year2001            -0.06173    0.03342   -1.85  0.06474 .
## Year2002             0.00381    0.03219    0.12  0.90578
## Year2003            -0.05014    0.03211   -1.56  0.11847
## Year2004            -0.06260    0.03176   -1.97  0.04872 *
## Year2005            -0.08710    0.03243   -2.69  0.00725 **
## Year2006            -0.04131    0.03154   -1.31  0.19036
```

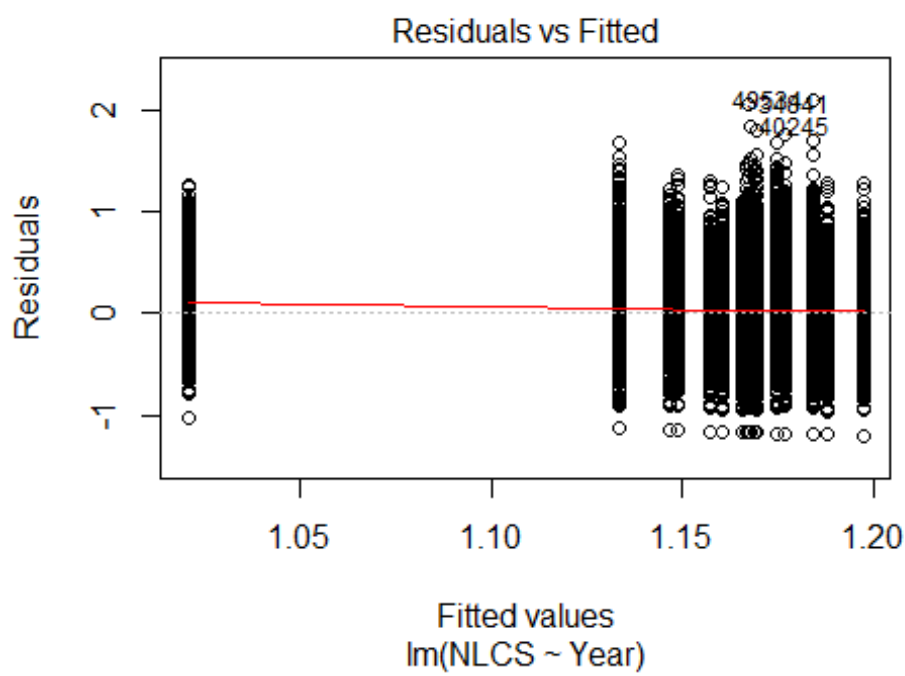
```
## Year2007           -0.04028    0.03108   -1.30  0.19506
## Year2008           -0.10645    0.03079   -3.46  0.00055 ***
## Year2009           -0.03642    0.03172   -1.15  0.25086
## Year2010           -0.07923    0.03309   -2.39  0.01666 *
## Year2011           -0.08591    0.03257   -2.64  0.00836 **
## Year2012           -0.07315    0.03365   -2.17  0.02975 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.494
## Multiple R-squared:  0.0176, Adjusted R-squared:  0.0163
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
##  1090 weights are ~= 1. The remaining 11465 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0005  0.8610  0.9500  0.8930  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          7.96e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##                   GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.022  1           1.011
## Year             1.022 16           1.001
```



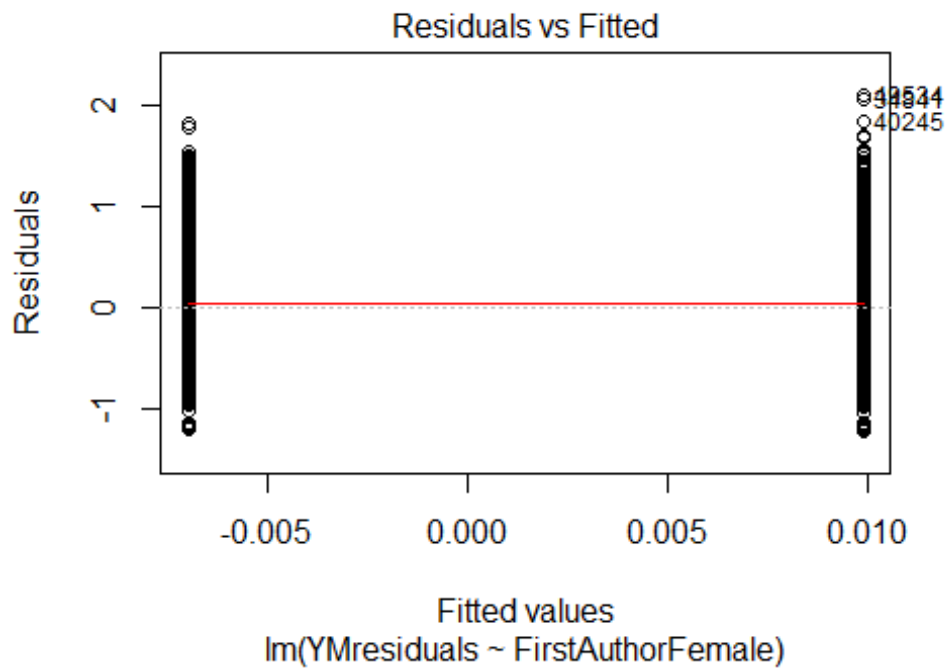

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.5793 -0.3337  0.0229  0.3367  2.2626
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)          1.3747     0.0263   52.25  < 2e-16 ***
## LastAuthorFemale1   -0.0261     0.0112   -2.35   0.0190 *
## Year1997             0.2046     0.0365    5.60  2.2e-08 ***
## Year1998             0.0738     0.0334    2.21   0.0272 *
## Year1999             0.0188     0.0334    0.56   0.5739
## Year2000            -0.0131     0.0335   -0.39   0.6958
## Year2001            -0.0583     0.0335   -1.74   0.0819 .
## Year2002             0.0059     0.0323    0.18   0.8549
## Year2003            -0.0450     0.0322   -1.40   0.1629
## Year2004            -0.0594     0.0318   -1.87   0.0618 .
## Year2005            -0.0803     0.0324   -2.48   0.0132 *
## Year2006            -0.0340     0.0316   -1.08   0.2814
```

```
## Year2007           -0.0313     0.0311   -1.01   0.3148
## Year2008           -0.0967     0.0308   -3.14   0.0017 **
## Year2009           -0.0293     0.0318   -0.92   0.3566
## Year2010           -0.0708     0.0331   -2.14   0.0326 *
## Year2011           -0.0757     0.0326   -2.32   0.0203 *
## Year2012           -0.0646     0.0337   -1.92   0.0552 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.494
## Multiple R-squared:  0.0165, Adjusted R-squared:  0.0151
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
##  1084 weights are ~= 1. The remaining 11471 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.002   0.860   0.949   0.893   0.985   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          7.96e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Sample size for the above analysis:  12555"
## [1] ""
## [1] ""
## [1] "####################################"
## [1] "Analysis of AJSC 2725"
## [1] "####################################"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 1960 2136 1982 2042 2249 2366 2404 1895 2300 2221 2399 2482 2761 3318 3202
## 2011 2012
## 3672 3597
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 1295 1167 1240 1390 1328 1079 1775 1399 1773 1698 1784 1879 2059 2441 2364
## 2011 2012
```

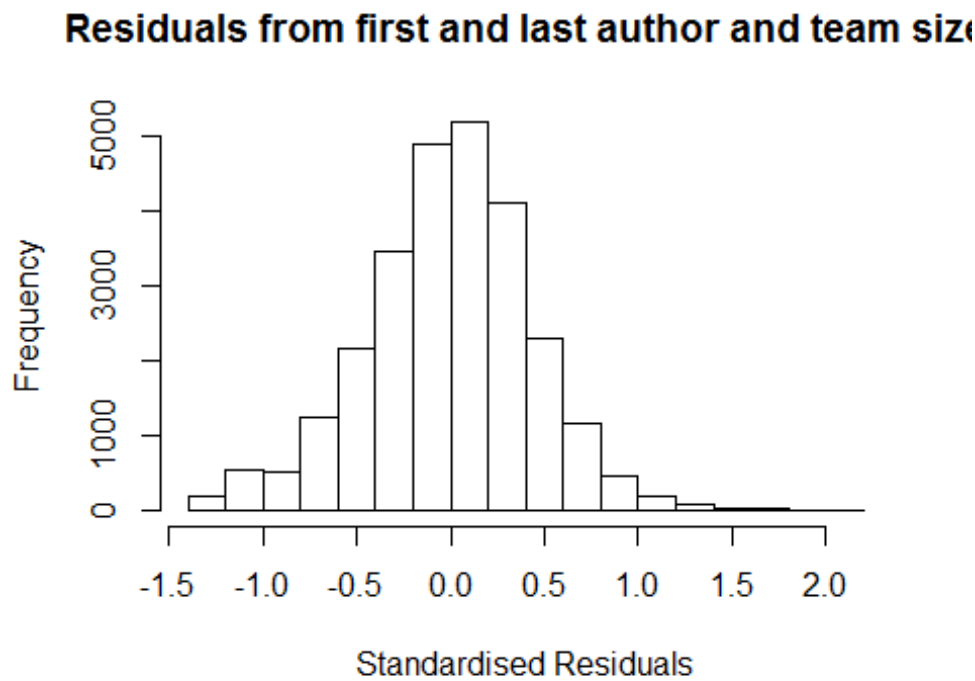
```
## 2755 2656
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 1176 1056 1116 1236 1185  964 1580 1213 1595 1489 1572 1661 1790 2117 2064
## 2011 2012
## 2411 2330
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
##  Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 610, df = 16, p-value <2e-16
```




```
##
##  Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 130, df = 1, p-value <2e-16
```

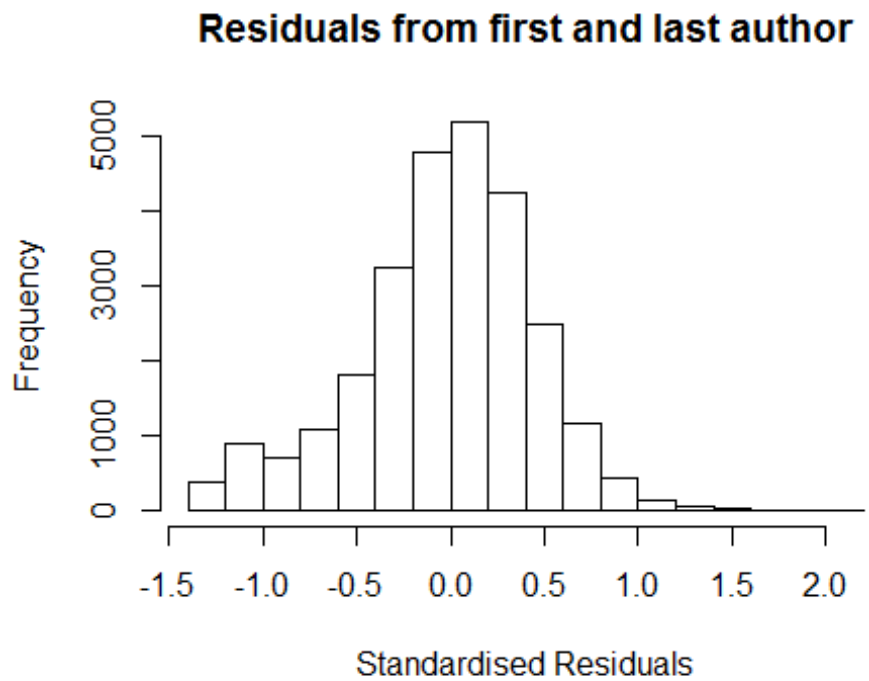



```
## [1] "Female first author team size 2018 geometric mean: 5.4654237735777"
## [1] "Male first author team size 2018 geometric mean: 4.82946958887189"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 850000, p-value = 2e-05
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 5.25392957856682"
## [1] "Male last author team size 2018 geometric mean: 5.07830535968098"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 740000, p-value = 0.3
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##                     GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.021  1           1.011
## LastAuthorFemale  1.010  1           1.005
## UniqueAuthors     1.037  4           1.005
## Year              1.054 16           1.002
```



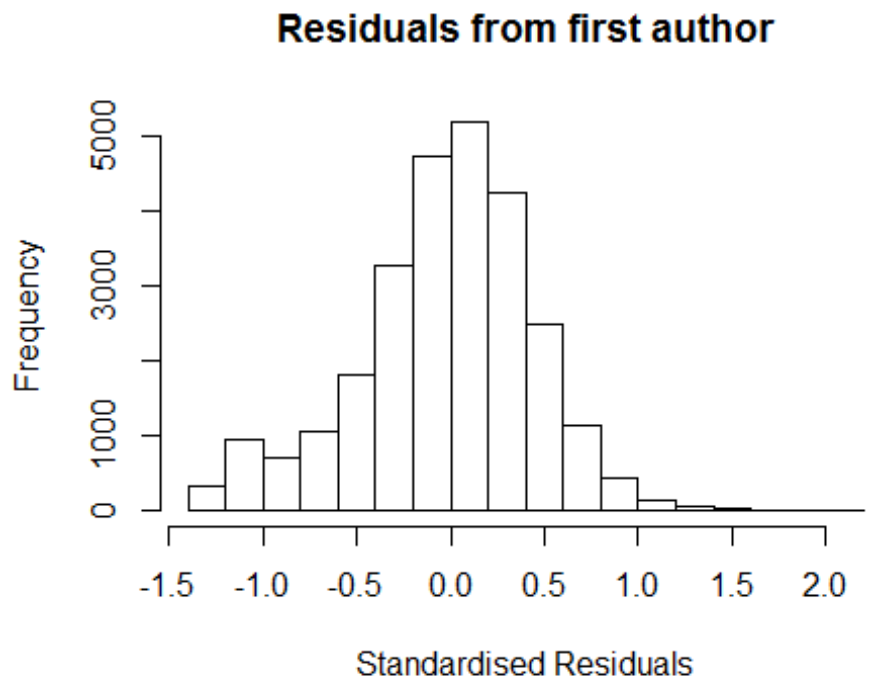

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
UniqueAuthors +
##     Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##     k.max = 1000))
##  \--> method = "MM"
## Residuals:
##    Min     1Q Median     3Q    Max
## -1.380 -0.273  0.010  0.274  2.138
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         0.71074    0.02358   30.14   <2e-16 ***
## FirstAuthorFemale1 -0.00597    0.00538   -1.11   0.2676
## LastAuthorFemale1  -0.01114    0.00602   -1.85   0.0642 .
## UniqueAuthors2      0.38271    0.01979   19.34   <2e-16 ***
## UniqueAuthors3      0.43834    0.01902   23.04   <2e-16 ***
## UniqueAuthors4      0.50300    0.01878   26.78   <2e-16 ***
## UniqueAuthors5      0.60943    0.01785   34.14   <2e-16 ***
## Year1997            0.06011    0.02375    2.53   0.0114 *
## Year1998            0.01071    0.02136    0.50   0.6160
## Year1999            0.00627    0.02079    0.30   0.7630
```

```
## Year2000            0.01955    0.02018    0.97   0.3327
## Year2001           -0.01696    0.02134   -0.79   0.4267
## Year2002           -0.02857    0.01910   -1.50   0.1347
## Year2003           -0.01350    0.01918   -0.70   0.4814
## Year2004           -0.13125    0.02012   -6.52    7e-11 ***
## Year2005           -0.05215    0.01913   -2.73   0.0064 **
## Year2006           -0.06103    0.01918   -3.18   0.0015 **
## Year2007           -0.05112    0.01910   -2.68   0.0074 **
## Year2008           -0.04158    0.01894   -2.20   0.0281 *
## Year2009           -0.04212    0.01881   -2.24   0.0252 *
## Year2010           -0.04028    0.01884   -2.14   0.0325 *
## Year2011           -0.01579    0.01864   -0.85   0.3969
## Year2012           -0.03981    0.01907   -2.09   0.0369 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.405
## Multiple R-squared:  0.134,  Adjusted R-squared:  0.134
## Convergence in 18 IRWLS iterations
##
## Robustness weights:
##  3 observations c(3507,16852,24077)
##   are outliers with |weight| = 0 ( < 3.8e-06);
##  2256 weights are ~= 1. The remaining 24296 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0155  0.8590  0.9500  0.8920  0.9860  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          3.77e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.028  1           1.014
## LastAuthorFemale  1.010  1           1.005
## Year              1.037 16           1.001
```



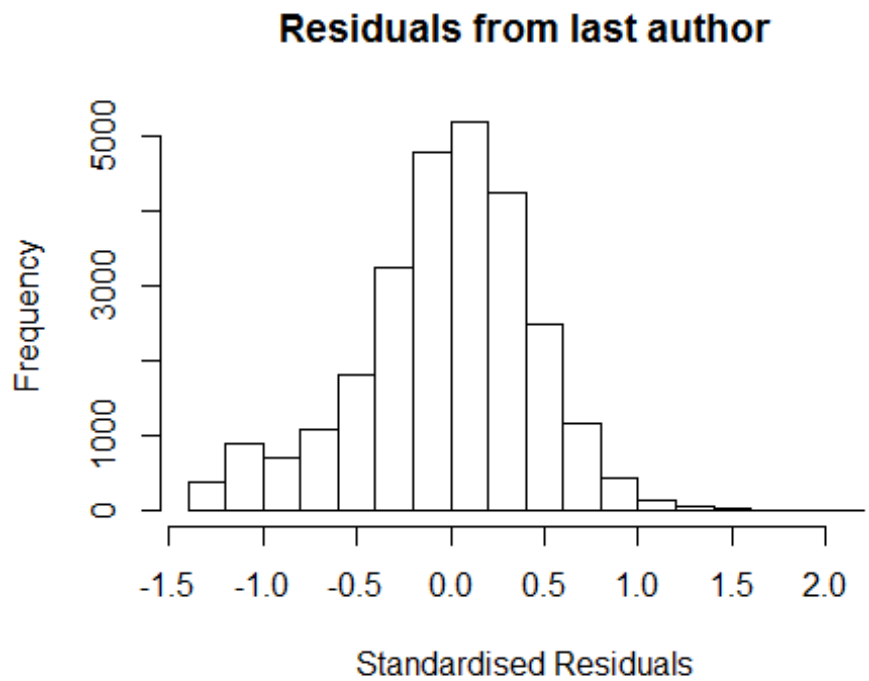

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.2427 -0.2808  0.0157  0.2796  2.0707
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)          1.16766    0.01782   65.51  < 2e-16 ***
## FirstAuthorFemale1   0.01557    0.00563    2.77  0.00567 **
## LastAuthorFemale1   -0.01467    0.00624   -2.35  0.01880 *
## Year1997             0.05945    0.02556    2.33  0.02002 *
## Year1998             0.01350    0.02298    0.59  0.55692
## Year1999             0.03379    0.02198    1.54  0.12420
## Year2000             0.05720    0.02137    2.68  0.00745 **
## Year2001             0.03762    0.02300    1.64  0.10186
## Year2002             0.02064    0.02036    1.01  0.31061
## Year2003             0.04046    0.02060    1.96  0.04957 *
## Year2004            -0.07921    0.02353   -3.37  0.00076 ***
## Year2005             0.01444    0.02051    0.70  0.48156
```

```
## Year2006           -0.00559    0.02042   -0.27  0.78423
## Year2007            0.01195    0.02036    0.59  0.55729
## Year2008            0.01929    0.02031    0.95  0.34210
## Year2009            0.01406    0.02009    0.70  0.48394
## Year2010            0.02041    0.02014    1.01  0.31101
## Year2011            0.04169    0.01988    2.10  0.03597 *
## Year2012            0.02762    0.02036    1.36  0.17494
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.412
## Multiple R-squared:  0.00472,    Adjusted R-squared:  0.00405
## Convergence in 15 IRWLS iterations
##
## Robustness weights:
##  2 observations c(16852,24077) are outliers with |weight| = 0 ( < 3.8e-
06);
##  2258 weights are ~= 1. The remaining 24295 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0159  0.8580  0.9490  0.8850  0.9860  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          3.77e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.029  1           1.014
## Year              1.029 16           1.001
```




```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.2384 -0.2812  0.0156  0.2797  2.0613
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)          1.16467    0.01775   65.61  < 2e-16 ***
## FirstAuthorFemale1   0.01428    0.00565    2.53  0.01142 *
## Year1997             0.05946    0.02555    2.33  0.01996 *
## Year1998             0.01362    0.02298    0.59  0.55329
## Year1999             0.03380    0.02198    1.54  0.12422
## Year2000             0.05702    0.02138    2.67  0.00766 **
## Year2001             0.03747    0.02299    1.63  0.10318
## Year2002             0.02046    0.02036    1.01  0.31483
## Year2003             0.04043    0.02060    1.96  0.04968 *
## Year2004            -0.07967    0.02353   -3.39  0.00071 ***
## Year2005             0.01442    0.02051    0.70  0.48196
## Year2006            -0.00602    0.02043   -0.29  0.76829
```

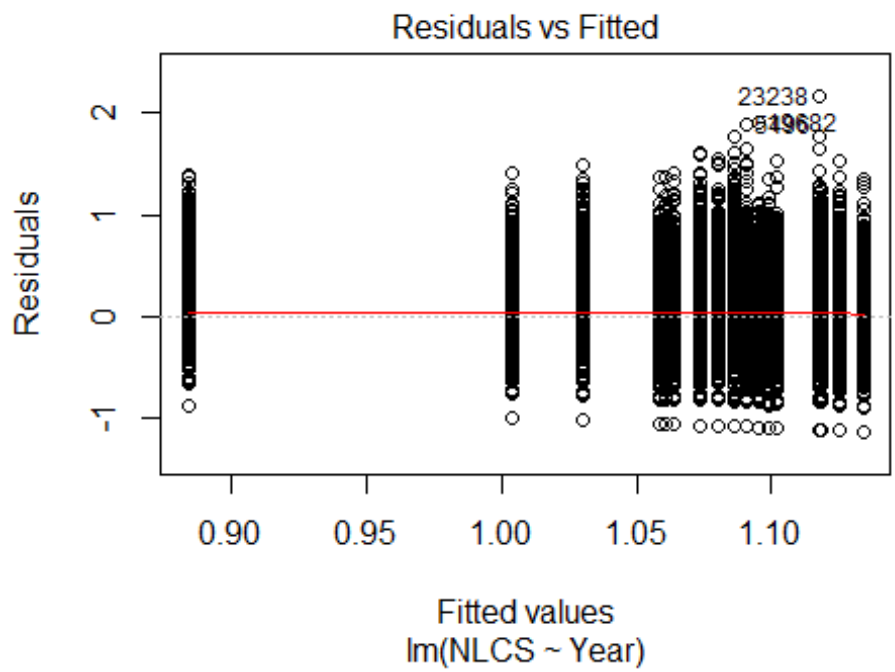
```
## Year2007            0.01126    0.02036    0.55  0.58025
## Year2008            0.01872    0.02031    0.92  0.35660
## Year2009            0.01368    0.02009    0.68  0.49592
## Year2010            0.01974    0.02014    0.98  0.32708
## Year2011            0.04075    0.01987    2.05  0.04029 *
## Year2012            0.02664    0.02036    1.31  0.19064
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.412
## Multiple R-squared:  0.00455,    Adjusted R-squared:  0.00391
## Convergence in 15 IRWLS iterations
##
## Robustness weights:
##  2 observations c(16852,24077) are outliers with |weight| = 0 ( < 3.8e-
06);
##  2221 weights are ~= 1. The remaining 24332 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0149  0.8580  0.9490  0.8850  0.9860  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          3.77e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##                   GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.011  1           1.005
## Year             1.011 16           1.000
```



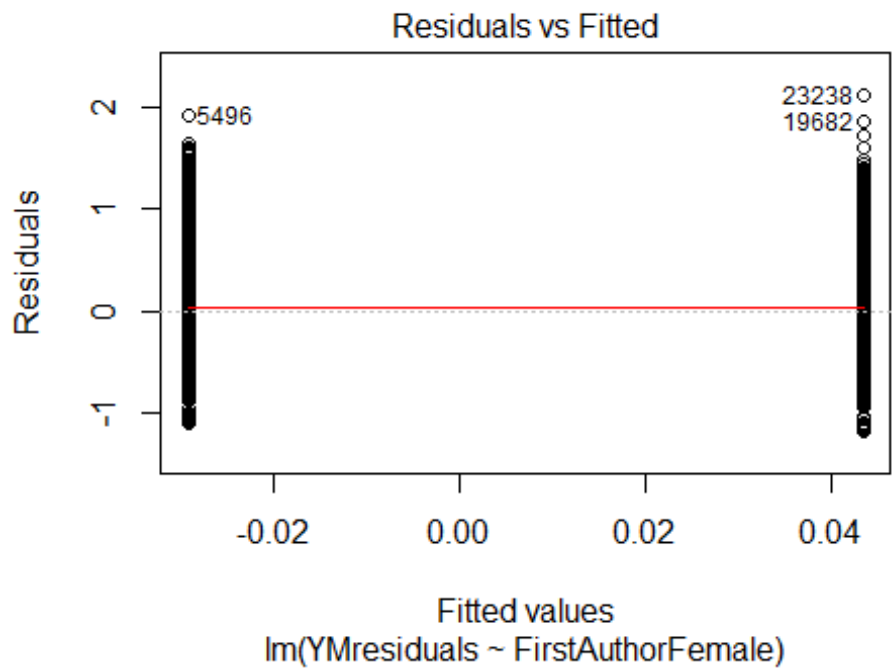

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.2325 -0.2809  0.0157  0.2795  2.0775
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.17226    0.01773   66.14   <2e-16 ***
## LastAuthorFemale1 -0.01295    0.00626   -2.07   0.0386 *
## Year1997            0.06026    0.02557    2.36   0.0184 *
## Year1998            0.01409    0.02298    0.61   0.5397
## Year1999            0.03436    0.02198    1.56   0.1181
## Year2000            0.05801    0.02138    2.71   0.0067 **
## Year2001            0.03846    0.02299    1.67   0.0944 .
## Year2002            0.02208    0.02035    1.08   0.2781
## Year2003            0.04147    0.02061    2.01   0.0442 *
## Year2004           -0.07685    0.02345   -3.28   0.0011 **
## Year2005            0.01615    0.02051    0.79   0.4309
## Year2006           -0.00390    0.02042   -0.19   0.8484
```

```
## Year2007           0.01364    0.02035    0.67   0.5028
## Year2008           0.02152    0.02029    1.06   0.2888
## Year2009           0.01622    0.02007    0.81   0.4192
## Year2010           0.02287    0.02013    1.14   0.2559
## Year2011           0.04416    0.01986    2.22   0.0262 *
## Year2012           0.03051    0.02034    1.50   0.1335
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.412
## Multiple R-squared:  0.00437,    Adjusted R-squared:  0.00373
## Convergence in 15 IRWLS iterations
##
## Robustness weights:
##  2 observations c(16852,24077) are outliers with |weight| = 0 ( < 3.8e-
06);
##  2242 weights are ~= 1. The remaining 24311 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0138  0.8580  0.9490  0.8850  0.9860  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          3.77e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                  psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Sample size for the above analysis:  26555"
## [1] ""
## [1] ""
## [1] "##################################"
## [1] "Analysis of AJSC 2726"
## [1] "##################################"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 1226 1349 1195 1135 1216 1328 1331 1027 1356 1109 1111 1213 1178 1312 1302
## 2011 2012
## 1628 1471
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
```

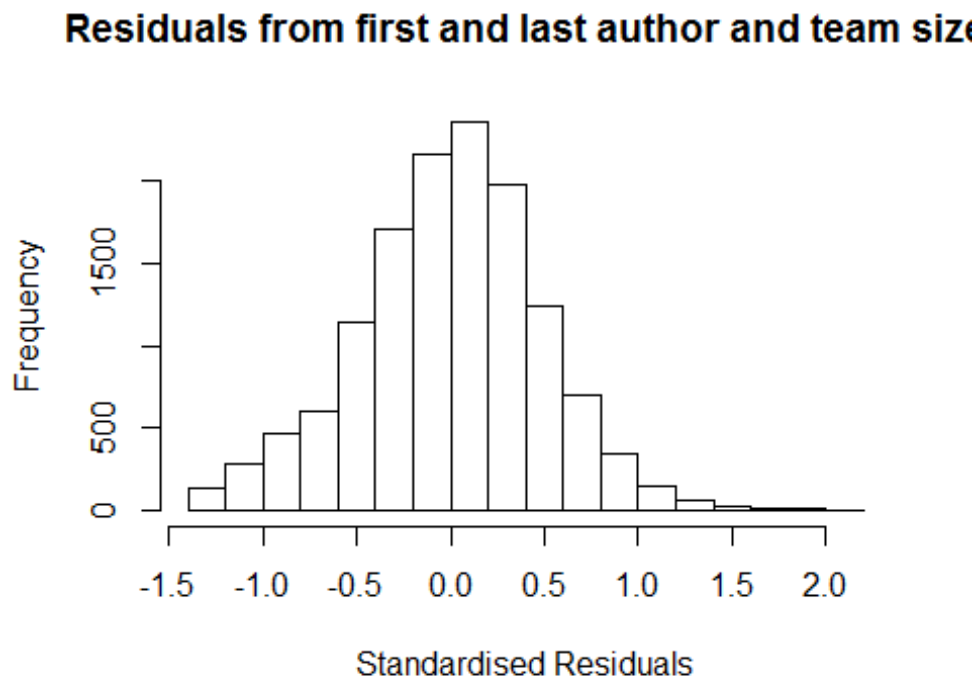
```
##  753  687  718  774  723  588  994  794 1070  864  841  923  873  988  966
## 2011 2012
## 1246 1135
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  690  625  652  709  662  541  899  711  974  781  738  826  769  868  836
## 2011 2012
## 1095  990
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
##  Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 380, df = 16, p-value <2e-16
```




```
##
##  Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 64, df = 1, p-value = 1e-15
```

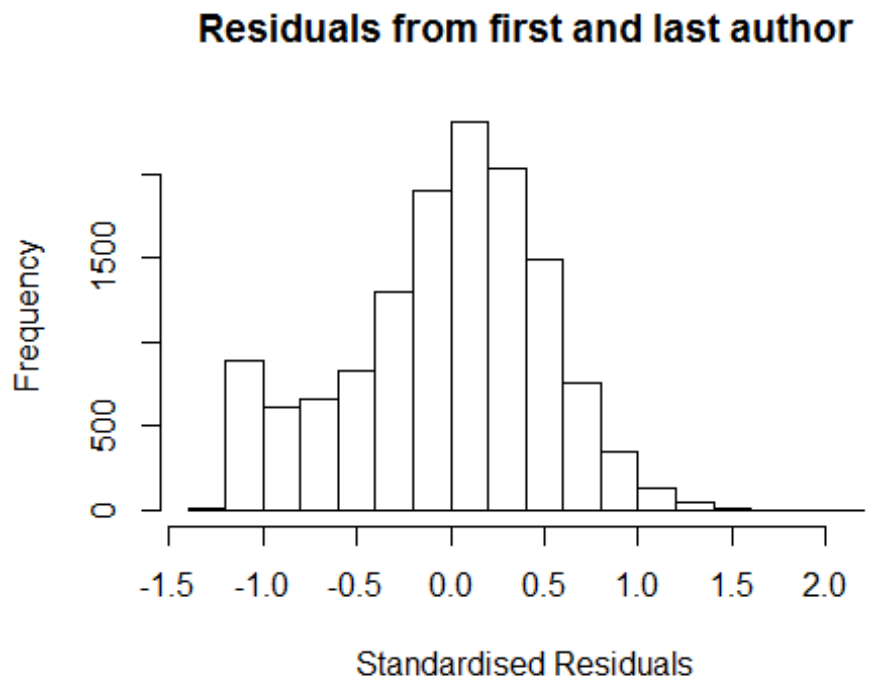



```
## [1] "Female first author team size 2018 geometric mean: 5.39140332661492"
## [1] "Male first author team size 2018 geometric mean: 4.66210883365903"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 160000, p-value = 6e-04
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 5.10619006270211"
## [1] "Male last author team size 2018 geometric mean: 4.93077762758731"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 130000, p-value = 0.3
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.043  1           1.021
## LastAuthorFemale  1.021  1           1.011
## UniqueAuthors     1.174  4           1.020
## Year              1.172 16           1.005
```



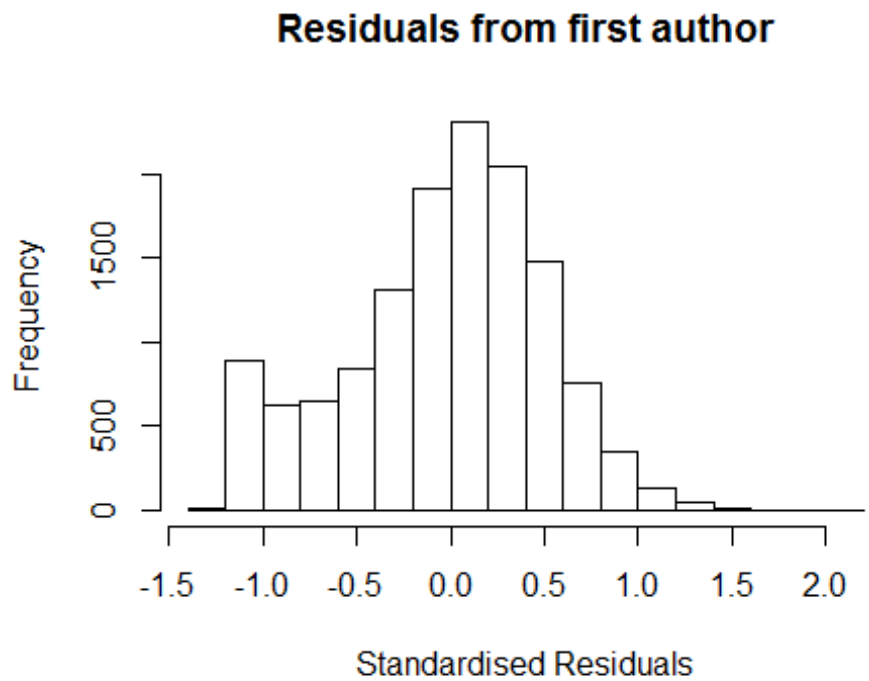

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
UniqueAuthors +
##     Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##     k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.3726 -0.3217  0.0137  0.3080  2.1366
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)          0.60031    0.02935   20.46  < 2e-16 ***
## FirstAuthorFemale1   0.02394    0.00879    2.72  0.00647 **
## LastAuthorFemale1    0.01470    0.00942    1.56  0.11869
## UniqueAuthors2       0.40024    0.02428   16.48  < 2e-16 ***
## UniqueAuthors3       0.48475    0.02308   21.00  < 2e-16 ***
## UniqueAuthors4       0.60426    0.02213   27.30  < 2e-16 ***
## UniqueAuthors5       0.74070    0.01977   37.46  < 2e-16 ***
## Year1997             0.00764    0.03681    0.21  0.83563
## Year1998            -0.06294    0.03215   -1.96  0.05029 .
## Year1999            -0.03458    0.03141   -1.10  0.27094
```

```
## Year2000           -0.04848    0.03087   -1.57  0.11636
## Year2001           -0.13759    0.03272   -4.21  2.6e-05 ***
## Year2002           -0.05909    0.02859   -2.07  0.03876 *
## Year2003           -0.06045    0.02808   -2.15  0.03136 *
## Year2004           -0.21941    0.02842   -7.72  1.2e-14 ***
## Year2005           -0.14077    0.02941   -4.79  1.7e-06 ***
## Year2006           -0.11204    0.02992   -3.75  0.00018 ***
## Year2007           -0.10302    0.02869   -3.59  0.00033 ***
## Year2008           -0.10295    0.02881   -3.57  0.00035 ***
## Year2009           -0.12020    0.02850   -4.22  2.5e-05 ***
## Year2010           -0.11682    0.02901   -4.03  5.7e-05 ***
## Year2011           -0.08312    0.02857   -2.91  0.00362 **
## Year2012           -0.10290    0.02856   -3.60  0.00032 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.435
## Multiple R-squared:  0.205,  Adjusted R-squared:  0.204
## Convergence in 15 IRWLS iterations
##
## Robustness weights:
##  observation 2086 is an outlier with |weight| = 0 ( < 7.5e-06);
##  1105 weights are ~= 1. The remaining 12260 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0028  0.8520  0.9420  0.8860  0.9840  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          7.48e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.024  1           1.012
## LastAuthorFemale  1.007  1           1.003
## Year              1.030 16           1.001
```



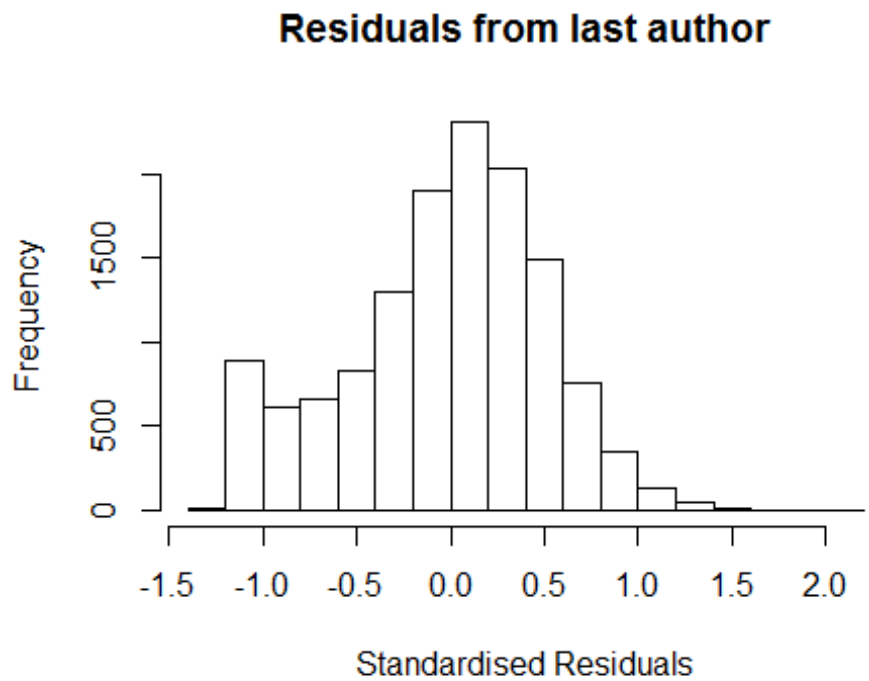

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.2180 -0.3417  0.0411  0.3451  2.0857
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)          1.06309    0.02740   38.80  < 2e-16 ***
## FirstAuthorFemale1   0.07856    0.00947    8.30  < 2e-16 ***
## LastAuthorFemale1    0.02344    0.01002    2.34    0.019 *
## Year1997             0.02884    0.04158    0.69    0.488
## Year1998            -0.04315    0.03616   -1.19    0.233
## Year1999             0.02734    0.03488    0.78    0.433
## Year2000            -0.01491    0.03457   -0.43    0.666
## Year2001            -0.07638    0.03854   -1.98    0.048 *
## Year2002             0.02383    0.03118    0.76    0.445
## Year2003             0.05288    0.03126    1.69    0.091 .
## Year2004            -0.21215    0.03757   -5.65  1.7e-08 ***
## Year2005            -0.02073    0.03288   -0.63    0.528
```

```
## Year2006           -0.02189    0.03304   -0.66    0.508
## Year2007            0.01724    0.03183    0.54    0.588
## Year2008            0.04180    0.03180    1.31    0.189
## Year2009            0.00328    0.03179    0.10    0.918
## Year2010           -0.00630    0.03222   -0.20    0.845
## Year2011            0.03021    0.03111    0.97    0.331
## Year2012            0.02984    0.03140    0.95    0.342
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.481
## Multiple R-squared:  0.0214, Adjusted R-squared:  0.0201
## Convergence in 15 IRWLS iterations
##
## Robustness weights:
##  1104 weights are ~= 1. The remaining 12262 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0205  0.8440  0.9440  0.8900  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          7.48e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.024  1           1.012
## Year              1.024 16           1.001
```




```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.2023 -0.3386  0.0413  0.3449  2.1004
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)          1.06811    0.02727   39.16  < 2e-16 ***
## FirstAuthorFemale1   0.08103    0.00951    8.52  < 2e-16 ***
## Year1997             0.02951    0.04159    0.71    0.478
## Year1998            -0.04314    0.03618   -1.19    0.233
## Year1999             0.02769    0.03485    0.79    0.427
## Year2000            -0.01480    0.03458   -0.43    0.669
## Year2001            -0.07645    0.03856   -1.98    0.047 *
## Year2002             0.02454    0.03119    0.79    0.432
## Year2003             0.05320    0.03129    1.70    0.089 .
## Year2004            -0.21257    0.03764   -5.65  1.7e-08 ***
## Year2005            -0.02077    0.03288   -0.63    0.528
## Year2006            -0.02010    0.03304   -0.61    0.543
```

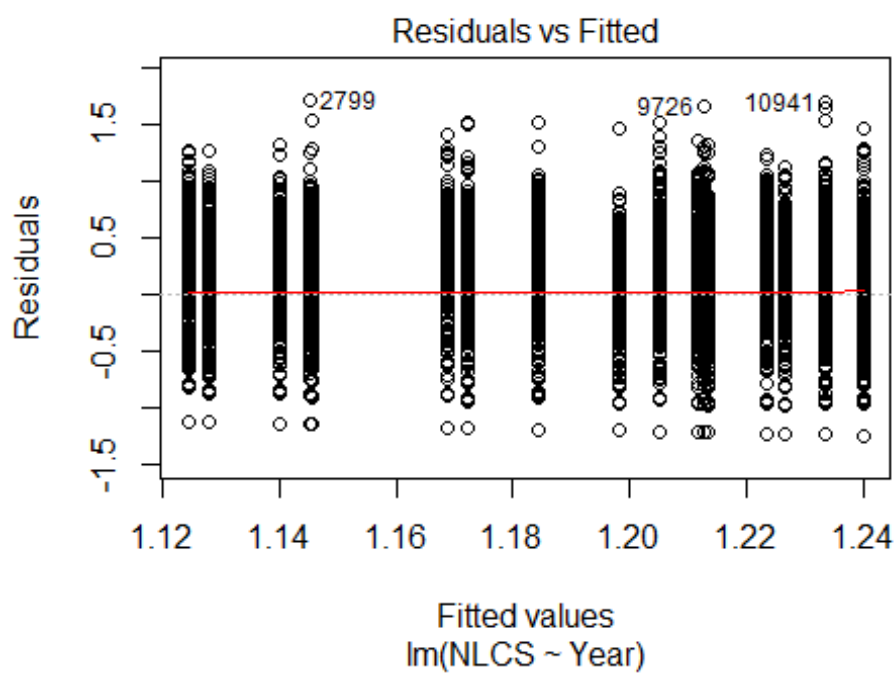
```
## Year2007            0.01854    0.03186    0.58    0.561
## Year2008            0.04340    0.03181    1.36    0.172
## Year2009            0.00405    0.03179    0.13    0.899
## Year2010           -0.00553    0.03222   -0.17    0.864
## Year2011            0.03144    0.03111    1.01    0.312
## Year2012            0.03062    0.03141    0.97    0.330
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.481
## Multiple R-squared:  0.0209, Adjusted R-squared:  0.0196
## Convergence in 15 IRWLS iterations
##
## Robustness weights:
##  1123 weights are ~= 1. The remaining 12243 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0172  0.8430  0.9440  0.8890  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          7.48e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##                   GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.008  1           1.004
## Year             1.008 16           1.000
```



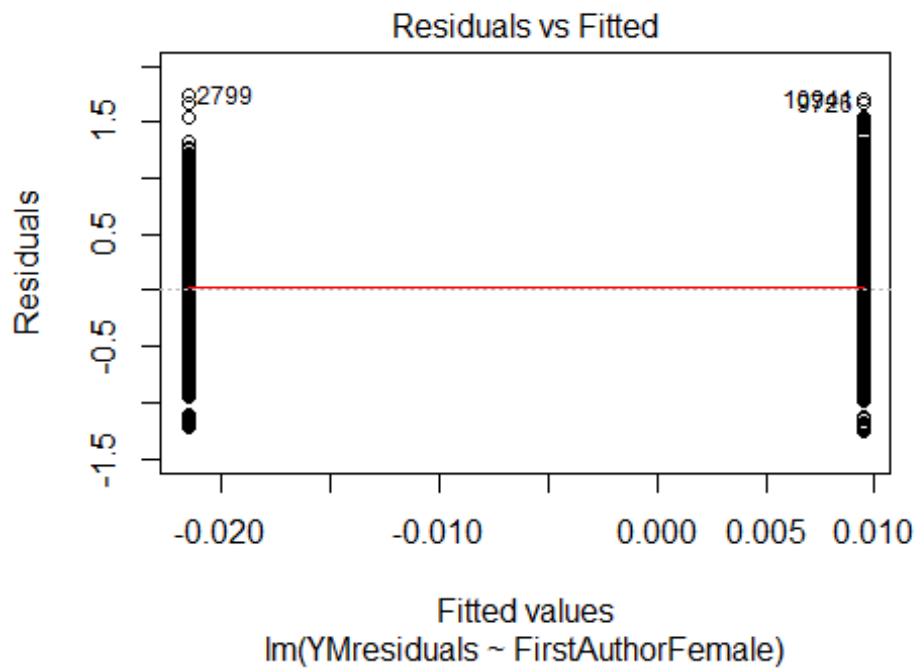

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.1786 -0.3424  0.0421  0.3423  2.1192
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.08482    0.02739   39.60  < 2e-16 ***
## LastAuthorFemale1   0.03325    0.01007    3.30  0.00096 ***
## Year1997            0.02982    0.04178    0.71  0.47548
## Year1998           -0.03872    0.03640   -1.06  0.28745
## Year1999            0.03002    0.03510    0.86  0.39230
## Year2000           -0.01110    0.03478   -0.32  0.74971
## Year2001           -0.07075    0.03867   -1.83  0.06735 .
## Year2002            0.02982    0.03129    0.95  0.34066
## Year2003            0.06049    0.03139    1.93  0.05398 .
## Year2004           -0.20447    0.03813   -5.36  8.3e-08 ***
## Year2005           -0.00994    0.03302   -0.30  0.76328
## Year2006           -0.01393    0.03319   -0.42  0.67468
```

```
## Year2007           0.02678    0.03190    0.84  0.40130
## Year2008           0.05309    0.03180    1.67  0.09504 .
## Year2009           0.01389    0.03191    0.44  0.66325
## Year2010           0.00518    0.03230    0.16  0.87255
## Year2011           0.04369    0.03117    1.40  0.16098
## Year2012           0.04315    0.03145    1.37  0.17008
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.481
## Multiple R-squared:  0.0157, Adjusted R-squared:  0.0145
## Convergence in 15 IRWLS iterations
##
## Robustness weights:
##  1126 weights are ~= 1. The remaining 12240 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0133  0.8400  0.9440  0.8890  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          7.48e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Sample size for the above analysis:  13366"
## [1] ""
## [1] ""
## [1] "####################################"
## [1] "Analysis of AJSC 2727"
## [1] "####################################"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  699  631  655  737  779  841  744  491  486  513  753  771  892  923  828
## 2011 2012
##  885  864
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  411  393  361  405  471  454  543  351  365  389  470  525  639  719  622
## 2011 2012
```

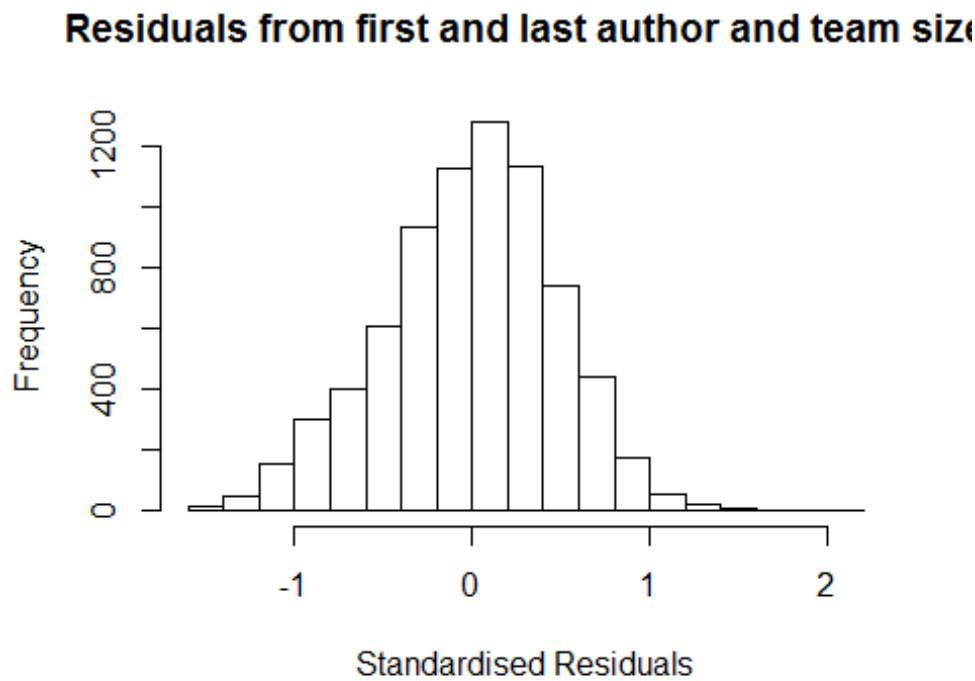
```
##  665  656
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  370  342  318  362  424  405  476  297  319  335  398  477  571  640  539
## 2011 2012
##  570  568
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
##  Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 81, df = 16, p-value = 1e-10
```




```
##
##  Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 1.7, df = 1, p-value = 0.2
```

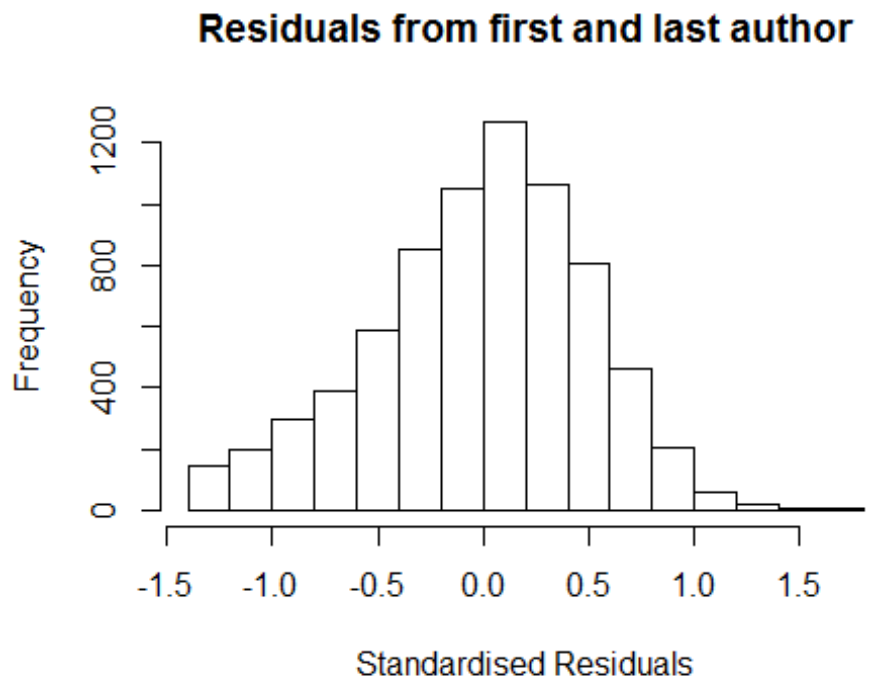



```
## [1] "Female first author team size 2018 geometric mean: 5.40892501034255"
## [1] "Male first author team size 2018 geometric mean: 4.94570670606375"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 27000, p-value = 0.3
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 4.8430962500231"
## [1] "Male last author team size 2018 geometric mean: 5.24399119658218"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 19000, p-value = 0.3
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##                     GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.043  1           1.021
## LastAuthorFemale  1.039  1           1.019
## UniqueAuthors     1.091  4           1.011
## Year              1.102 16           1.003
```



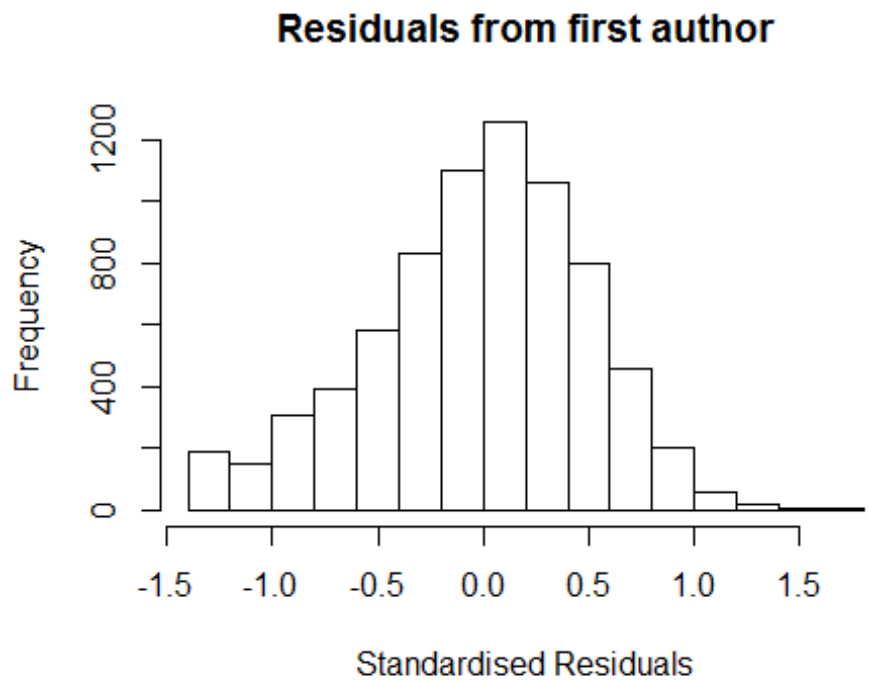

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
UniqueAuthors +
##     Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##     k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.5068 -0.3201  0.0212  0.3182  2.0254
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         0.99297    0.03396   29.24  < 2e-16 ***
## FirstAuthorFemale1 -0.03552    0.01265   -2.81  0.00500 **
## LastAuthorFemale1  -0.06845    0.01437   -4.76  1.9e-06 ***
## UniqueAuthors2      0.18335    0.02388    7.68  1.8e-14 ***
## UniqueAuthors3      0.23688    0.02360   10.04  < 2e-16 ***
## UniqueAuthors4      0.35288    0.02297   15.36  < 2e-16 ***
## UniqueAuthors5      0.51380    0.01935   26.56  < 2e-16 ***
## Year1997           -0.03617    0.04252   -0.85  0.39502
## Year1998           -0.06706    0.04323   -1.55  0.12094
## Year1999           -0.11349    0.04142   -2.74  0.00617 **
```

```
## Year2000           -0.03997    0.03751   -1.07  0.28663
## Year2001           -0.00506    0.03712   -0.14  0.89168
## Year2002           -0.11163    0.03645   -3.06  0.00220 **
## Year2003           -0.09840    0.03892   -2.53  0.01149 *
## Year2004           -0.13221    0.03869   -3.42  0.00064 ***
## Year2005           -0.11605    0.03914   -2.97  0.00304 **
## Year2006           -0.17587    0.03890   -4.52  6.3e-06 ***
## Year2007           -0.03649    0.03762   -0.97  0.33220
## Year2008           -0.06397    0.03743   -1.71  0.08748 .
## Year2009           -0.04139    0.03650   -1.13  0.25678
## Year2010           -0.07102    0.03759   -1.89  0.05891 .
## Year2011           -0.13033    0.03767   -3.46  0.00054 ***
## Year2012           -0.18220    0.03831   -4.76  2.0e-06 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.471
## Multiple R-squared:  0.14,   Adjusted R-squared:  0.138
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
##  590 weights are ~= 1. The remaining 6821 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.025   0.867   0.951   0.904   0.985   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.35e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.018  1           1.009
## LastAuthorFemale  1.024  1           1.012
## Year              1.018 16           1.001
```



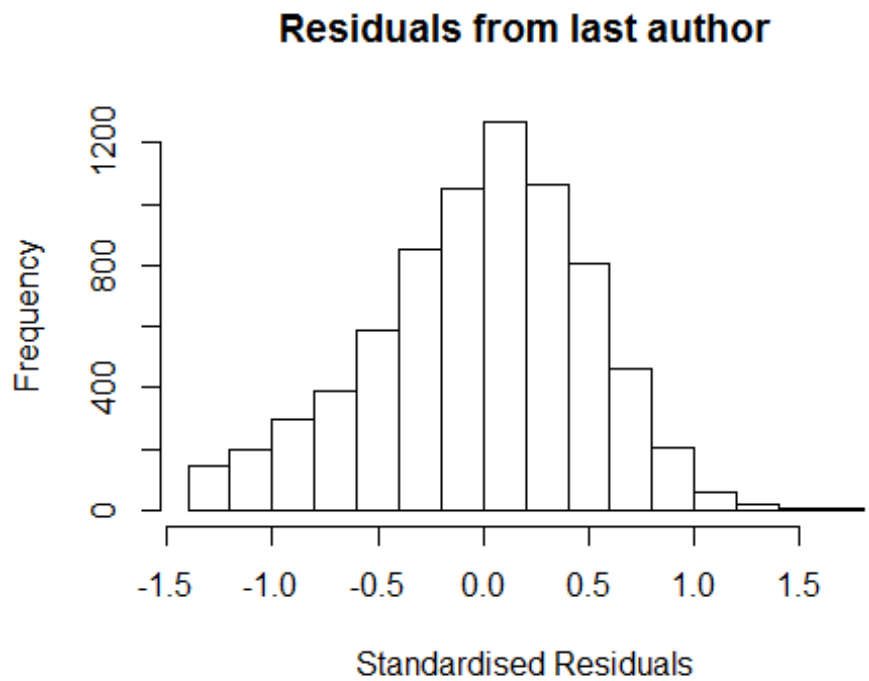

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.2842 -0.3380  0.0303  0.3423  1.7450
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.24479    0.03280   37.95  < 2e-16 ***
## FirstAuthorFemale1 -0.01494    0.01345   -1.11    0.267
## LastAuthorFemale1  -0.09123    0.01577   -5.78  7.6e-09 ***
## Year1997           -0.01672    0.04595   -0.36    0.716
## Year1998           -0.04559    0.04576   -1.00    0.319
## Year1999           -0.06914    0.04356   -1.59    0.113
## Year2000            0.00289    0.04029    0.07    0.943
## Year2001            0.03092    0.03975    0.78    0.437
## Year2002            0.00186    0.03912    0.05    0.962
## Year2003           -0.00209    0.04061   -0.05    0.959
## Year2004           -0.03708    0.04069   -0.91    0.362
## Year2005           -0.03437    0.04228   -0.81    0.416
```

```
## Year2006           -0.08660    0.04129   -2.10    0.036 *
## Year2007            0.03503    0.03979    0.88    0.379
## Year2008            0.02111    0.03954    0.53    0.593
## Year2009            0.03946    0.03920    1.01    0.314
## Year2010            0.03221    0.04029    0.80    0.424
## Year2011           -0.02247    0.04040   -0.56    0.578
## Year2012           -0.07620    0.04065   -1.87    0.061 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.496
## Multiple R-squared:  0.0117, Adjusted R-squared:  0.00929
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
##  594 weights are ~= 1. The remaining 6817 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.189   0.865   0.949   0.901   0.985   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.35e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.006  1           1.003
## Year              1.006 16           1.000
```




```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.2716 -0.3375  0.0301  0.3444  1.7281
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.22870    0.03247   37.84   <2e-16 ***
## FirstAuthorFemale1 -0.03197    0.01358   -2.35    0.019 *
## Year1997           -0.01815    0.04571   -0.40    0.691
## Year1998           -0.04266    0.04556   -0.94    0.349
## Year1999           -0.06488    0.04334   -1.50    0.134
## Year2000            0.00334    0.04014    0.08    0.934
## Year2001            0.03091    0.03965    0.78    0.436
## Year2002            0.00636    0.03890    0.16    0.870
## Year2003            0.00046    0.04062    0.01    0.991
## Year2004           -0.03711    0.04056   -0.92    0.360
## Year2005           -0.02867    0.04218   -0.68    0.497
## Year2006           -0.08293    0.04114   -2.02    0.044 *
```

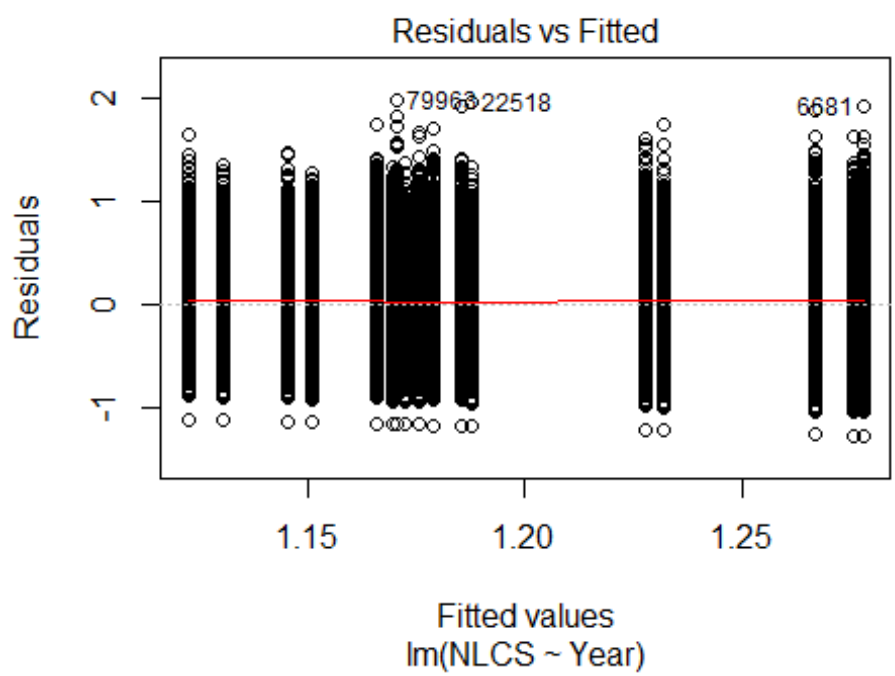
```
## Year2007            0.03642    0.03977    0.92    0.360
## Year2008            0.02174    0.03948    0.55    0.582
## Year2009            0.04289    0.03908    1.10    0.272
## Year2010            0.03441    0.04025    0.85    0.393
## Year2011           -0.01950    0.04035   -0.48    0.629
## Year2012           -0.07147    0.04059   -1.76    0.078 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.496
## Multiple R-squared:  0.00667,    Adjusted R-squared:  0.00438
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
##  609 weights are ~= 1. The remaining 6802 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.199   0.865   0.949   0.900   0.986   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.35e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##                   GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.012  1           1.006
## Year             1.012 16           1.000
```



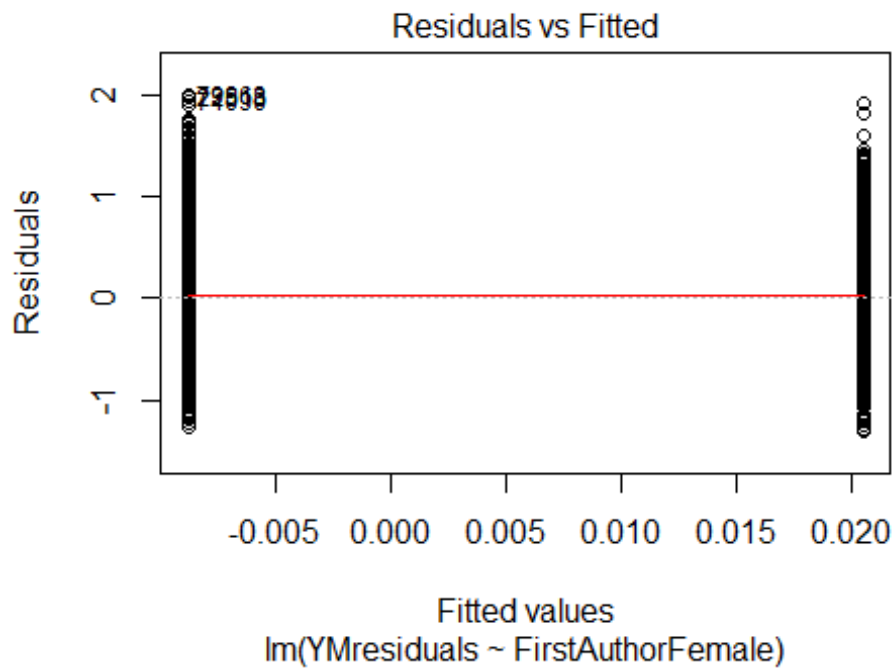

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.2805 -0.3345  0.0313  0.3438  1.7524
##
## Coefficients:
##                   Estimate Std. Error t value Pr(>|t|)
## (Intercept)        1.24171    0.03269   37.98   <2e-16 ***
## LastAuthorFemale1 -0.09483    0.01578   -6.01    2e-09 ***
## Year1997          -0.01602    0.04594   -0.35    0.727
## Year1998          -0.04575    0.04575   -1.00    0.317
## Year1999          -0.06932    0.04357   -1.59    0.112
## Year2000           0.00284    0.04029    0.07    0.944
## Year2001           0.03064    0.03973    0.77    0.441
## Year2002           0.00204    0.03914    0.05    0.959
## Year2003          -0.00335    0.04064   -0.08    0.934
## Year2004          -0.03744    0.04070   -0.92    0.358
## Year2005          -0.03559    0.04225   -0.84    0.400
## Year2006          -0.08738    0.04128   -2.12    0.034 *
```

```
## Year2007           0.03437    0.03977    0.86    0.388
## Year2008           0.02002    0.03954    0.51    0.613
## Year2009           0.03875    0.03921    0.99    0.323
## Year2010           0.03073    0.04028    0.76    0.445
## Year2011          -0.02422    0.04038   -0.60    0.549
## Year2012          -0.07773    0.04062   -1.91    0.056 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.496
## Multiple R-squared:  0.0115, Adjusted R-squared:  0.00922
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
##  613 weights are ~= 1. The remaining 6798 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.185   0.864   0.949   0.901   0.985   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.35e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Sample size for the above analysis:  7411"
## [1] ""
## [1] ""
## [1] "####################################"
## [1] "Analysis of AJSC 2728"
## [1] "####################################"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 4077 3925 3866 3597 3967 4098 3648 3265 3437 3952 4425 4781 5081 5454 5526
## 2011 2012
## 6128 5798
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 2734 2631 2592 2472 2367 2125 2633 2468 2556 2951 3107 3565 3783 4025 4130
## 2011 2012
```

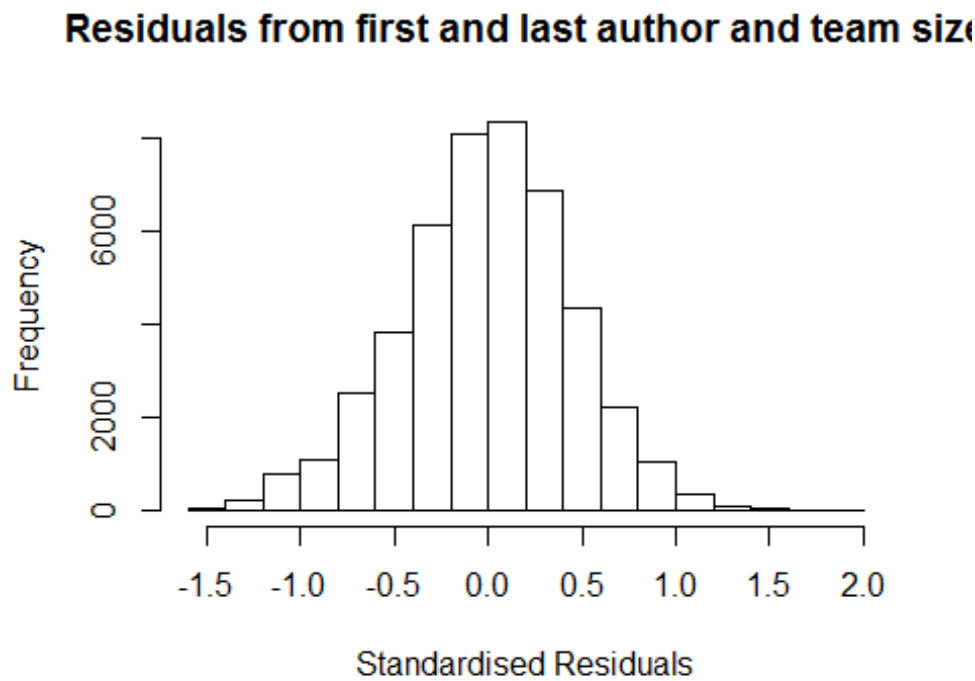
```
## 4564 4438
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 2427 2284 2229 2089 2021 1828 2256 2071 2187 2488 2665 3080 3286 3473 3607
## 2011 2012
## 3974 3900
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
##  Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 170, df = 16, p-value <2e-16
```




```
##
##  Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 140, df = 1, p-value <2e-16
```

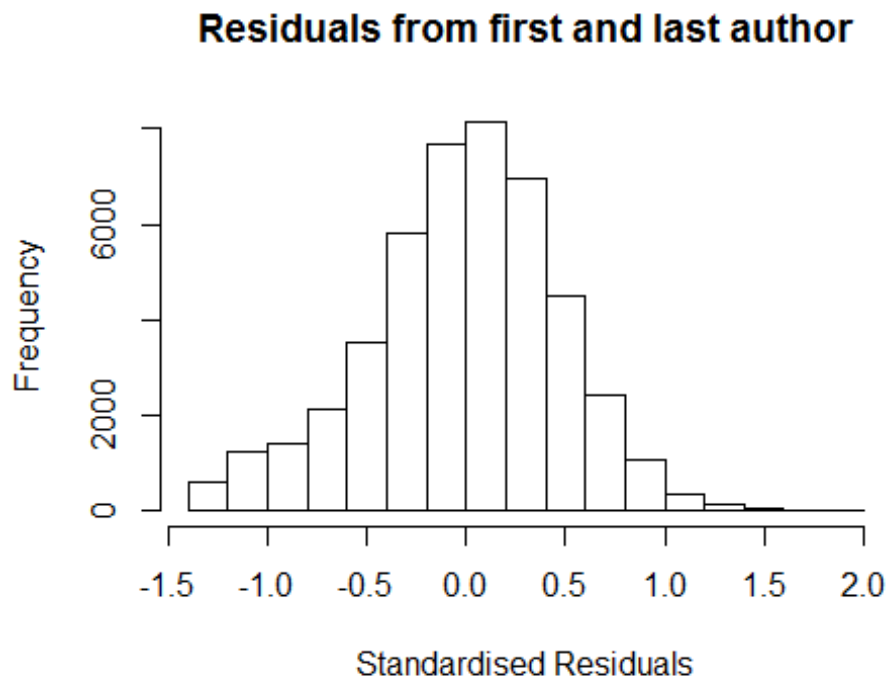



```
## [1] "Female first author team size 2018 geometric mean: 5.24414653710737"
## [1] "Male first author team size 2018 geometric mean: 5.19328623769271"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 2800000, p-value = 0.4
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 5.1388086497858"
## [1] "Male last author team size 2018 geometric mean: 5.23825034845302"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 2200000, p-value = 0.2
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##                     GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.029  1           1.015
## LastAuthorFemale  1.029  1           1.014
## UniqueAuthors     1.051  4           1.006
## Year              1.058 16           1.002
```



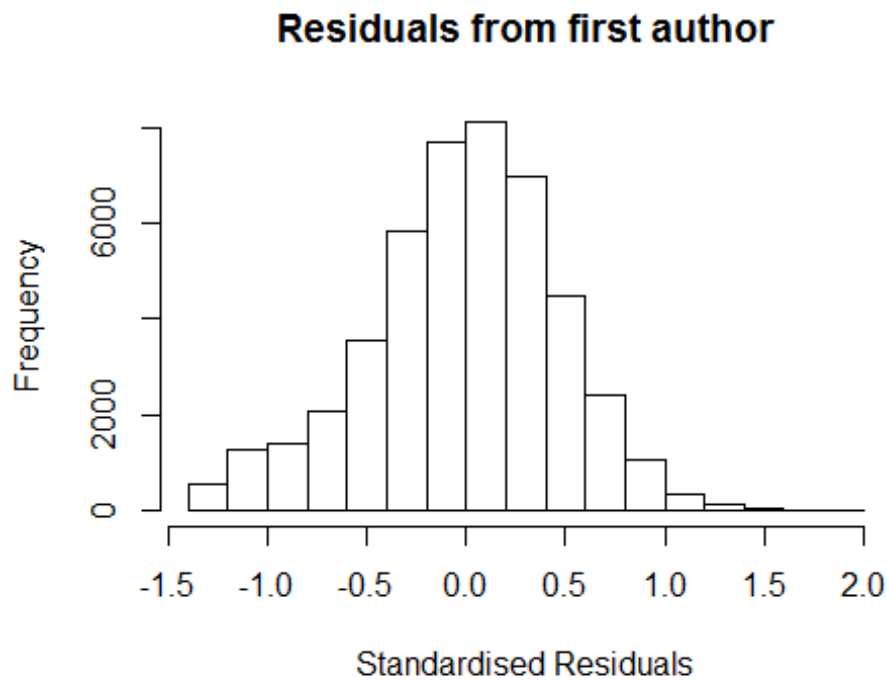

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
UniqueAuthors +
##     Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##     k.max = 1000))
##  \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.47555 -0.29114  0.00745  0.29465  1.92603
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         0.85997    0.01387   62.00  < 2e-16 ***
## FirstAuthorFemale1  0.01665    0.00457    3.65  0.00027 ***
## LastAuthorFemale1   0.01476    0.00518    2.85  0.00438 **
## UniqueAuthors2      0.27188    0.01176   23.12  < 2e-16 ***
## UniqueAuthors3      0.35217    0.01123   31.37  < 2e-16 ***
## UniqueAuthors4      0.42183    0.01119   37.70  < 2e-16 ***
## UniqueAuthors5      0.55447    0.01056   52.53  < 2e-16 ***
## Year1997            0.04446    0.01468    3.03  0.00246 **
## Year1998            0.03120    0.01411    2.21  0.02699 *
## Year1999            0.02232    0.01412    1.58  0.11395
```

```
## Year2000           -0.02221    0.01414   -1.57  0.11614
## Year2001           -0.05853    0.01408   -4.16  3.2e-05 ***
## Year2002           -0.08198    0.01354   -6.05  1.4e-09 ***
## Year2003           -0.13239    0.01388   -9.54  < 2e-16 ***
## Year2004           -0.11809    0.01335   -8.85  < 2e-16 ***
## Year2005           -0.10400    0.01310   -7.94  2.1e-15 ***
## Year2006           -0.15168    0.01318  -11.50  < 2e-16 ***
## Year2007           -0.12513    0.01257   -9.96  < 2e-16 ***
## Year2008           -0.10910    0.01263   -8.64  < 2e-16 ***
## Year2009           -0.10074    0.01259   -8.00  1.3e-15 ***
## Year2010           -0.10679    0.01265   -8.44  < 2e-16 ***
## Year2011           -0.12562    0.01256  -10.00  < 2e-16 ***
## Year2012           -0.12288    0.01261   -9.74  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.434
## Multiple R-squared:  0.13,   Adjusted R-squared:  0.13
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
##  3909 weights are ~= 1. The remaining 41956 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0105  0.8660  0.9500  0.9000  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          2.18e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.021  1           1.011
## LastAuthorFemale  1.019  1           1.009
## Year              1.013 16           1.000
```



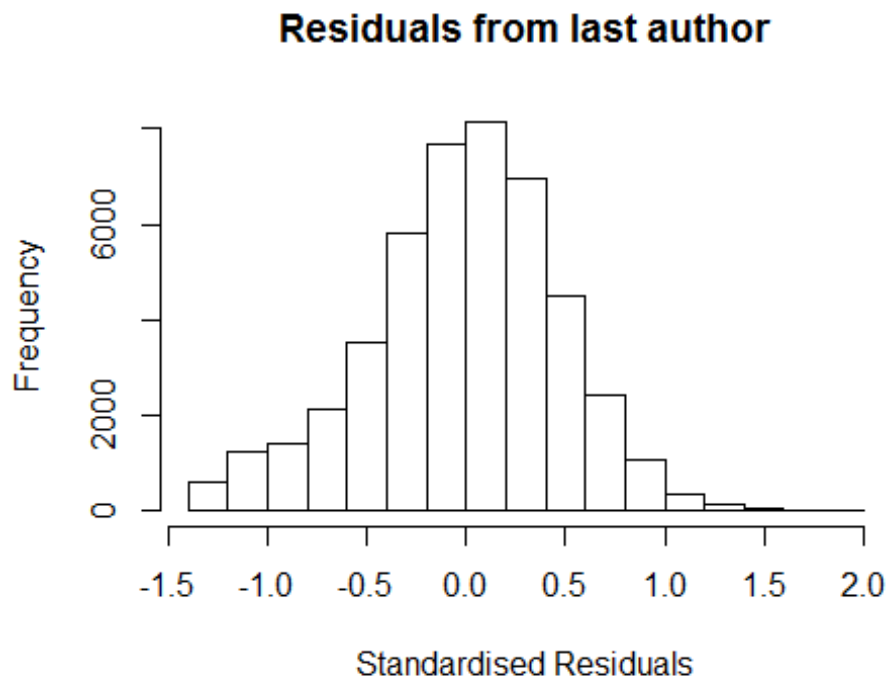

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.3156 -0.3009  0.0135  0.3041  1.9764
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)          1.23689    0.01051  117.64  < 2e-16 ***
## FirstAuthorFemale1   0.03016    0.00475    6.35  2.2e-10 ***
## LastAuthorFemale1   -0.00718    0.00537   -1.34  0.18097
## Year1997             0.04859    0.01536    3.16  0.00156 **
## Year1998             0.03649    0.01464    2.49  0.01268 *
## Year1999             0.02599    0.01473    1.76  0.07765 .
## Year2000            -0.00748    0.01479   -0.51  0.61322
## Year2001            -0.05119    0.01482   -3.45  0.00055 ***
## Year2002            -0.06048    0.01433   -4.22  2.5e-05 ***
## Year2003            -0.10087    0.01486   -6.79  1.1e-11 ***
## Year2004            -0.08830    0.01408   -6.27  3.6e-10 ***
## Year2005            -0.06496    0.01382   -4.70  2.6e-06 ***
```

```
## Year2006           -0.11251    0.01388   -8.11  5.3e-16 ***
## Year2007           -0.08598    0.01326   -6.48  9.0e-11 ***
## Year2008           -0.05695    0.01331   -4.28  1.9e-05 ***
## Year2009           -0.04624    0.01324   -3.49  0.00048 ***
## Year2010           -0.04897    0.01337   -3.66  0.00025 ***
## Year2011           -0.06539    0.01313   -4.98  6.4e-07 ***
## Year2012           -0.05709    0.01318   -4.33  1.5e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.449
## Multiple R-squared:  0.00893,    Adjusted R-squared:  0.00854
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  3932 weights are ~= 1. The remaining 41933 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0134  0.8640  0.9500  0.8950  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          2.18e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.009  1           1.004
## Year              1.009 16           1.000
```




```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.3133 -0.3007  0.0134  0.3043  1.9706
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)          1.23595    0.01049  117.79  < 2e-16 ***
## FirstAuthorFemale1   0.02897    0.00474    6.12  9.6e-10 ***
## Year1997             0.04842    0.01536    3.15  0.00162 **
## Year1998             0.03641    0.01464    2.49  0.01286 *
## Year1999             0.02573    0.01473    1.75  0.08060 .
## Year2000            -0.00772    0.01479   -0.52  0.60185
## Year2001            -0.05151    0.01482   -3.48  0.00051 ***
## Year2002            -0.06068    0.01433   -4.23  2.3e-05 ***
## Year2003            -0.10107    0.01485   -6.80  1.0e-11 ***
## Year2004            -0.08865    0.01407   -6.30  3.0e-10 ***
## Year2005            -0.06526    0.01382   -4.72  2.3e-06 ***
## Year2006            -0.11268    0.01388   -8.12  4.7e-16 ***
```

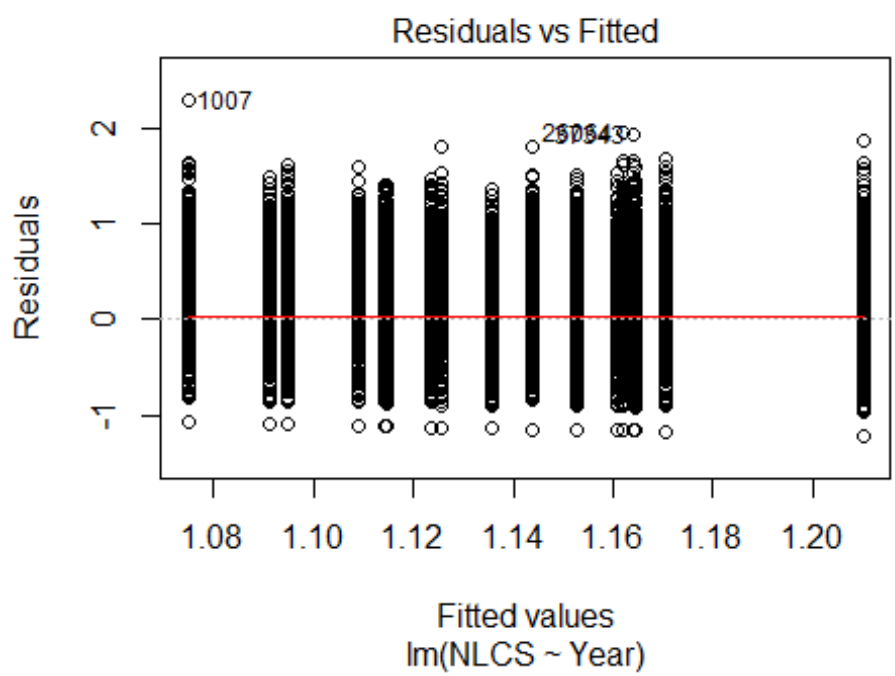
```
## Year2007           -0.08644    0.01325   -6.52  6.9e-11 ***
## Year2008           -0.05721    0.01330   -4.30  1.7e-05 ***
## Year2009           -0.04651    0.01324   -3.51  0.00044 ***
## Year2010           -0.04928    0.01337   -3.69  0.00023 ***
## Year2011           -0.06584    0.01312   -5.02  5.2e-07 ***
## Year2012           -0.05755    0.01317   -4.37  1.3e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.449
## Multiple R-squared:  0.0089, Adjusted R-squared:  0.00853
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  3916 weights are ~= 1. The remaining 41949 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0146  0.8640  0.9500  0.8950  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          2.18e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##                   GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.006  1           1.003
## Year             1.006 16           1.000
```



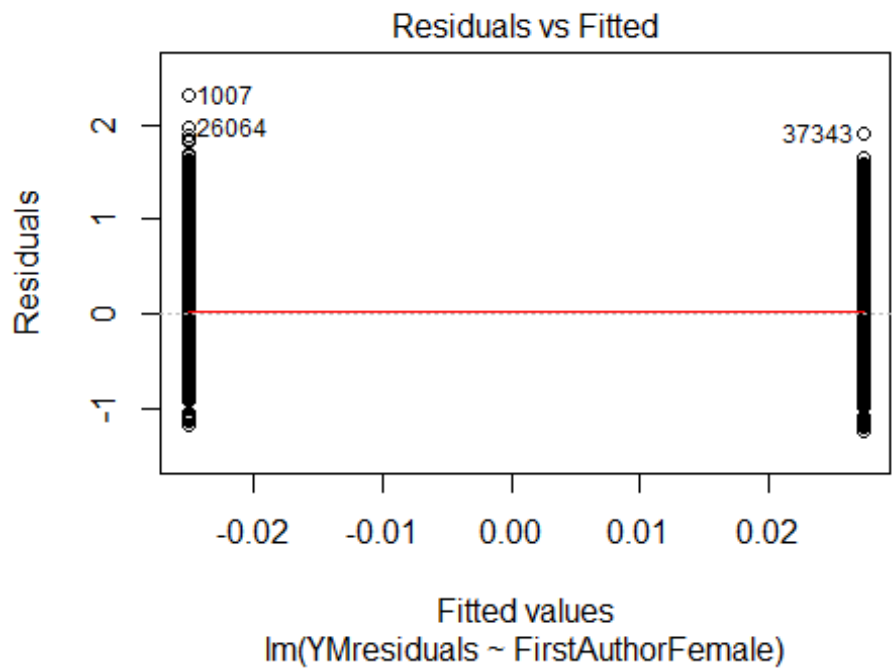

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.2915 -0.3016  0.0126  0.3050  1.9603
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)        1.243297   0.010439  119.10  < 2e-16 ***
## LastAuthorFemale1 -0.000962   0.005339   -0.18  0.85704
## Year1997           0.048180   0.015351    3.14  0.00170 **
## Year1998           0.036582   0.014624    2.50  0.01237 *
## Year1999           0.025654   0.014721    1.74  0.08139 .
## Year2000          -0.006511   0.014779   -0.44  0.65956
## Year2001          -0.050442   0.014819   -3.40  0.00066 ***
## Year2002          -0.059314   0.014327   -4.14  3.5e-05 ***
## Year2003          -0.099502   0.014844   -6.70  2.1e-11 ***
## Year2004          -0.086647   0.014078   -6.15  7.6e-10 ***
## Year2005          -0.063082   0.013810   -4.57  4.9e-06 ***
## Year2006          -0.110264   0.013868   -7.95  1.9e-15 ***
```

```
## Year2007          -0.083934   0.013256   -6.33  2.4e-10 ***
## Year2008          -0.054900   0.013296   -4.13  3.6e-05 ***
## Year2009          -0.043806   0.013239   -3.31  0.00094 ***
## Year2010          -0.046693   0.013372   -3.49  0.00048 ***
## Year2011          -0.062775   0.013117   -4.79  1.7e-06 ***
## Year2012          -0.053661   0.013165   -4.08  4.6e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.449
## Multiple R-squared:  0.0081, Adjusted R-squared:  0.00773
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  3958 weights are ~= 1. The remaining 41907 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.017   0.864   0.950   0.895   0.985   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          2.18e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Sample size for the above analysis:  45865"
## [1] ""
## [1] ""
## [1] "####################################"
## [1] "Analysis of AJSC 2729"
## [1] "####################################"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 2058 2071 1876 1844 1947 1868 1734 1451 1449 1642 1883 2150 2232 2295 2300
## 2011 2012
## 2498 2452
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 1097 1183 1045 1214 1280 1188 1268 1061 1082 1242 1326 1526 1640 1712 1676
## 2011 2012
```

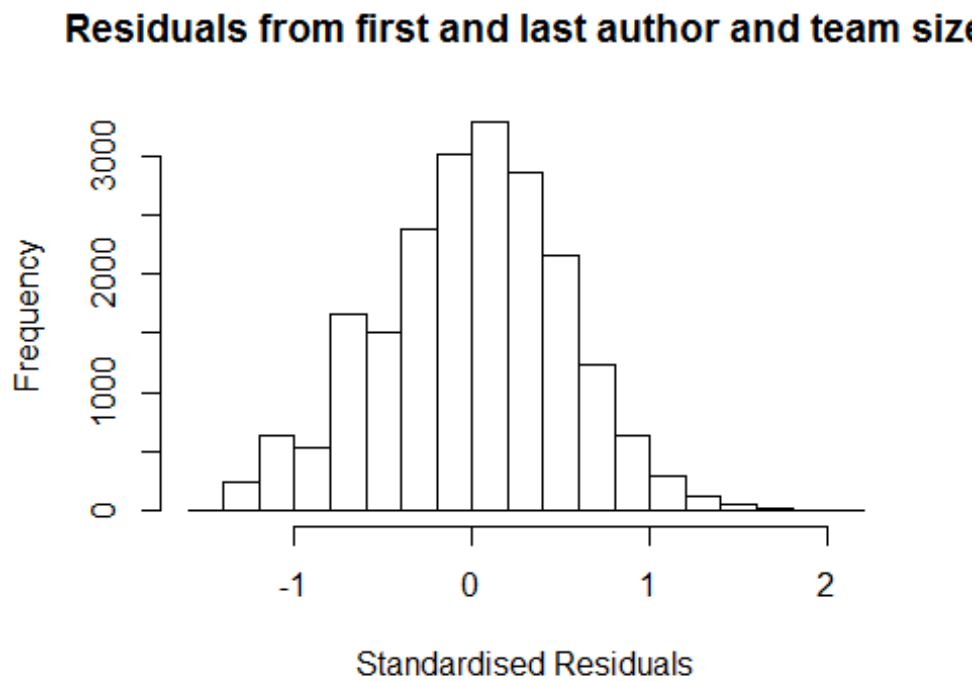
```
## 1890 1882
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  984 1044  923 1078 1139 1084 1153  912  971 1088 1171 1354 1448 1503 1453
## 2011 2012
## 1671 1676
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
##  Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 110, df = 16, p-value <2e-16
```




```
##
##  Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 33, df = 1, p-value = 1e-08
```

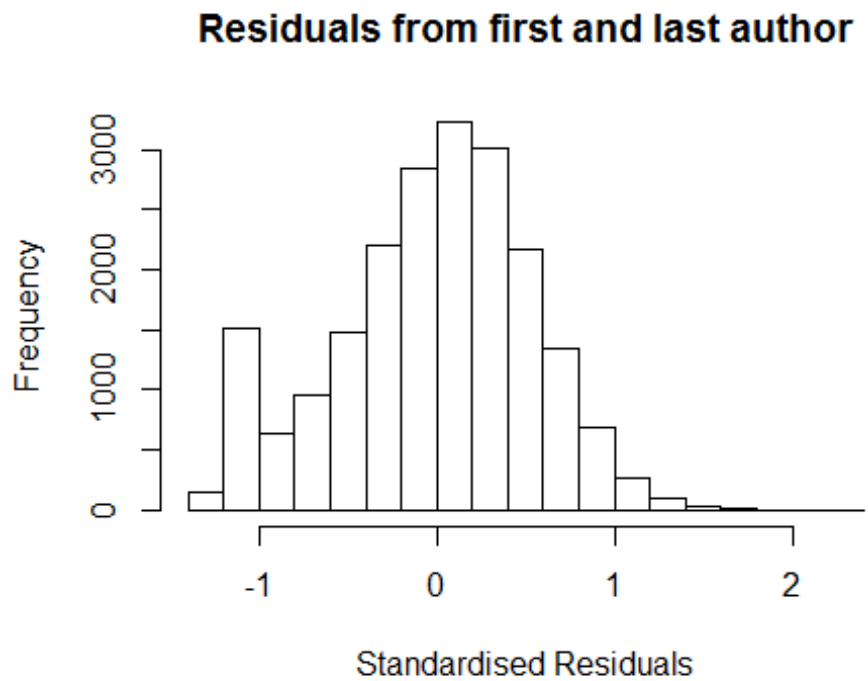



```
## [1] "Female first author team size 2018 geometric mean: 4.46580351520726"
## [1] "Male first author team size 2018 geometric mean: 4.17395742106163"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 380000, p-value = 0.2
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 4.29520229167562"
## [1] "Male last author team size 2018 geometric mean: 4.47960011554947"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 420000, p-value = 0.1
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.107  1           1.052
## LastAuthorFemale  1.072  1           1.035
## UniqueAuthors     1.069  4           1.008
## Year              1.127 16           1.004
```



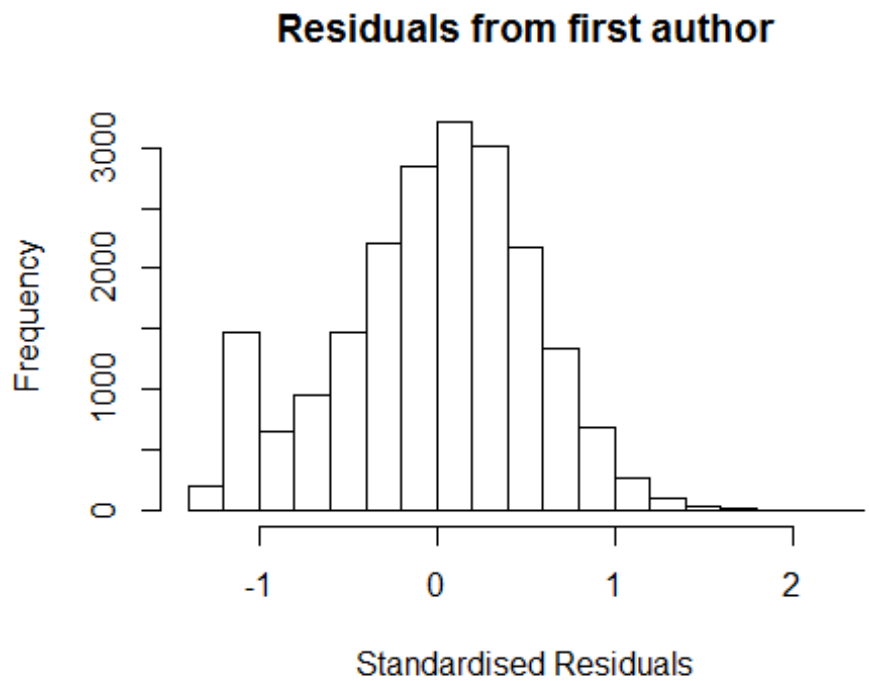

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
UniqueAuthors +
##     Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##     k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.4081 -0.3492  0.0209  0.3496  2.0641
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)          0.67944    0.02343   29.00  < 2e-16 ***
## FirstAuthorFemale1   0.02925    0.00788    3.71  0.00021 ***
## LastAuthorFemale1    0.02100    0.00809    2.60  0.00943 **
## UniqueAuthors2       0.32280    0.01847   17.48  < 2e-16 ***
## UniqueAuthors3       0.43095    0.01694   25.44  < 2e-16 ***
## UniqueAuthors4       0.48851    0.01686   28.98  < 2e-16 ***
## UniqueAuthors5       0.61741    0.01536   40.19  < 2e-16 ***
## Year1997             0.01017    0.02640    0.39  0.70014
## Year1998             0.01063    0.02721    0.39  0.69593
## Year1999             0.02868    0.02632    1.09  0.27588
```

```
## Year2000            0.06845    0.02554    2.68  0.00736 **
## Year2001            0.01098    0.02646    0.42  0.67813
## Year2002           -0.02025    0.02554   -0.79  0.42790
## Year2003           -0.01094    0.02554   -0.43  0.66832
## Year2004           -0.05898    0.02560   -2.30  0.02124 *
## Year2005           -0.00439    0.02430   -0.18  0.85651
## Year2006            0.06096    0.02462    2.48  0.01327 *
## Year2007            0.01665    0.02484    0.67  0.50281
## Year2008           -0.00330    0.02407   -0.14  0.89089
## Year2009           -0.05844    0.02383   -2.45  0.01421 *
## Year2010           -0.02706    0.02367   -1.14  0.25303
## Year2011           -0.03385    0.02377   -1.42  0.15442
## Year2012           -0.01775    0.02427   -0.73  0.46455
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.506
## Multiple R-squared:  0.128,  Adjusted R-squared:  0.127
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
##  1691 weights are ~= 1. The remaining 18961 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0583  0.8570  0.9480  0.9010  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          4.84e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.087  1           1.042
## LastAuthorFemale  1.049  1           1.024
## Year              1.070 16           1.002
```



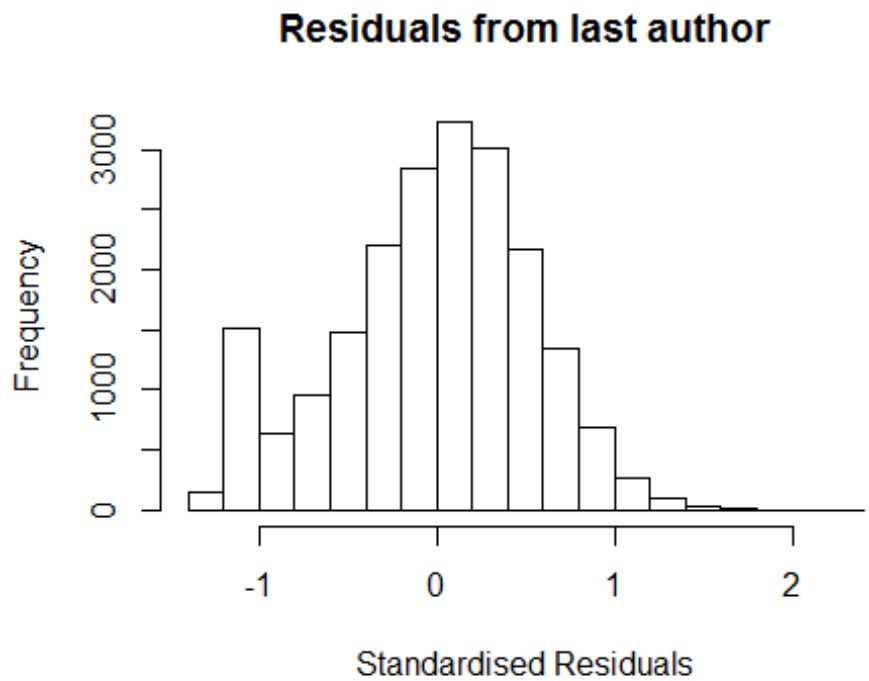

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.2503 -0.3491  0.0374  0.3605  2.2951
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.065854   0.021341   49.94  < 2e-16 ***
## FirstAuthorFemale1  0.056853   0.008317    6.84  8.4e-12 ***
## LastAuthorFemale1  -0.002353   0.008512   -0.28  0.78221
## Year1997            0.026311   0.028475    0.92  0.35550
## Year1998            0.027682   0.029669    0.93  0.35083
## Year1999            0.052901   0.028593    1.85  0.06431 .
## Year2000            0.097394   0.027579    3.53  0.00041 ***
## Year2001            0.050603   0.028926    1.75  0.08024 .
## Year2002            0.021658   0.027275    0.79  0.42717
## Year2003            0.060661   0.027799    2.18  0.02911 *
## Year2004           -0.000615   0.027834   -0.02  0.98238
## Year2005            0.086934   0.026080    3.33  0.00086 ***
```

```
## Year2006            0.129932   0.026381    4.93  8.5e-07 ***
## Year2007            0.087659   0.026647    3.29  0.00100 **
## Year2008            0.077555   0.025877    3.00  0.00273 **
## Year2009            0.034787   0.025609    1.36  0.17436
## Year2010            0.070918   0.025391    2.79  0.00523 **
## Year2011            0.061554   0.025410    2.42  0.01543 *
## Year2012            0.079065   0.025873    3.06  0.00225 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.525
## Multiple R-squared:  0.00662,    Adjusted R-squared:  0.00575
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
##  1768 weights are ~= 1. The remaining 18884 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0163  0.8570  0.9490  0.8960  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          4.84e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.058  1           1.029
## Year              1.058 16           1.002
```




```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.2515 -0.3492  0.0374  0.3598  2.2956
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.065394   0.021277   50.07  < 2e-16 ***
## FirstAuthorFemale1  0.056309   0.008227    6.84  7.9e-12 ***
## Year1997            0.026283   0.028475    0.92  0.35602
## Year1998            0.027668   0.029671    0.93  0.35109
## Year1999            0.052874   0.028591    1.85  0.06442 .
## Year2000            0.097265   0.027568    3.53  0.00042 ***
## Year2001            0.050596   0.028927    1.75  0.08029 .
## Year2002            0.021610   0.027271    0.79  0.42812
## Year2003            0.060643   0.027799    2.18  0.02916 *
## Year2004           -0.000662   0.027830   -0.02  0.98101
## Year2005            0.086856   0.026075    3.33  0.00087 ***
## Year2006            0.129812   0.026374    4.92  8.6e-07 ***
```

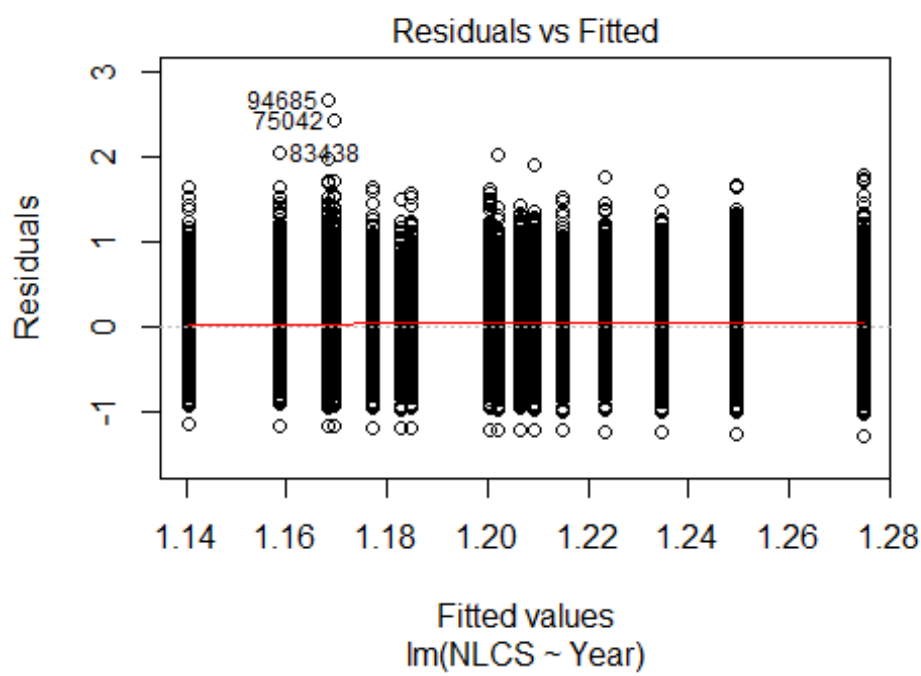
```
## Year2007            0.087530   0.026638    3.29  0.00102 **
## Year2008            0.077383   0.025863    2.99  0.00277 **
## Year2009            0.034630   0.025595    1.35  0.17607
## Year2010            0.070703   0.025368    2.79  0.00532 **
## Year2011            0.061321   0.025384    2.42  0.01571 *
## Year2012            0.078773   0.025825    3.05  0.00229 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.525
## Multiple R-squared:  0.00662,    Adjusted R-squared:  0.0058
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
##  1769 weights are ~= 1. The remaining 18883 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0162  0.8570  0.9490  0.8960  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          4.84e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##                   GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.021  1           1.010
## Year             1.021 16           1.001
```



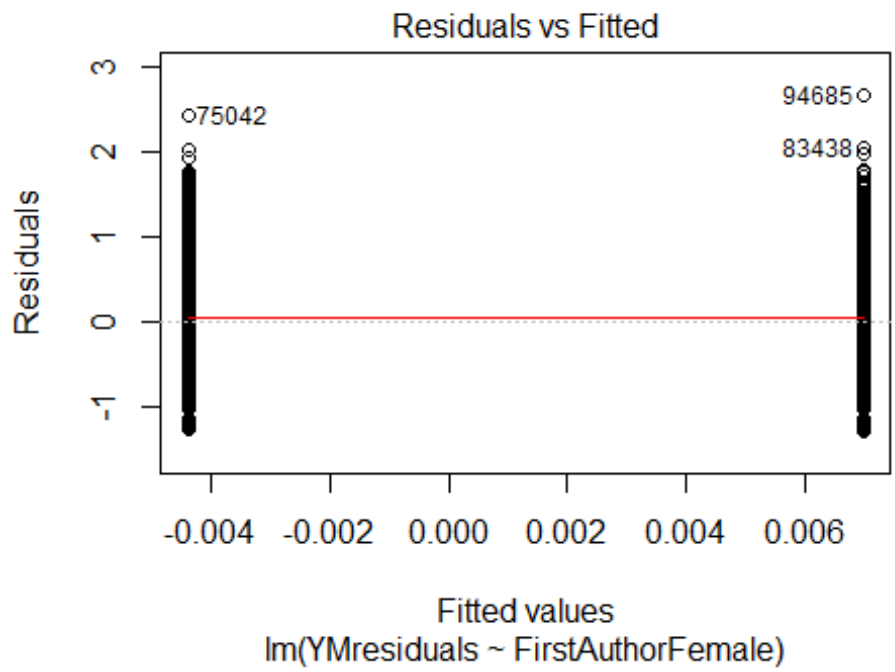

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.2312 -0.3525  0.0344  0.3620  2.2824
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.07856    0.02126   50.73  < 2e-16 ***
## LastAuthorFemale1   0.01155    0.00841    1.37  0.17004
## Year1997            0.03134    0.02844    1.10  0.27042
## Year1998            0.03118    0.02967    1.05  0.29339
## Year1999            0.05636    0.02857    1.97  0.04855 *
## Year2000            0.10401    0.02748    3.78  0.00015 ***
## Year2001            0.05731    0.02891    1.98  0.04745 *
## Year2002            0.02965    0.02722    1.09  0.27601
## Year2003            0.06812    0.02778    2.45  0.01421 *
## Year2004            0.00711    0.02780    0.26  0.79812
## Year2005            0.09561    0.02601    3.68  0.00024 ***
## Year2006            0.14112    0.02630    5.37  8.1e-08 ***
```

```
## Year2007           0.09908    0.02660    3.73  0.00020 ***
## Year2008           0.09232    0.02570    3.59  0.00033 ***
## Year2009           0.04865    0.02548    1.91  0.05627 .
## Year2010           0.08589    0.02523    3.40  0.00066 ***
## Year2011           0.07767    0.02522    3.08  0.00207 **
## Year2012           0.09645    0.02568    3.76  0.00017 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.525
## Multiple R-squared:  0.00428,    Adjusted R-squared:  0.00346
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
##  1718 weights are ~= 1. The remaining 18934 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0191  0.8580  0.9490  0.8970  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          4.84e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Sample size for the above analysis:  20652"
## [1] ""
## [1] ""
## [1] "####################################"
## [1] "Analysis of AJSC 2730"
## [1] "####################################"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 3722 3697 3824 3880 3998 4385 4290 4028 4349 4822 5281 5574 5936 5844 5698
## 2011 2012
## 6375 6432
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 2245 2328 2461 2445 2302 2092 3021 2880 3043 3413 3760 3985 4255 4162 4029
## 2011 2012
```

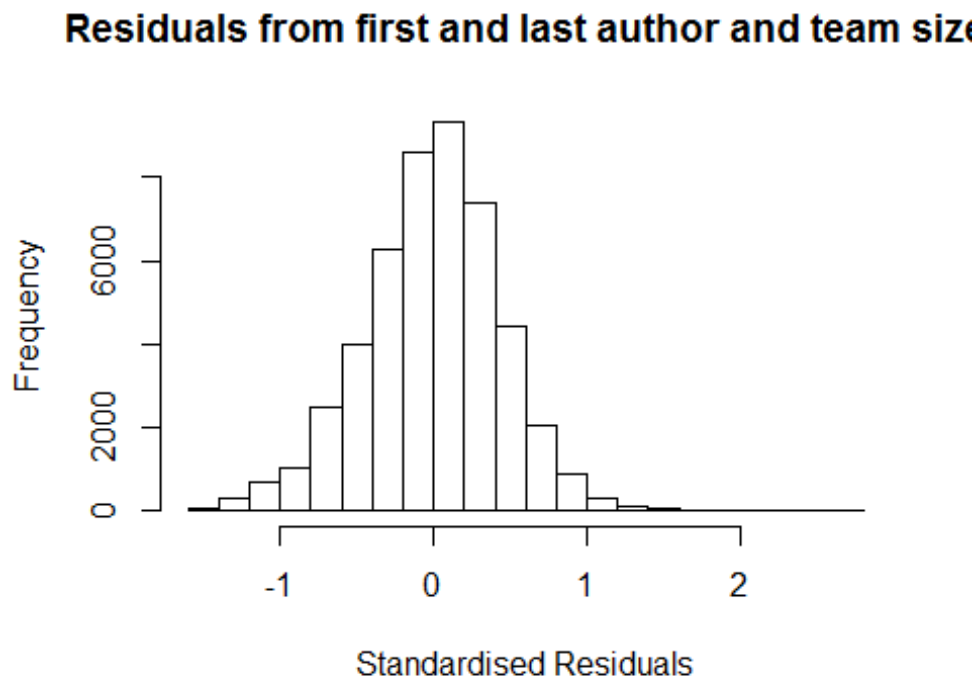
```
## 4590 4672
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 1949 2034 2155 2109 1958 1814 2609 2464 2580 2918 3236 3393 3644 3535 3464
## 2011 2012
## 3905 4052
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
##  Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 440, df = 16, p-value <2e-16
```




```
##
##  Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 120, df = 1, p-value <2e-16
```

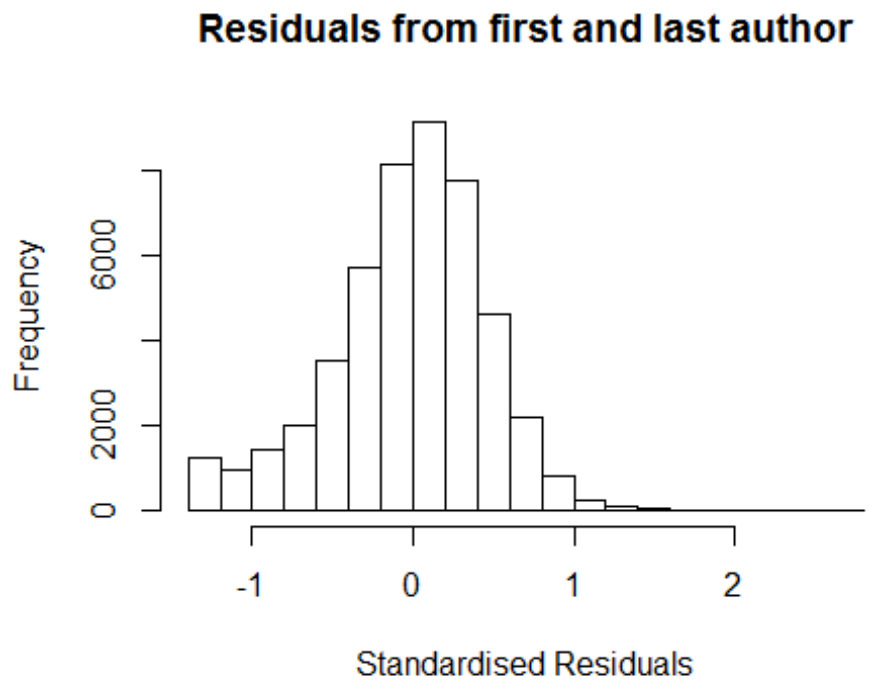



```
## [1] "Female first author team size 2018 geometric mean: 6.44681422193723"
## [1] "Male first author team size 2018 geometric mean: 6.63621458855587"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 2400000, p-value = 0.02
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 6.01450820669504"
## [1] "Male last author team size 2018 geometric mean: 6.87222545606976"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 2100000, p-value = 3e-12
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.037  1           1.018
## LastAuthorFemale  1.037  1           1.018
## UniqueAuthors     1.030  4           1.004
## Year              1.047 16           1.001
```



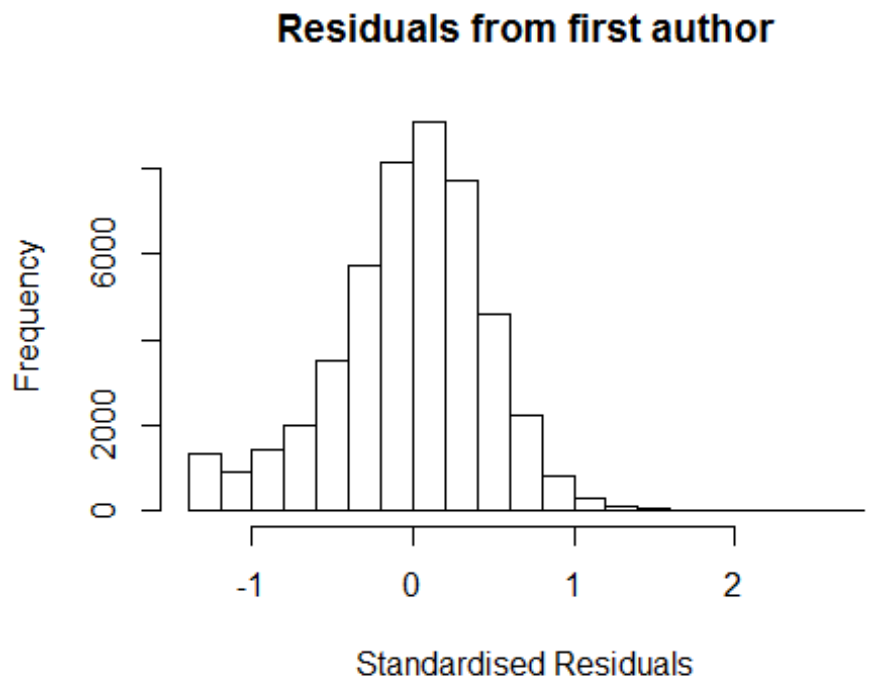

```
## [1] "List of  1 outliers with residuals above  2.5"
##          ScopusId  NLCS Year OneField Fields residuals
## 94685 84855792427 3.831 2012     2720      2     2.771
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
UniqueAuthors +
##     Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##     k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.4535 -0.2782  0.0117  0.2763  2.7708
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         0.74888    0.01675   44.70  < 2e-16 ***
## FirstAuthorFemale1 -0.00111    0.00410   -0.27    0.787
## LastAuthorFemale1  -0.01102    0.00460   -2.40    0.017 *
## UniqueAuthors2      0.38602    0.01398   27.61  < 2e-16 ***
## UniqueAuthors3      0.49235    0.01308   37.65  < 2e-16 ***
## UniqueAuthors4      0.55726    0.01254   44.43  < 2e-16 ***
## UniqueAuthors5      0.70457    0.01131   62.27  < 2e-16 ***
## Year1997           -0.01534    0.01638   -0.94    0.349
## Year1998           -0.06844    0.01588   -4.31  1.6e-05 ***
## Year1999           -0.07220    0.01578   -4.58  4.8e-06 ***
```

```
## Year2000           -0.10288    0.01591   -6.47  1.0e-10 ***
## Year2001           -0.09741    0.01660   -5.87  4.4e-09 ***
## Year2002           -0.11074    0.01507   -7.35  2.1e-13 ***
## Year2003           -0.15797    0.01463  -10.80  < 2e-16 ***
## Year2004           -0.19362    0.01468  -13.19  < 2e-16 ***
## Year2005           -0.17212    0.01454  -11.84  < 2e-16 ***
## Year2006           -0.14008    0.01438   -9.74  < 2e-16 ***
## Year2007           -0.12159    0.01444   -8.42  < 2e-16 ***
## Year2008           -0.12937    0.01434   -9.02  < 2e-16 ***
## Year2009           -0.13150    0.01440   -9.13  < 2e-16 ***
## Year2010           -0.16975    0.01457  -11.65  < 2e-16 ***
## Year2011           -0.17246    0.01435  -12.02  < 2e-16 ***
## Year2012           -0.17996    0.01448  -12.43  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.411
## Multiple R-squared:  0.184,  Adjusted R-squared:  0.183
## Convergence in 15 IRWLS iterations
##
## Robustness weights:
##  4 observations c(9738,37373,41599,47466)
##   are outliers with |weight| = 0 ( < 2.1e-06);
##  4031 weights are ~= 1. The remaining 43784 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0017  0.8600  0.9500  0.8970  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          2.09e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.022  1           1.011
## LastAuthorFemale  1.020  1           1.010
## Year              1.021 16           1.001
```



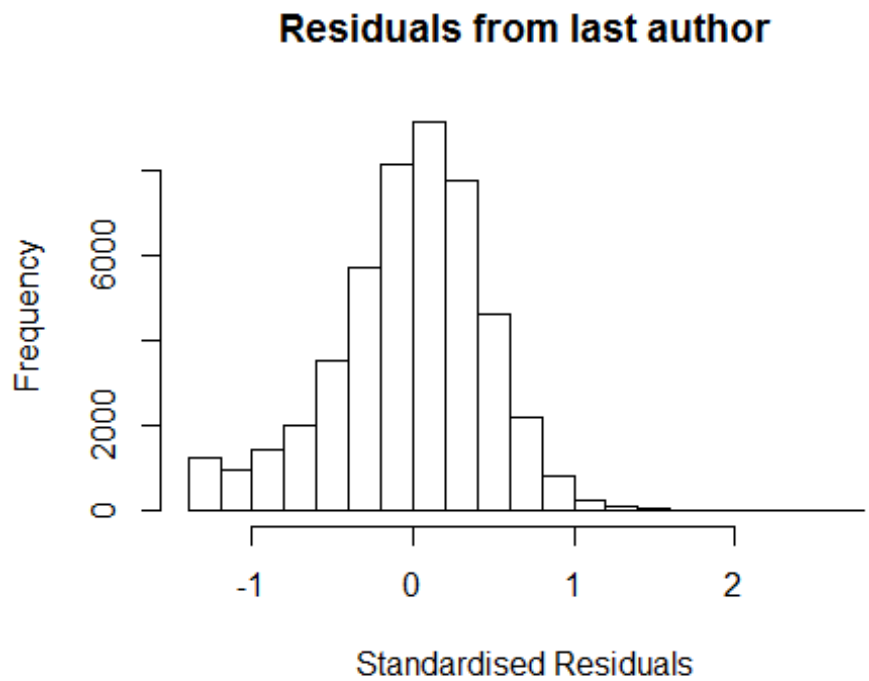

```
## [1] "List of  1 outliers with residuals above  2.5"
##          ScopusId  NLCS Year OneField Fields residuals
## 94685 84855792427 3.831 2012     2720      2      2.62
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.3293 -0.2935  0.0209  0.2878  2.6198
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)          1.31005    0.01304  100.50  < 2e-16 ***
## FirstAuthorFemale1  0.01632    0.00434    3.76  0.00017 ***
## LastAuthorFemale1  -0.03542    0.00491   -7.22  5.3e-13 ***
## Year1997            0.00293    0.01698    0.17  0.86281
## Year1998           -0.05093    0.01656   -3.08  0.00210 **
## Year1999           -0.05606    0.01629   -3.44  0.00058 ***
## Year2000           -0.08125    0.01655   -4.91  9.2e-07 ***
## Year2001           -0.08742    0.01738   -5.03  4.9e-07 ***
## Year2002           -0.07050    0.01561   -4.52  6.3e-06 ***
## Year2003           -0.10632    0.01528   -6.96  3.5e-12 ***
## Year2004           -0.14803    0.01558   -9.50  < 2e-16 ***
## Year2005           -0.11557    0.01531   -7.55  4.5e-14 ***
```

```
## Year2006           -0.09995    0.01501   -6.66  2.8e-11 ***
## Year2007           -0.07346    0.01515   -4.85  1.2e-06 ***
## Year2008           -0.07873    0.01498   -5.25  1.5e-07 ***
## Year2009           -0.07654    0.01503   -5.09  3.6e-07 ***
## Year2010           -0.11188    0.01530   -7.31  2.7e-13 ***
## Year2011           -0.12789    0.01504   -8.50  < 2e-16 ***
## Year2012           -0.11515    0.01515   -7.60  3.0e-14 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.427
## Multiple R-squared:  0.00768,    Adjusted R-squared:  0.00731
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
##  4 observations c(9738,37373,41599,47466)
##   are outliers with |weight| = 0 ( < 2.1e-06);
##  3936 weights are ~= 1. The remaining 43879 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0008  0.8600  0.9500  0.8880  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          2.09e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.013  1           1.006
## Year              1.013 16           1.000
```




```
## [1] "List of  1 outliers with residuals above  2.5"
##          ScopusId  NLCS Year OneField Fields residuals
## 94685 84855792427 3.831 2012     2720      2      2.62
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.3179 -0.2942  0.0206  0.2878  2.6341
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)          1.30435    0.01300  100.37  < 2e-16 ***
## FirstAuthorFemale1   0.01093    0.00434    2.52   0.0118 *
## Year1997             0.00267    0.01697    0.16   0.8748
## Year1998            -0.05103    0.01655   -3.08   0.0021 **
## Year1999            -0.05586    0.01628   -3.43   0.0006 ***
## Year2000            -0.08175    0.01655   -4.94  7.8e-07 ***
## Year2001            -0.08721    0.01737   -5.02  5.2e-07 ***
## Year2002            -0.07084    0.01560   -4.54  5.6e-06 ***
## Year2003            -0.10655    0.01527   -6.98  3.1e-12 ***
## Year2004            -0.14872    0.01557   -9.55  < 2e-16 ***
## Year2005            -0.11714    0.01530   -7.66  2.0e-14 ***
## Year2006            -0.10093    0.01501   -6.72  1.8e-11 ***
```

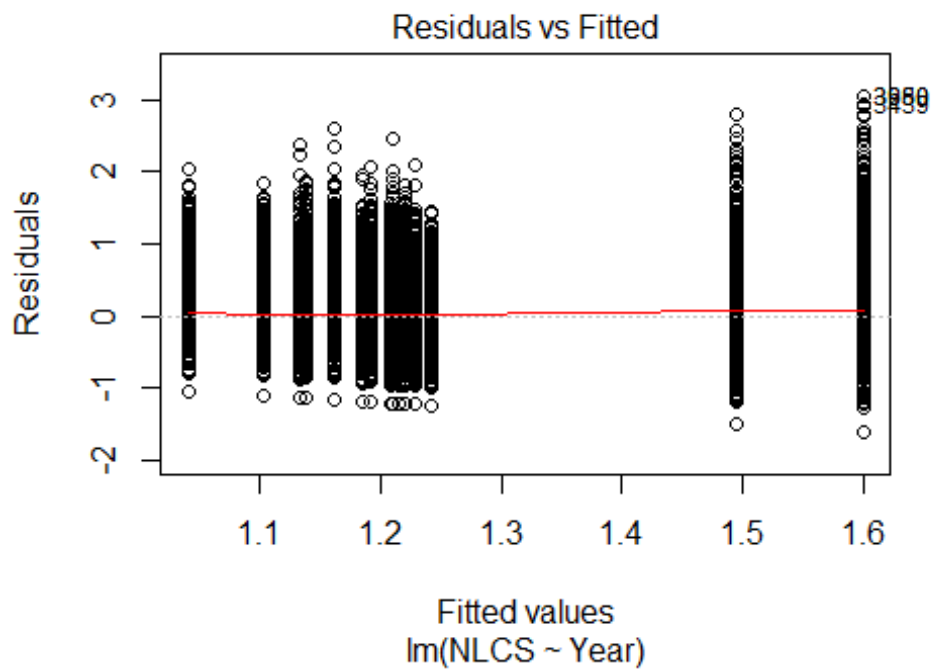
```
## Year2007           -0.07517    0.01514   -4.97  6.9e-07 ***
## Year2008           -0.08058    0.01498   -5.38  7.5e-08 ***
## Year2009           -0.07891    0.01502   -5.25  1.5e-07 ***
## Year2010           -0.11414    0.01530   -7.46  8.7e-14 ***
## Year2011           -0.13078    0.01502   -8.70  < 2e-16 ***
## Year2012           -0.11842    0.01513   -7.82  5.2e-15 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.427
## Multiple R-squared:  0.00659,    Adjusted R-squared:  0.00624
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
##  4 observations c(9738,37373,41599,47466)
##   are outliers with |weight| = 0 ( < 2.1e-06);
##  3931 weights are ~= 1. The remaining 43884 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0025  0.8600  0.9490  0.8880  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          2.09e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##                   GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.011  1           1.005
## Year             1.011 16           1.000
```



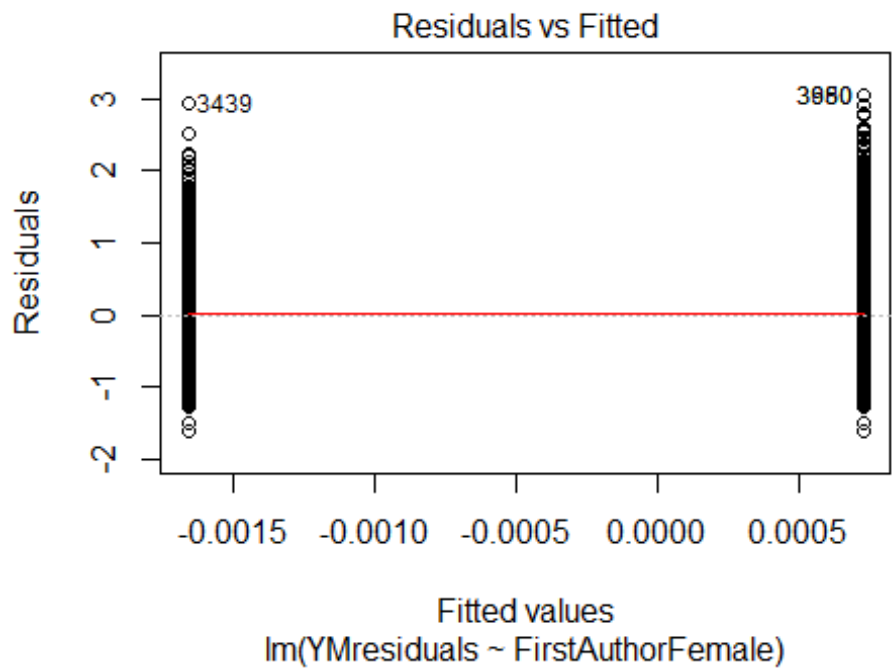

```
## [1] "List of  1 outliers with residuals above  2.5"
##          ScopusId  NLCS Year OneField Fields residuals
## 94685 84855792427 3.831 2012     2720      2      2.62
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.3175 -0.2932  0.0207  0.2874  2.6295
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.31446    0.01298  101.29  < 2e-16 ***
## LastAuthorFemale1 -0.03231    0.00489   -6.60  4.1e-11 ***
## Year1997            0.00304    0.01699    0.18  0.85813
## Year1998          -0.05031    0.01656   -3.04  0.00239 **
## Year1999          -0.05555    0.01629   -3.41  0.00065 ***
## Year2000          -0.08078    0.01656   -4.88  1.1e-06 ***
## Year2001          -0.08672    0.01739   -4.99  6.2e-07 ***
## Year2002          -0.06982    0.01562   -4.47  7.9e-06 ***
## Year2003          -0.10528    0.01529   -6.89  5.8e-12 ***
## Year2004          -0.14711    0.01559   -9.44  < 2e-16 ***
## Year2005          -0.11412    0.01531   -7.45  9.3e-14 ***
## Year2006          -0.09817    0.01501   -6.54  6.3e-11 ***
```

```
## Year2007          -0.07185    0.01515   -4.74  2.1e-06 ***
## Year2008          -0.07685    0.01498   -5.13  2.9e-07 ***
## Year2009          -0.07434    0.01503   -4.95  7.6e-07 ***
## Year2010          -0.10987    0.01530   -7.18  7.0e-13 ***
## Year2011          -0.12585    0.01503   -8.37  < 2e-16 ***
## Year2012          -0.11300    0.01514   -7.46  8.6e-14 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.427
## Multiple R-squared:  0.00741,    Adjusted R-squared:  0.00706
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
##  4 observations c(9738,37373,41599,47466)
##   are outliers with |weight| = 0 ( < 2.1e-06);
##  3933 weights are ~= 1. The remaining 43882 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0005  0.8600  0.9500  0.8880  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          2.09e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                  psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Sample size for the above analysis:  47819"
## [1] ""
## [1] ""
## [1] "##################################"
## [1] "Analysis of AJSC 2731"
## [1] "##################################"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 4336 3395 1662 1714 2010 1998 2080 1857 1590 1715 2059 2046 2040 2066 2063
## 2011 2012
## 2245 1988
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
```

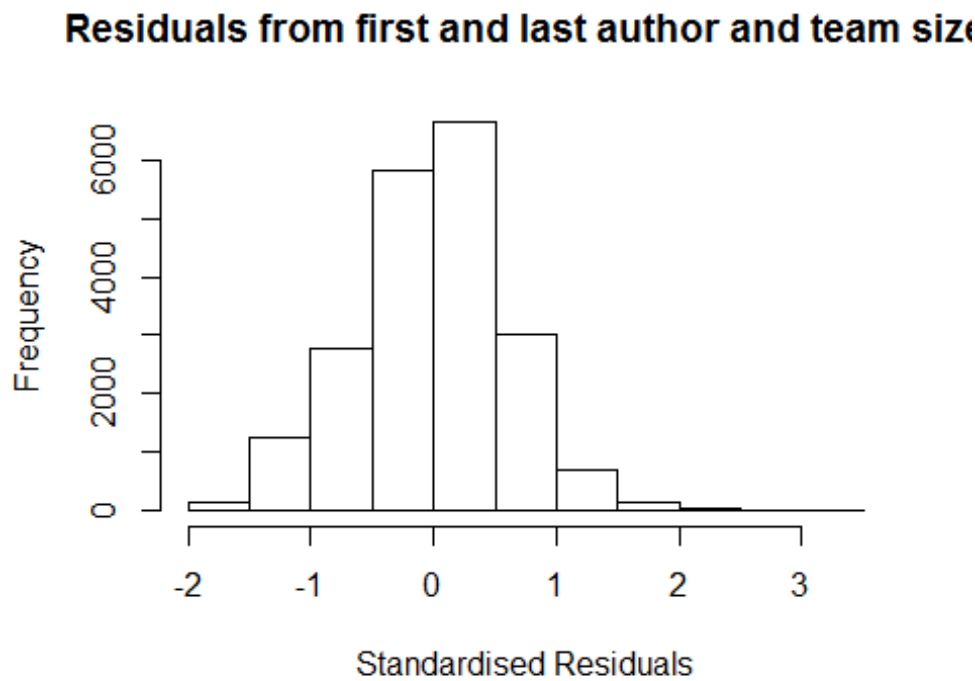
```
## 1137 1111 1151 1203 1301 1032 1628 1475 1208 1334 1566 1562 1564 1559 1595
## 2011 2012
## 1691 1506
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  986  969 1015 1040 1155  877 1427 1279 1034 1160 1348 1350 1356 1356 1371
## 2011 2012
## 1460 1294
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
##  Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 1600, df = 16, p-value <2e-16
```




```
##
##  Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 33, df = 1, p-value = 8e-09
```



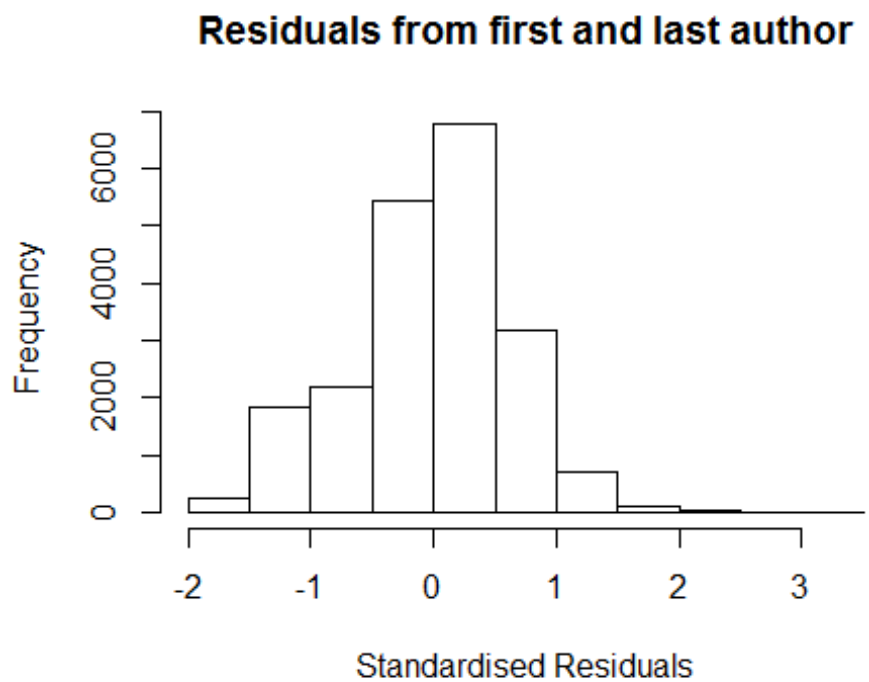

```
## [1] "Female first author team size 2018 geometric mean: 4.3659705928614"
## [1] "Male first author team size 2018 geometric mean: 4.02583013897902"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 2e+05, p-value = 0.1
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 4.31897243298424"
## [1] "Male last author team size 2018 geometric mean: 4.08939438766376"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 170000, p-value = 0.08
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.033  1           1.016
## LastAuthorFemale  1.018  1           1.009
## UniqueAuthors     1.034  4           1.004
## Year              1.049 16           1.001
```




```
## [1] "List of  9 outliers with residuals above  2.5"
##        ScopusId  NLCS Year OneField Fields residuals
## 93   0029908754 3.875 1996     2731      1     2.659
## 3439 0029757496 4.541 1996     2731      1     2.632
## 3521 0029816808 4.380 1996     2731      1     2.744
## 3680 0029915446 4.649 1996     2731      1     2.710
## 3950 0030067257 4.642 1996     2731      1     2.706
## 4545 0029980777 4.398 1996     2731      3     3.182
## 5628 0030779611 4.287 1997     2731      2     2.865
## 5635 0030832881 4.070 1997     2731      2     2.648
## 7534 0031001942 3.956 1997     2731      1     2.881
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
UniqueAuthors +
##     Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##     k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.9394 -0.3803  0.0175  0.3864  3.1818
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)          1.21622    0.04521   26.90  < 2e-16 ***
## FirstAuthorFemale1 -0.03063    0.00908   -3.37  0.00075 ***
```

```
## LastAuthorFemale1  -0.00311    0.01040   -0.30  0.76520
## UniqueAuthors2      0.34657    0.02074   16.71  < 2e-16 ***
## UniqueAuthors3      0.42296    0.01992   21.23  < 2e-16 ***
## UniqueAuthors4      0.53495    0.01997   26.79  < 2e-16 ***
## UniqueAuthors5      0.72323    0.01843   39.23  < 2e-16 ***
## Year1997           -0.14089    0.05106   -2.76  0.00580 **
## Year1998           -0.42697    0.04515   -9.46  < 2e-16 ***
## Year1999           -0.45668    0.04463  -10.23  < 2e-16 ***
## Year2000           -0.59064    0.04856  -12.16  < 2e-16 ***
## Year2001           -0.44156    0.04632   -9.53  < 2e-16 ***
## Year2002           -0.52162    0.04563  -11.43  < 2e-16 ***
## Year2003           -0.55714    0.04578  -12.17  < 2e-16 ***
## Year2004           -0.50185    0.04483  -11.19  < 2e-16 ***
## Year2005           -0.45885    0.04475  -10.25  < 2e-16 ***
## Year2006           -0.48996    0.04417  -11.09  < 2e-16 ***
## Year2007           -0.49900    0.04424  -11.28  < 2e-16 ***
## Year2008           -0.47044    0.04440  -10.60  < 2e-16 ***
## Year2009           -0.50973    0.04467  -11.41  < 2e-16 ***
## Year2010           -0.52252    0.04487  -11.65  < 2e-16 ***
## Year2011           -0.57800    0.04460  -12.96  < 2e-16 ***
## Year2012           -0.56847    0.04506  -12.62  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.553
## Multiple R-squared:  0.16,   Adjusted R-squared:  0.159
## Convergence in 21 IRWLS iterations
##
## Robustness weights:
##  9 observations c(54,492,507,541,610,926,1064,1070,1583)
##   are outliers with |weight| = 0 ( < 4.9e-06);
##  1678 weights are ~= 1. The remaining 18790 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0154  0.8570  0.9470  0.8940  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          4.88e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                    psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
```

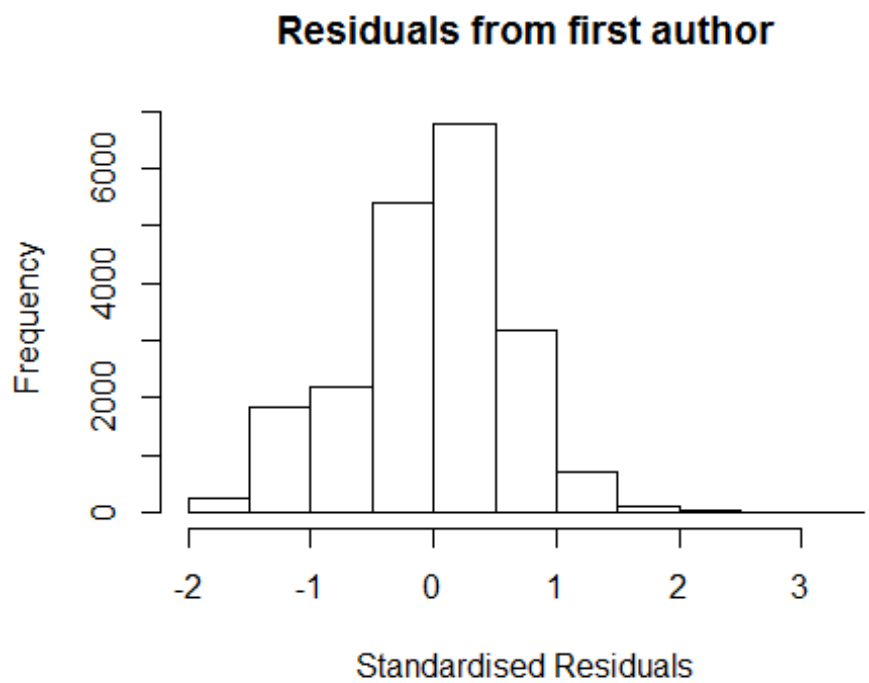
```
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##                     GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.023  1           1.011
## LastAuthorFemale  1.010  1           1.005
## Year              1.025 16           1.001
```




```
## [1] "List of  9 outliers with residuals above  2.5"
##          ScopusId  NLCS Year OneField Fields residuals
## 3439   0029757496 4.541 1996     2731      1     2.898
## 3451   0029761332 4.196 1996     2731      1     2.556
## 3521   0029816808 4.380 1996     2731      1     2.747
## 3680   0029915446 4.649 1996     2731      1     3.009
## 3950   0030067257 4.642 1996     2731      1     3.009
## 4545   0029980777 4.398 1996     2731      3     2.758
## 5628   0030779611 4.287 1997     2731      2     2.765
## 5635   0030832881 4.070 1997     2731      2     2.548
## 40505 84863320414 3.758 2012     2731      1     2.593
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.6426 -0.3936  0.0386  0.4012  3.0089
```

```
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.64008    0.04461   36.76  < 2e-16 ***
## FirstAuthorFemale1  0.00250    0.00971    0.26    0.797
## LastAuthorFemale1  -0.00681    0.01132   -0.60    0.547
## Year1997           -0.11774    0.05345   -2.20    0.028 *
## Year1998           -0.38810    0.04758   -8.16  3.6e-16 ***
## Year1999           -0.42876    0.04712   -9.10  < 2e-16 ***
## Year2000           -0.58235    0.05177  -11.25  < 2e-16 ***
## Year2001           -0.39492    0.04900   -8.06  8.1e-16 ***
## Year2002           -0.48727    0.04869  -10.01  < 2e-16 ***
## Year2003           -0.52669    0.04924  -10.70  < 2e-16 ***
## Year2004           -0.43786    0.04722   -9.27  < 2e-16 ***
## Year2005           -0.40114    0.04719   -8.50  < 2e-16 ***
## Year2006           -0.41098    0.04661   -8.82  < 2e-16 ***
## Year2007           -0.42775    0.04672   -9.16  < 2e-16 ***
## Year2008           -0.39179    0.04703   -8.33  < 2e-16 ***
## Year2009           -0.43164    0.04734   -9.12  < 2e-16 ***
## Year2010           -0.44117    0.04761   -9.27  < 2e-16 ***
## Year2011           -0.49618    0.04728  -10.49  < 2e-16 ***
## Year2012           -0.47475    0.04782   -9.93  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.581
## Multiple R-squared:  0.0363, Adjusted R-squared:  0.0355
## Convergence in 22 IRWLS iterations
##
## Robustness weights:
##  6 observations c(492,507,541,610,926,1064)
##   are outliers with |weight| = 0 ( < 4.9e-06);
##  1740 weights are ~= 1. The remaining 18731 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0087  0.8430  0.9490  0.8910  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          4.88e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
```

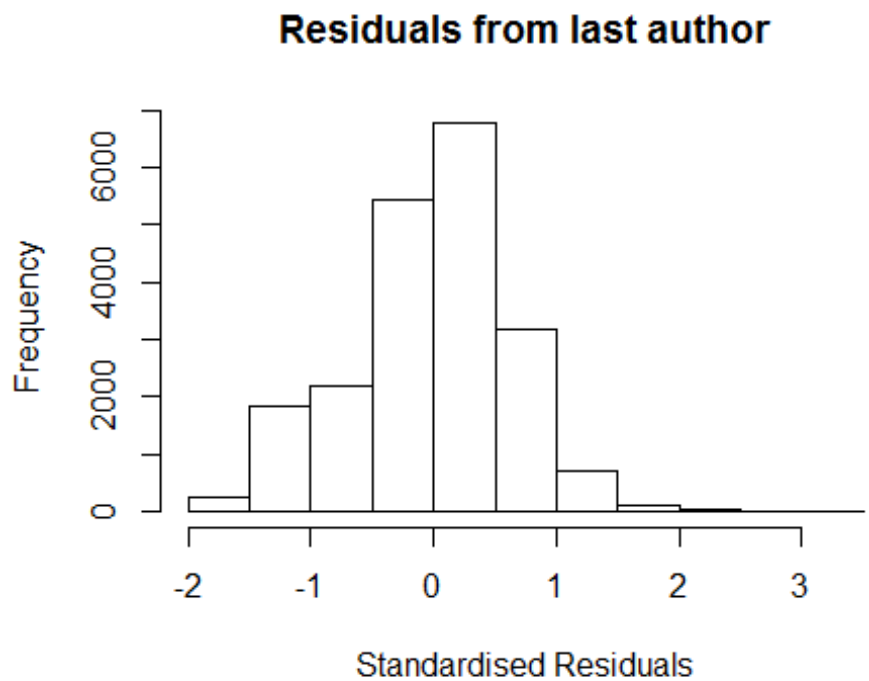
```
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##                   GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.02  1           1.010
## Year              1.02 16           1.001
```




```
## [1] "List of  9 outliers with residuals above  2.5"
##          ScopusId  NLCS Year OneField Fields residuals
## 3439   0029757496 4.541 1996     2731      1     2.898
## 3451   0029761332 4.196 1996     2731      1     2.556
## 3521   0029816808 4.380 1996     2731      1     2.747
## 3680   0029915446 4.649 1996     2731      1     3.009
## 3950   0030067257 4.642 1996     2731      1     3.009
## 4545   0029980777 4.398 1996     2731      3     2.758
## 5628   0030779611 4.287 1997     2731      2     2.765
## 5635   0030832881 4.070 1997     2731      2     2.548
## 40505 84863320414 3.758 2012     2731      1     2.593
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.6407 -0.3932  0.0395  0.4010  3.0100
##
```

```
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.63899    0.04456   36.78  < 2e-16 ***
## FirstAuthorFemale1  0.00170    0.00973    0.17    0.861
## Year1997           -0.11756    0.05344   -2.20    0.028 *
## Year1998           -0.38819    0.04757   -8.16  3.5e-16 ***
## Year1999           -0.42885    0.04712   -9.10  < 2e-16 ***
## Year2000           -0.58263    0.05179  -11.25  < 2e-16 ***
## Year2001           -0.39480    0.04900   -8.06  8.2e-16 ***
## Year2002           -0.48745    0.04869  -10.01  < 2e-16 ***
## Year2003           -0.52680    0.04924  -10.70  < 2e-16 ***
## Year2004           -0.43807    0.04721   -9.28  < 2e-16 ***
## Year2005           -0.40130    0.04719   -8.50  < 2e-16 ***
## Year2006           -0.41106    0.04661   -8.82  < 2e-16 ***
## Year2007           -0.42804    0.04672   -9.16  < 2e-16 ***
## Year2008           -0.39199    0.04703   -8.34  < 2e-16 ***
## Year2009           -0.43192    0.04734   -9.12  < 2e-16 ***
## Year2010           -0.44145    0.04760   -9.27  < 2e-16 ***
## Year2011           -0.49635    0.04727  -10.50  < 2e-16 ***
## Year2012           -0.47488    0.04782   -9.93  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.581
## Multiple R-squared:  0.0363, Adjusted R-squared:  0.0355
## Convergence in 22 IRWLS iterations
##
## Robustness weights:
##  6 observations c(492,507,541,610,926,1064)
##   are outliers with |weight| = 0 ( < 4.9e-06);
##  1735 weights are ~= 1. The remaining 18736 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0085  0.8430  0.9490  0.8910  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          4.88e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
```

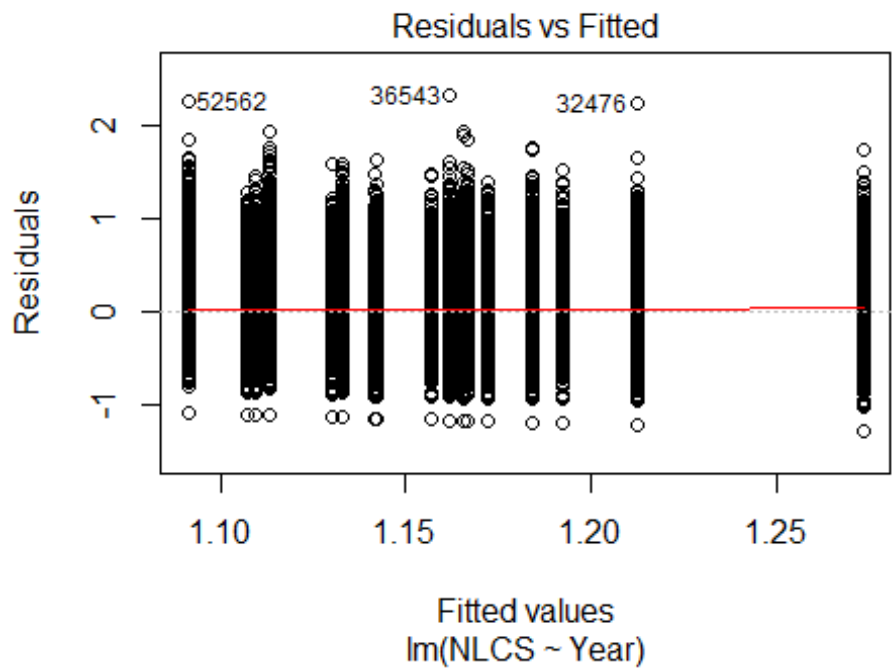
```
##                    GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.007  1           1.003
## Year             1.007 16           1.000
```



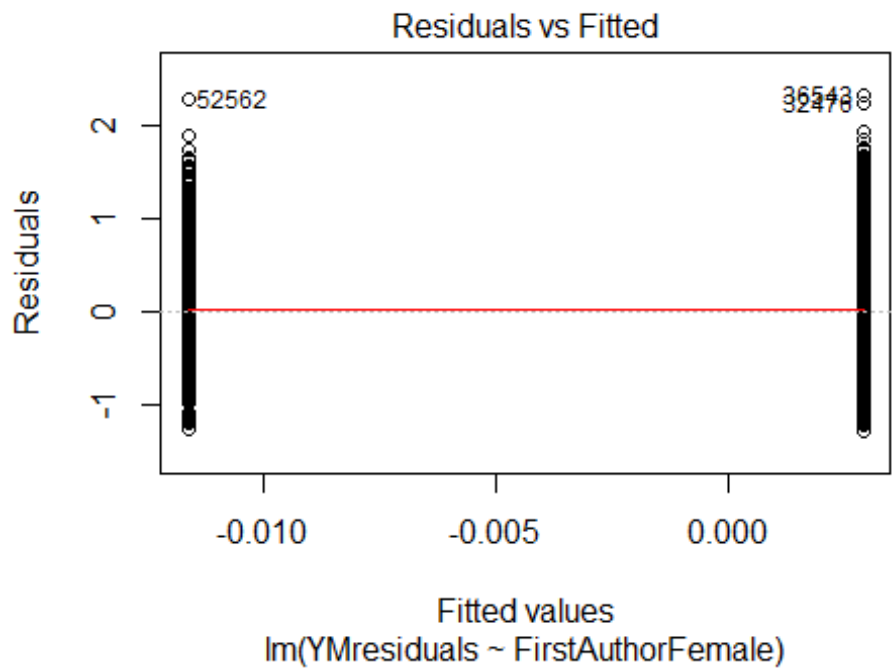

```
## [1] "List of  9 outliers with residuals above  2.5"
##          ScopusId  NLCS Year OneField Fields residuals
## 3439   0029757496 4.541 1996     2731      1     2.898
## 3451   0029761332 4.196 1996     2731      1     2.556
## 3521   0029816808 4.380 1996     2731      1     2.747
## 3680   0029915446 4.649 1996     2731      1     3.009
## 3950   0030067257 4.642 1996     2731      1     3.009
## 4545   0029980777 4.398 1996     2731      3     2.758
## 5628   0030779611 4.287 1997     2731      2     2.765
## 5635   0030832881 4.070 1997     2731      2     2.548
## 40505 84863320414 3.758 2012     2731      1     2.593
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.6407 -0.3933  0.0383  0.4019  3.0083
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
```

```
## (Intercept)        1.64067    0.04452   36.85  < 2e-16 ***
## LastAuthorFemale1 -0.00645    0.01132   -0.57    0.569
## Year1997          -0.11775    0.05346   -2.20    0.028 *
## Year1998          -0.38819    0.04758   -8.16  3.6e-16 ***
## Year1999          -0.42877    0.04712   -9.10  < 2e-16 ***
## Year2000          -0.58234    0.05178  -11.25  < 2e-16 ***
## Year2001          -0.39492    0.04901   -8.06  8.2e-16 ***
## Year2002          -0.48728    0.04870  -10.01  < 2e-16 ***
## Year2003          -0.52659    0.04925  -10.69  < 2e-16 ***
## Year2004          -0.43772    0.04722   -9.27  < 2e-16 ***
## Year2005          -0.40100    0.04720   -8.50  < 2e-16 ***
## Year2006          -0.41083    0.04662   -8.81  < 2e-16 ***
## Year2007          -0.42766    0.04673   -9.15  < 2e-16 ***
## Year2008          -0.39161    0.04703   -8.33  < 2e-16 ***
## Year2009          -0.43142    0.04734   -9.11  < 2e-16 ***
## Year2010          -0.44093    0.04761   -9.26  < 2e-16 ***
## Year2011          -0.49592    0.04728  -10.49  < 2e-16 ***
## Year2012          -0.47449    0.04783   -9.92  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.581
## Multiple R-squared:  0.0364, Adjusted R-squared:  0.0355
## Convergence in 22 IRWLS iterations
##
## Robustness weights:
##  6 observations c(492,507,541,610,926,1064)
##   are outliers with |weight| = 0 ( < 4.9e-06);
##  1733 weights are ~= 1. The remaining 18738 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0088  0.8430  0.9490  0.8910  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          4.88e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Sample size for the above analysis:  20477"
## [1] ""
## [1] ""
```

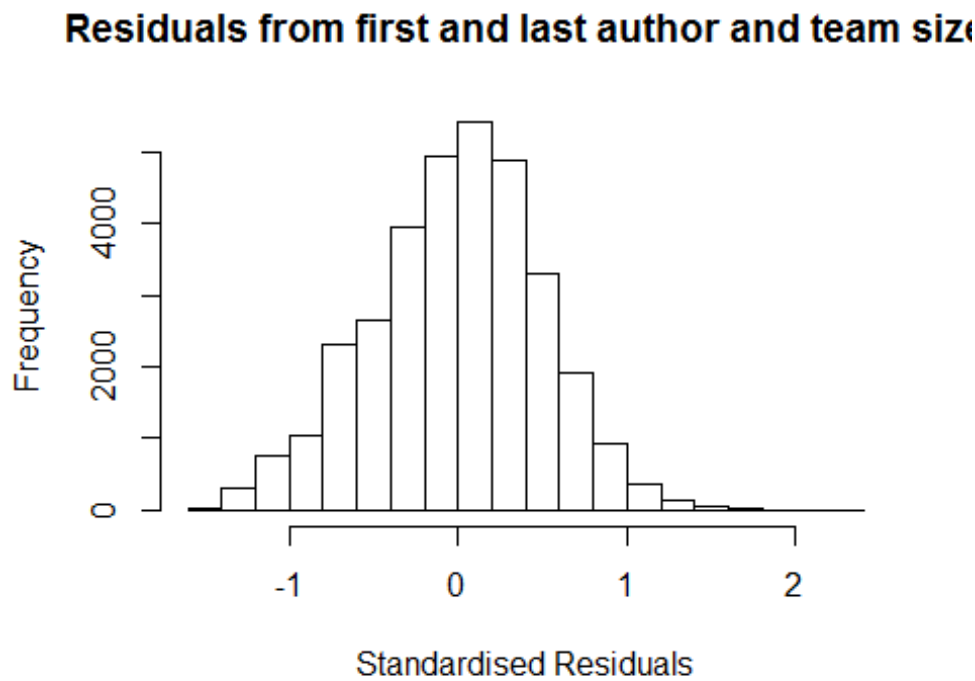
```
## [1] "##################################"
## [1] "Analysis of AJSC 2732"
## [1] "##################################"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 2302 2358 2426 2069 2354 2628 2401 2119 2441 2617 2866 3198 3440 3565 3814
## 2011 2012
## 3864 4123
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 1570 1534 1581 1416 1507 1496 1916 1672 1980 2115 2287 2586 2817 2917 3097
## 2011 2012
## 3160 3326
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 1418 1381 1424 1270 1355 1364 1719 1489 1756 1887 2014 2304 2497 2596 2777
## 2011 2012
## 2825 2984
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
##  Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 250, df = 16, p-value <2e-16
```




```
##
##  Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 0.83, df = 1, p-value = 0.4
```

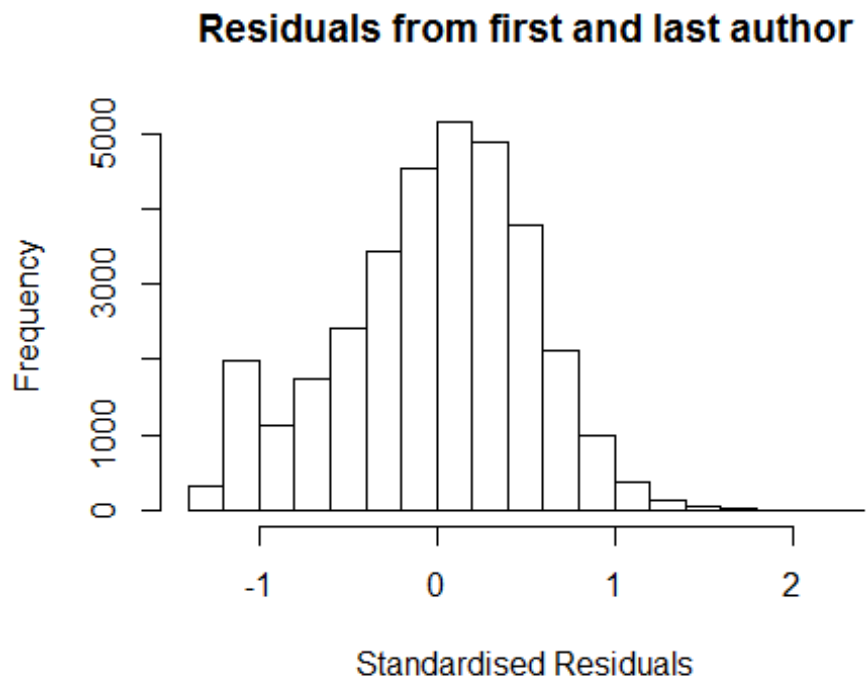



```
## [1] "Female first author team size 2018 geometric mean: 4.32556177780545"
## [1] "Male first author team size 2018 geometric mean: 4.5203268838237"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 1300000, p-value = 0.01
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 4.02088339927396"
## [1] "Male last author team size 2018 geometric mean: 4.56857102876394"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 9e+05, p-value = 8e-08
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##                     GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.045  1           1.022
## LastAuthorFemale  1.039  1           1.020
## UniqueAuthors     1.027  4           1.003
## Year              1.039 16           1.001
```



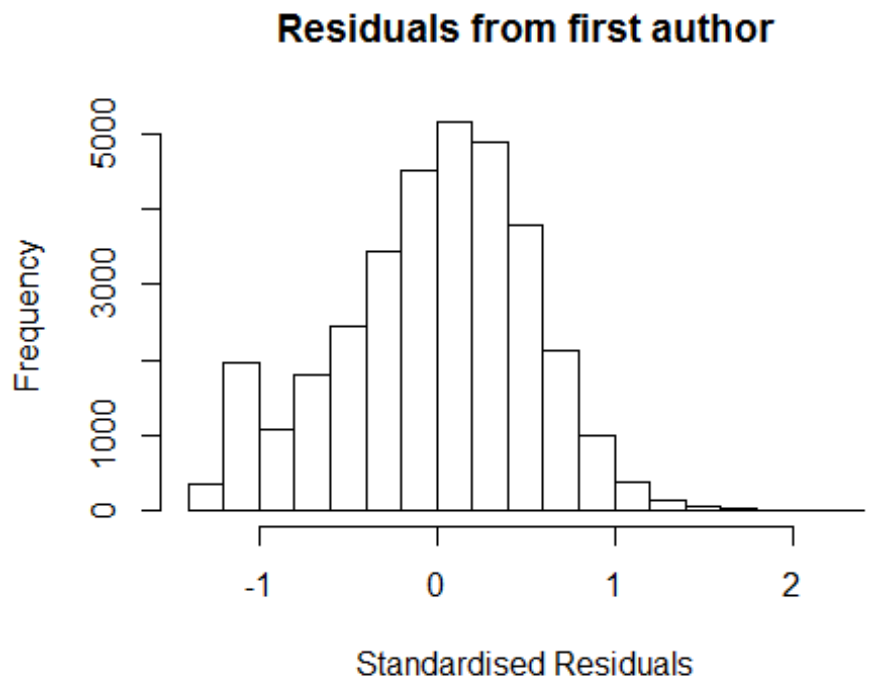

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
UniqueAuthors +
##     Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##     k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.5545 -0.3369  0.0207  0.3329  2.2300
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         0.76548    0.01693   45.22  < 2e-16 ***
## FirstAuthorFemale1 -0.01067    0.00730   -1.46   0.1436
## LastAuthorFemale1  -0.00611    0.00772   -0.79   0.4286
## UniqueAuthors2      0.34128    0.01245   27.41  < 2e-16 ***
## UniqueAuthors3      0.46786    0.01198   39.07  < 2e-16 ***
## UniqueAuthors4      0.57544    0.01188   48.44  < 2e-16 ***
## UniqueAuthors5      0.70024    0.01099   63.70  < 2e-16 ***
## Year1997            0.08879    0.02024    4.39  1.2e-05 ***
## Year1998           -0.01701    0.02005   -0.85   0.3960
## Year1999           -0.02165    0.01992   -1.09   0.2771
```

```
## Year2000           -0.03392    0.01953   -1.74   0.0824 .
## Year2001           -0.04219    0.01917   -2.20   0.0278 *
## Year2002           -0.08762    0.01819   -4.82  1.5e-06 ***
## Year2003           -0.11191    0.01881   -5.95  2.7e-09 ***
## Year2004           -0.12186    0.01813   -6.72  1.8e-11 ***
## Year2005           -0.10199    0.01793   -5.69  1.3e-08 ***
## Year2006           -0.10924    0.01770   -6.17  6.8e-10 ***
## Year2007           -0.04725    0.01773   -2.66   0.0077 **
## Year2008           -0.07668    0.01770   -4.33  1.5e-05 ***
## Year2009           -0.08010    0.01793   -4.47  8.0e-06 ***
## Year2010           -0.11215    0.01739   -6.45  1.1e-10 ***
## Year2011           -0.14470    0.01757   -8.24  < 2e-16 ***
## Year2012           -0.18126    0.01778  -10.19  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.493
## Multiple R-squared:  0.168,  Adjusted R-squared:  0.167
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
##  2758 weights are ~= 1. The remaining 30302 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0046  0.8630  0.9500  0.9040  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          3.02e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.021  1           1.010
## LastAuthorFemale  1.016  1           1.008
## Year              1.012 16           1.000
```



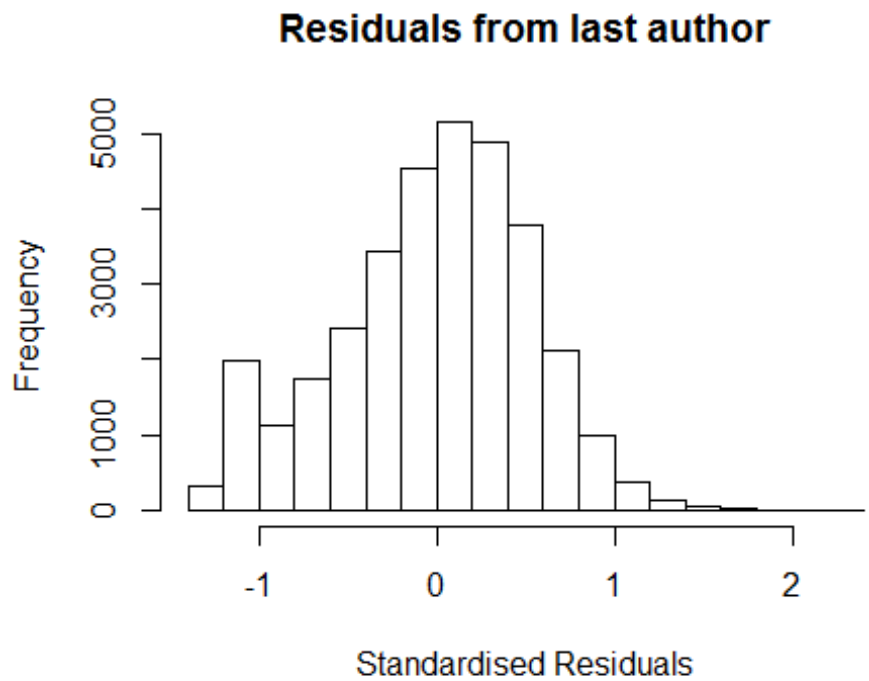

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.3102 -0.3603  0.0397  0.3633  2.3257
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)          1.20573    0.01585   76.05  < 2e-16 ***
## FirstAuthorFemale1 -0.00811    0.00785   -1.03  0.30115
## LastAuthorFemale1  -0.03065    0.00838   -3.66  0.00026 ***
## Year1997            0.10448    0.02262    4.62  3.9e-06 ***
## Year1998           -0.01197    0.02212   -0.54  0.58827
## Year1999            0.00869    0.02195    0.40  0.69203
## Year2000           -0.02151    0.02124   -1.01  0.31124
## Year2001           -0.02494    0.02106   -1.18  0.23634
## Year2002           -0.04668    0.02027   -2.30  0.02129 *
## Year2003           -0.08061    0.02099   -3.84  0.00012 ***
## Year2004           -0.07711    0.02006   -3.84  0.00012 ***
## Year2005           -0.04268    0.01982   -2.15  0.03133 *
```

```
## Year2006           -0.04049    0.01937   -2.09  0.03660 *
## Year2007            0.02788    0.01935    1.44  0.14965
## Year2008           -0.01474    0.01960   -0.75  0.45206
## Year2009           -0.01368    0.01989   -0.69  0.49155
## Year2010           -0.04490    0.01942   -2.31  0.02080 *
## Year2011           -0.07246    0.01939   -3.74  0.00019 ***
## Year2012           -0.10307    0.01972   -5.23  1.7e-07 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.521
## Multiple R-squared:  0.00753,    Adjusted R-squared:  0.00698
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  2602 weights are ~= 1. The remaining 30458 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0087  0.8620  0.9480  0.8990  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          3.02e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.009  1           1.005
## Year              1.009 16           1.000
```




```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.3072 -0.3587  0.0384  0.3642  2.2993
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)          1.20259    0.01583   75.99  < 2e-16 ***
## FirstAuthorFemale1 -0.01422    0.00787   -1.81  0.07096 .
## Year1997             0.10461    0.02262    4.62  3.8e-06 ***
## Year1998            -0.01245    0.02211   -0.56  0.57343
## Year1999             0.00764    0.02192    0.35  0.72739
## Year2000            -0.02198    0.02123   -1.04  0.30043
## Year2001            -0.02594    0.02105   -1.23  0.21785
## Year2002            -0.04752    0.02027   -2.34  0.01906 *
## Year2003            -0.08176    0.02100   -3.89  9.9e-05 ***
## Year2004            -0.07788    0.02006   -3.88  0.00010 ***
## Year2005            -0.04370    0.01981   -2.21  0.02741 *
## Year2006            -0.04137    0.01936   -2.14  0.03257 *
```

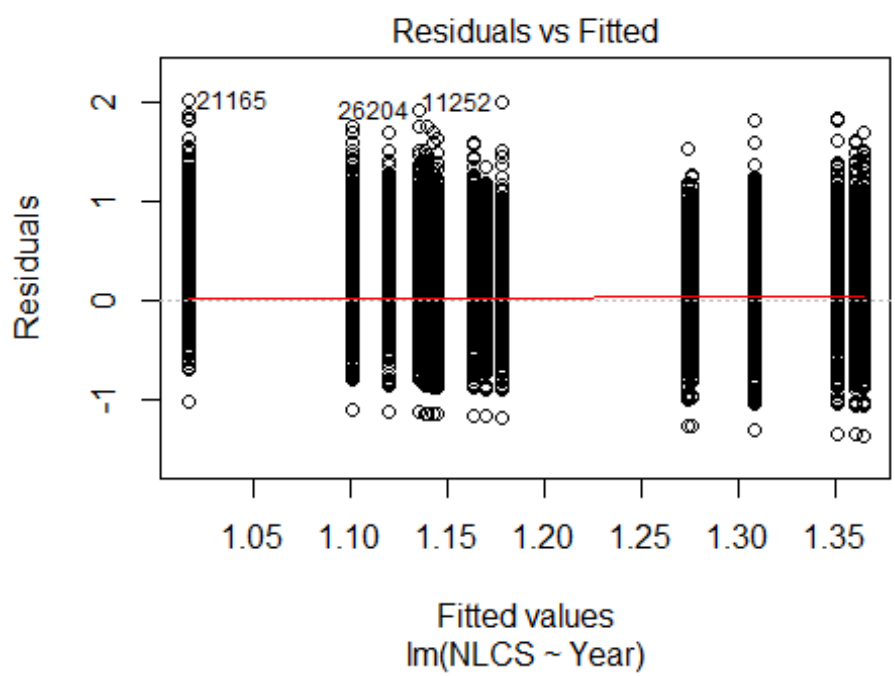
```
## Year2007            0.02728    0.01934    1.41  0.15850
## Year2008           -0.01587    0.01959   -0.81  0.41793
## Year2009           -0.01459    0.01988   -0.73  0.46315
## Year2010           -0.04594    0.01941   -2.37  0.01796 *
## Year2011           -0.07399    0.01937   -3.82  0.00013 ***
## Year2012           -0.10378    0.01971   -5.27  1.4e-07 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.521
## Multiple R-squared:  0.00709,    Adjusted R-squared:  0.00658
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  2662 weights are ~= 1. The remaining 30398 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.013   0.862   0.948   0.899   0.985   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          3.02e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##                   GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.004  1           1.002
## Year             1.004 16           1.000
```



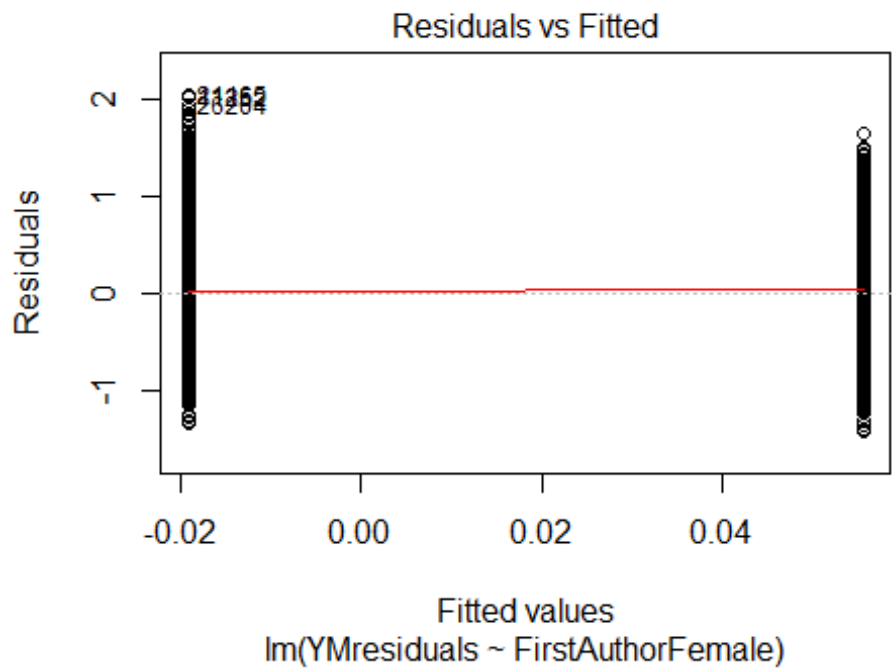

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.3094 -0.3599  0.0389  0.3636  2.3288
##
## Coefficients:
##                   Estimate Std. Error t value Pr(>|t|)
## (Intercept)        1.20479    0.01584   76.07  < 2e-16 ***
## LastAuthorFemale1 -0.03248    0.00839   -3.87  0.00011 ***
## Year1997           0.10458    0.02263    4.62  3.8e-06 ***
## Year1998          -0.01194    0.02212   -0.54  0.58935
## Year1999           0.00864    0.02196    0.39  0.69379
## Year2000          -0.02158    0.02125   -1.02  0.30993
## Year2001          -0.02503    0.02106   -1.19  0.23469
## Year2002          -0.04687    0.02028   -2.31  0.02083 *
## Year2003          -0.08109    0.02099   -3.86  0.00011 ***
## Year2004          -0.07750    0.02006   -3.86  0.00011 ***
## Year2005          -0.04302    0.01983   -2.17  0.03003 *
## Year2006          -0.04092    0.01937   -2.11  0.03460 *
```

```
## Year2007           0.02748    0.01935    1.42  0.15567
## Year2008          -0.01513    0.01961   -0.77  0.44019
## Year2009          -0.01438    0.01988   -0.72  0.46933
## Year2010          -0.04541    0.01941   -2.34  0.01932 *
## Year2011          -0.07313    0.01938   -3.77  0.00016 ***
## Year2012          -0.10379    0.01970   -5.27  1.4e-07 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.521
## Multiple R-squared:  0.00749,    Adjusted R-squared:  0.00698
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  2642 weights are ~= 1. The remaining 30418 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0083  0.8620  0.9480  0.8990  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          3.02e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Sample size for the above analysis:  33060"
## [1] ""
## [1] ""
## [1] "####################################"
## [1] "Analysis of AJSC 2733"
## [1] "####################################"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 1436 1421 1507 1538 1638 1535 1407 1184 1181 1177 1281 1455 1469 1856 1593
## 2011 2012
## 1786 1607
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  908  979  959  967 1124  903 1169  984  973  976 1085 1191 1214 1519 1301
## 2011 2012
```

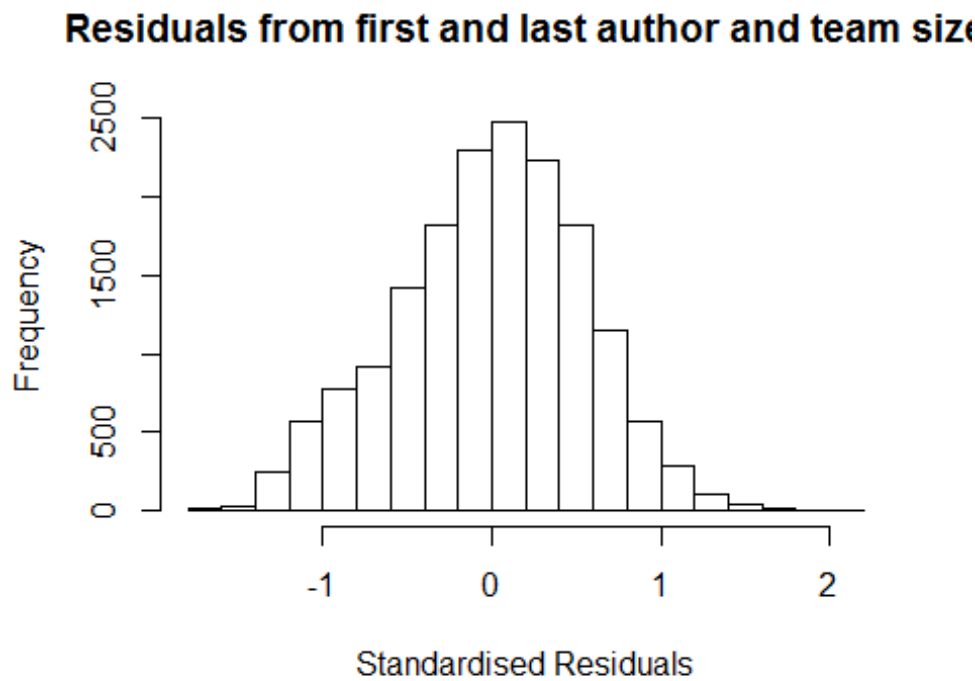
```
## 1480 1312
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  804  866  849  847 1015  791 1025  880  851  860  975 1052 1078 1332 1161
## 2011 2012
## 1295 1130
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
##  Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 88, df = 16, p-value = 5e-12
```




```
##
##  Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 8.3, df = 1, p-value = 0.004
```

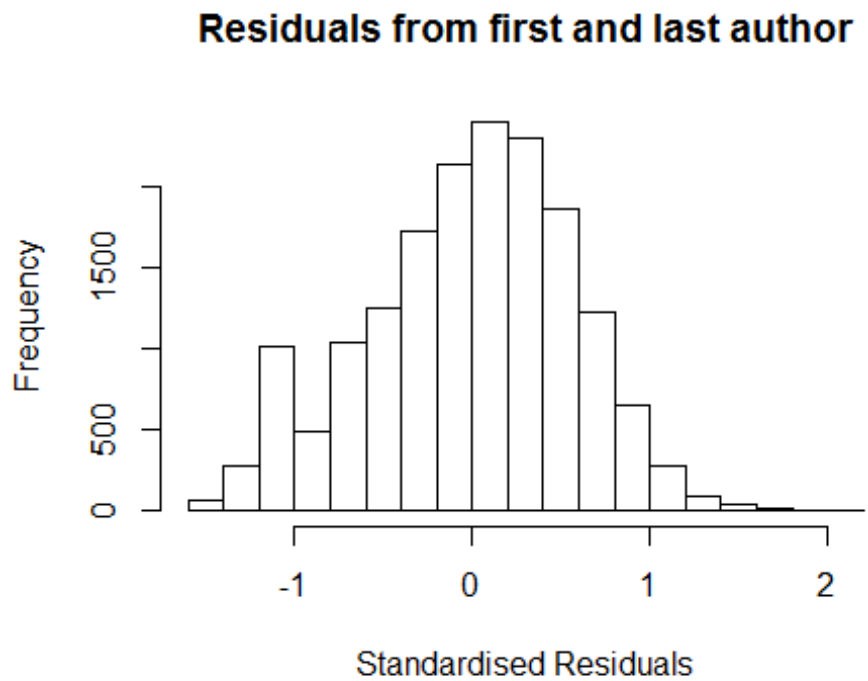



```
## [1] "Female first author team size 2018 geometric mean: 4.20051582103467"
## [1] "Male first author team size 2018 geometric mean: 4.07849729751915"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 240000, p-value = 0.5
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 3.95026476645819"
## [1] "Male last author team size 2018 geometric mean: 4.1871265707537"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 190000, p-value = 0.03
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.069  1           1.034
## LastAuthorFemale  1.062  1           1.031
## UniqueAuthors     1.035  4           1.004
## Year              1.054 16           1.002
```



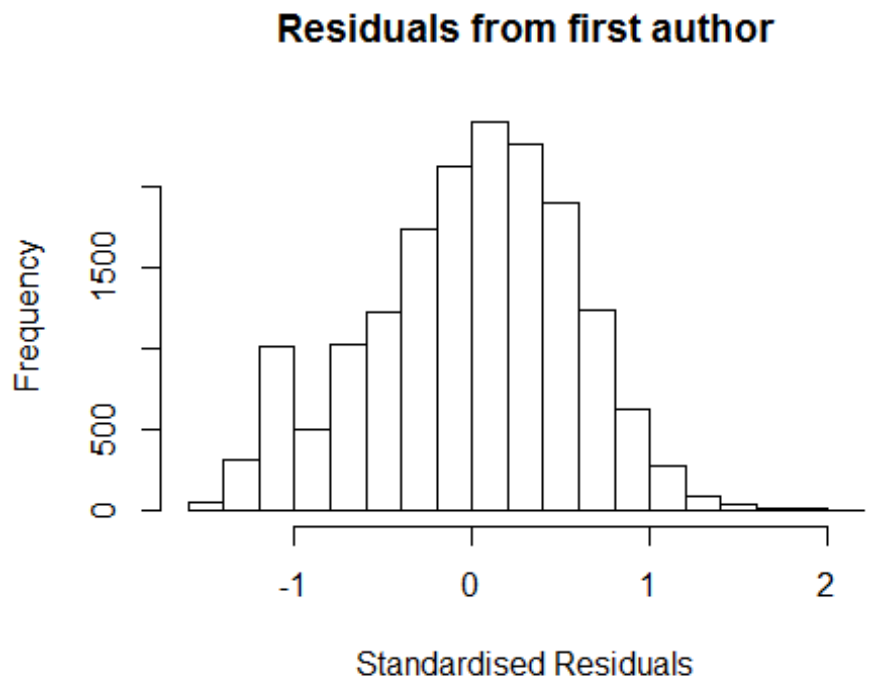

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
UniqueAuthors +
##     Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##     k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.6530 -0.3763  0.0256  0.3773  2.0708
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.07143    0.02601   41.19  < 2e-16 ***
## FirstAuthorFemale1  0.05274    0.01025    5.14  2.7e-07 ***
## LastAuthorFemale1   0.07122    0.01107    6.43  1.3e-10 ***
## UniqueAuthors2      0.16767    0.01967    8.53  < 2e-16 ***
## UniqueAuthors3      0.23582    0.01914   12.32  < 2e-16 ***
## UniqueAuthors4      0.35379    0.01926   18.37  < 2e-16 ***
## UniqueAuthors5      0.54148    0.01830   29.58  < 2e-16 ***
## Year1997            0.00781    0.02887    0.27   0.7867
## Year1998           -0.00221    0.02744   -0.08   0.9357
## Year1999           -0.03115    0.02828   -1.10   0.2707
```

```
## Year2000           -0.08477    0.02690   -3.15   0.0016 **
## Year2001           -0.08681    0.02734   -3.18   0.0015 **
## Year2002           -0.19709    0.02691   -7.32  2.5e-13 ***
## Year2003           -0.20391    0.02827   -7.21  5.7e-13 ***
## Year2004           -0.24837    0.02865   -8.67  < 2e-16 ***
## Year2005           -0.24814    0.02811   -8.83  < 2e-16 ***
## Year2006           -0.25598    0.02719   -9.41  < 2e-16 ***
## Year2007           -0.25814    0.02682   -9.63  < 2e-16 ***
## Year2008           -0.25492    0.02640   -9.66  < 2e-16 ***
## Year2009           -0.38231    0.02807  -13.62  < 2e-16 ***
## Year2010           -0.29542    0.02666  -11.08  < 2e-16 ***
## Year2011           -0.34749    0.02632  -13.20  < 2e-16 ***
## Year2012           -0.30679    0.02727  -11.25  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.548
## Multiple R-squared:  0.12,   Adjusted R-squared:  0.119
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
##  1385 weights are ~= 1. The remaining 15426 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.123   0.868   0.949   0.906   0.986   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          5.95e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.070  1           1.034
## LastAuthorFemale  1.063  1           1.031
## Year              1.028 16           1.001
```



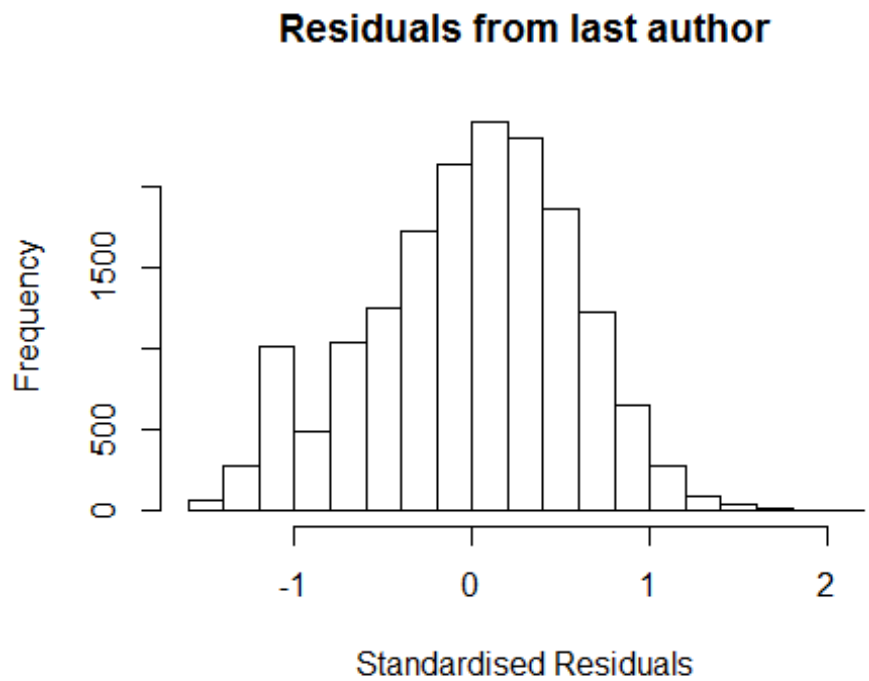

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.4890 -0.3865  0.0337  0.3956  2.0156
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)          1.34169    0.02150   62.40  < 2e-16 ***
## FirstAuthorFemale1   0.06762    0.01071    6.31  2.8e-10 ***
## LastAuthorFemale1    0.06799    0.01152    5.90  3.6e-09 ***
## Year1997             0.00423    0.03006    0.14    0.888
## Year1998             0.01170    0.02874    0.41    0.684
## Year1999            -0.03212    0.02952   -1.09    0.277
## Year2000            -0.06699    0.02786   -2.40    0.016 *
## Year2001            -0.07503    0.02828   -2.65    0.008 **
## Year2002            -0.17625    0.02808   -6.28  3.5e-10 ***
## Year2003            -0.18749    0.02955   -6.34  2.3e-10 ***
## Year2004            -0.24426    0.03052   -8.00  1.3e-15 ***
## Year2005            -0.23594    0.02963   -7.96  1.8e-15 ***
```

```
## Year2006           -0.22987    0.02839   -8.10  6.0e-16 ***
## Year2007           -0.21977    0.02800   -7.85  4.5e-15 ***
## Year2008           -0.20953    0.02760   -7.59  3.3e-14 ***
## Year2009           -0.33685    0.02910  -11.58  < 2e-16 ***
## Year2010           -0.22702    0.02786   -8.15  3.9e-16 ***
## Year2011           -0.27531    0.02735  -10.07  < 2e-16 ***
## Year2012           -0.23856    0.02810   -8.49  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.568
## Multiple R-squared:  0.0356, Adjusted R-squared:  0.0346
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  1349 weights are ~= 1. The remaining 15462 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.181   0.865   0.948   0.904   0.985   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          5.95e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.018  1           1.009
## Year              1.018 16           1.001
```




```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.4452 -0.3826  0.0355  0.3968  2.0068
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)           1.3485     0.0215   62.80  < 2e-16 ***
## FirstAuthorFemale1    0.0831     0.0104    7.95  2.0e-15 ***
## Year1997              0.0071     0.0301    0.24    0.813
## Year1998              0.0135     0.0287    0.47    0.637
## Year1999             -0.0304     0.0295   -1.03    0.303
## Year2000             -0.0642     0.0278   -2.30    0.021 *
## Year2001             -0.0700     0.0282   -2.48    0.013 *
## Year2002             -0.1743     0.0281   -6.21  5.5e-10 ***
## Year2003             -0.1840     0.0296   -6.22  5.1e-10 ***
## Year2004             -0.2407     0.0305   -7.89  3.2e-15 ***
## Year2005             -0.2315     0.0297   -7.80  6.4e-15 ***
## Year2006             -0.2248     0.0284   -7.92  2.4e-15 ***
```

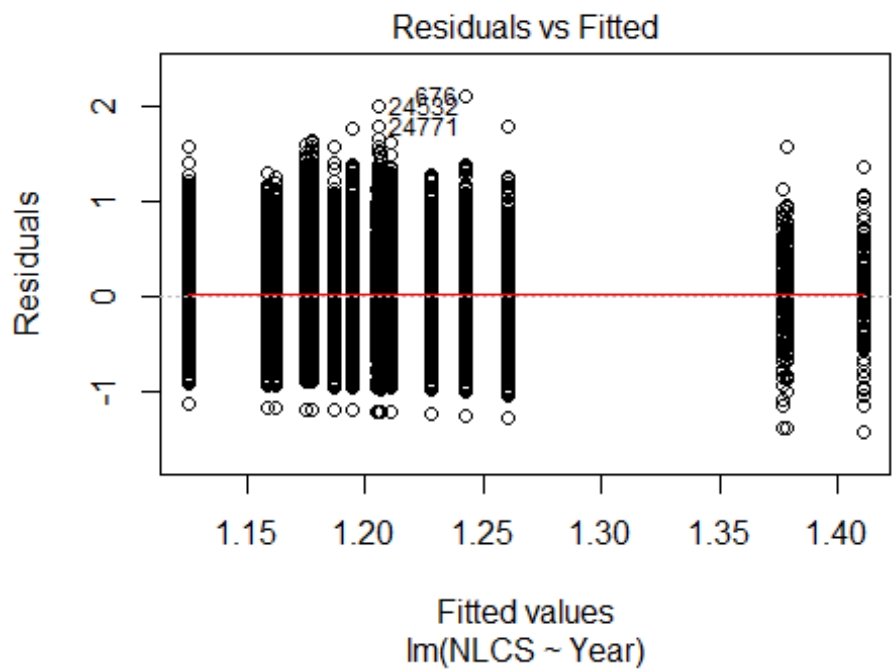
```
## Year2007            -0.2165     0.0280   -7.74  1.0e-14 ***
## Year2008            -0.2057     0.0276   -7.46  8.8e-14 ***
## Year2009            -0.3358     0.0291  -11.54  < 2e-16 ***
## Year2010            -0.2240     0.0279   -8.04  9.4e-16 ***
## Year2011            -0.2718     0.0273   -9.95  < 2e-16 ***
## Year2012            -0.2334     0.0281   -8.32  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.568
## Multiple R-squared:  0.0335, Adjusted R-squared:  0.0325
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  1382 weights are ~= 1. The remaining 15429 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.186   0.865   0.947   0.904   0.985   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          5.95e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##                   GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.012  1           1.006
## Year             1.012 16           1.000
```



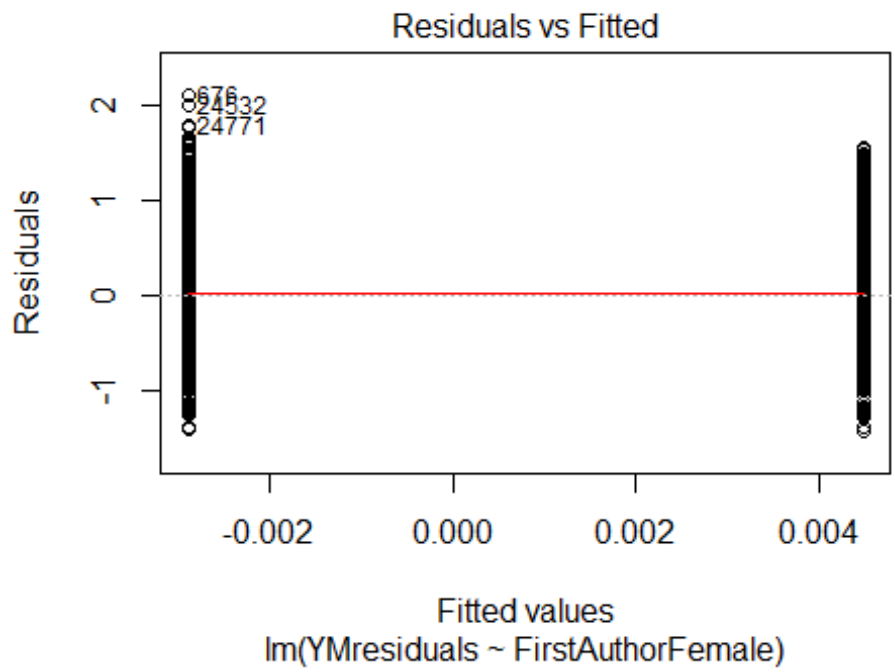

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.4498 -0.3817  0.0401  0.3980  2.0039
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.35365    0.02142   63.18  < 2e-16 ***
## LastAuthorFemale1   0.08598    0.01124    7.65  2.2e-14 ***
## Year1997            0.00121    0.03006    0.04   0.9680
## Year1998            0.01017    0.02878    0.35   0.7239
## Year1999           -0.03387    0.02953   -1.15   0.2513
## Year2000           -0.06731    0.02793   -2.41   0.0160 *
## Year2001           -0.07584    0.02828   -2.68   0.0073 **
## Year2002           -0.17655    0.02816   -6.27  3.7e-10 ***
## Year2003           -0.18624    0.02960   -6.29  3.2e-10 ***
## Year2004           -0.24294    0.03054   -7.96  1.9e-15 ***
## Year2005           -0.23574    0.02966   -7.95  2.0e-15 ***
## Year2006           -0.22804    0.02846   -8.01  1.2e-15 ***
```

```
## Year2007          -0.21593    0.02804   -7.70  1.4e-14 ***
## Year2008          -0.20591    0.02766   -7.44  1.0e-13 ***
## Year2009          -0.33214    0.02912  -11.41  < 2e-16 ***
## Year2010          -0.22093    0.02793   -7.91  2.7e-15 ***
## Year2011          -0.26885    0.02738   -9.82  < 2e-16 ***
## Year2012          -0.23369    0.02818   -8.29  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.569
## Multiple R-squared:  0.0332, Adjusted R-squared:  0.0322
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  1338 weights are ~= 1. The remaining 15473 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.189   0.864   0.948   0.905   0.985   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          5.95e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Sample size for the above analysis:  16811"
## [1] ""
## [1] ""
## [1] "####################################"
## [1] "Analysis of AJSC 2734"
## [1] "####################################"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 1910 1882  181  182  204 2036  207 1622 1693 1695 1889 1999 2074 2192 2437
## 2011 2012
## 2392 2262
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 1370 1400  126  137  115 1061  155 1202 1243 1236 1427 1498 1581 1663 1875
## 2011 2012
```

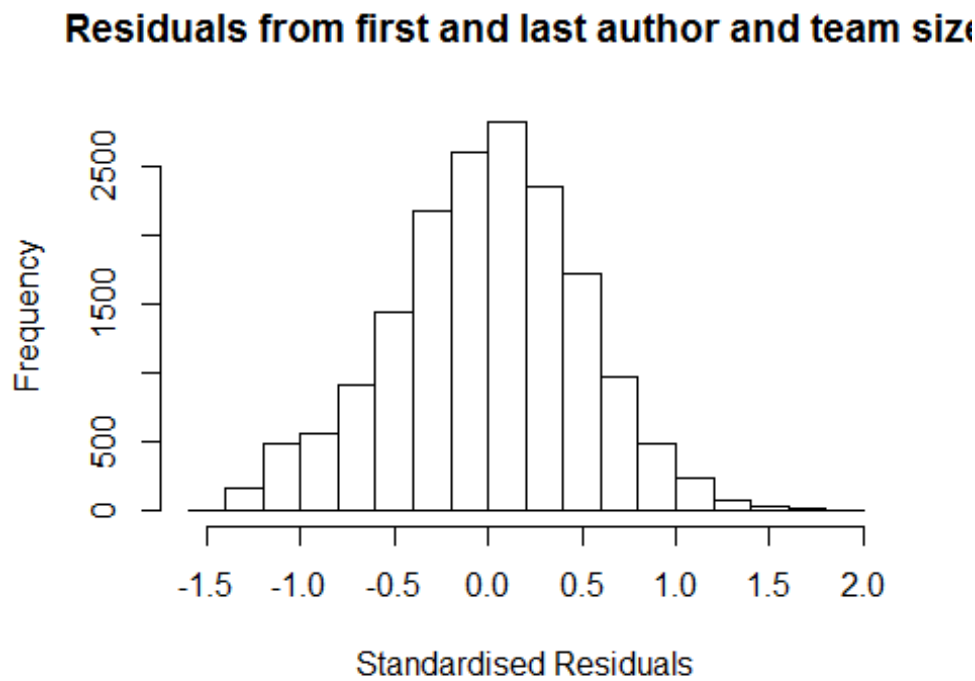
```
## 1860 1764
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 1223 1246  113  124  102  924  140 1047 1039 1036 1247 1281 1368 1432 1586
## 2011 2012
## 1601 1505
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
##  Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 93, df = 16, p-value = 6e-13
```




```
##
##  Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 40, df = 1, p-value = 3e-10
```

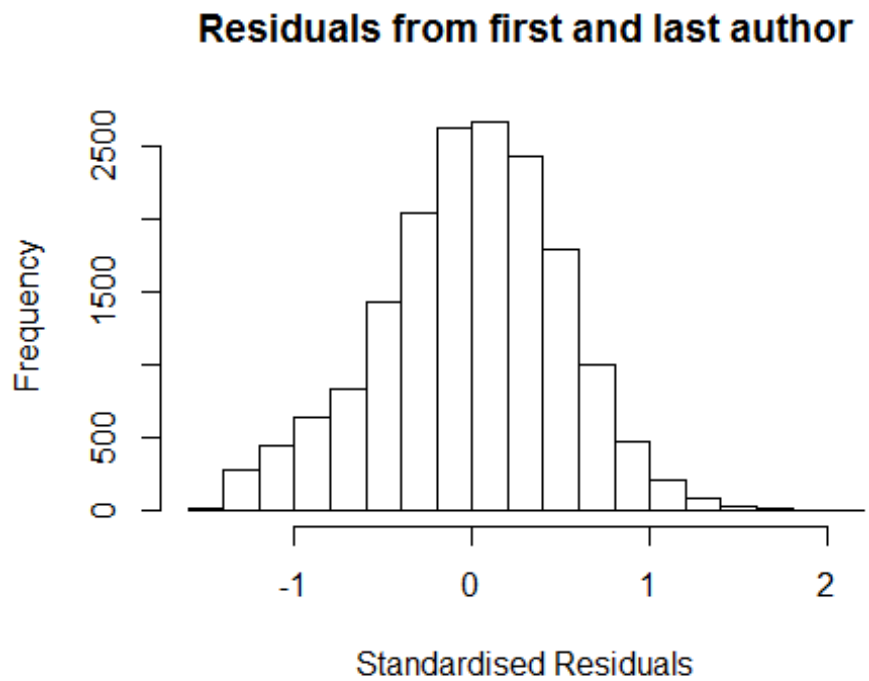



```
## [1] "Female first author team size 2018 geometric mean: 3.69683342855458"
## [1] "Male first author team size 2018 geometric mean: 3.61047274682779"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 3e+05, p-value = 0.6
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 3.50337612940185"
## [1] "Male last author team size 2018 geometric mean: 3.75042924282853"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 270000, p-value = 0.1
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.038  1           1.019
## LastAuthorFemale  1.031  1           1.015
## UniqueAuthors     1.040  4           1.005
## Year              1.060 16           1.002
```



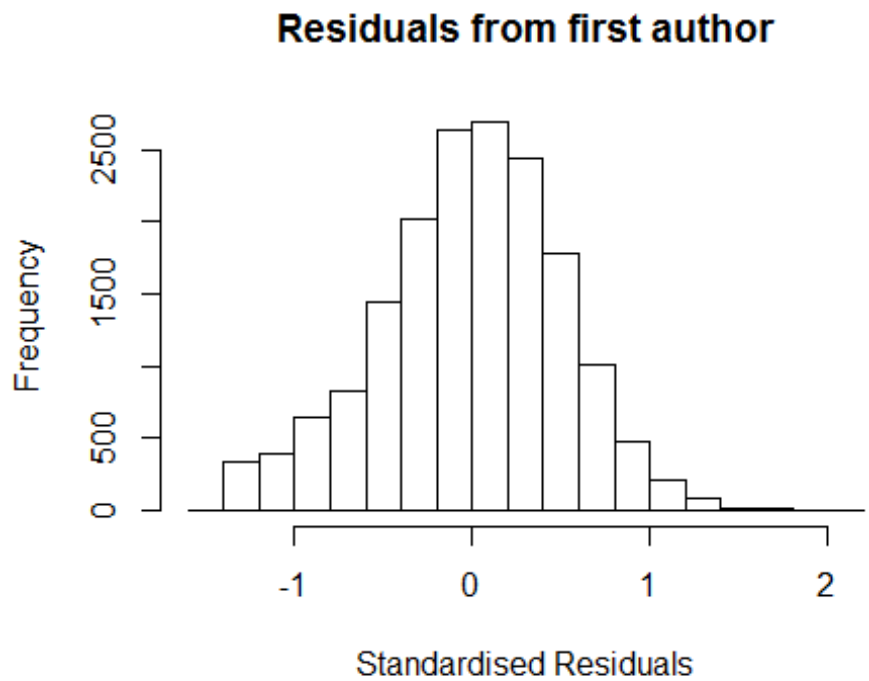

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
UniqueAuthors +
##     Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##     k.max = 1000))
##  \--> method = "MM"
## Residuals:
##    Min     1Q Median     3Q    Max
## -1.461 -0.324  0.013  0.330  1.956
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.06682    0.02010   53.08  < 2e-16 ***
## FirstAuthorFemale1 -0.00657    0.00810   -0.81  0.41713
## LastAuthorFemale1  -0.01706    0.00881   -1.94  0.05283 .
## UniqueAuthors2      0.11683    0.01683    6.94  4.0e-12 ***
## UniqueAuthors3      0.19323    0.01622   11.91  < 2e-16 ***
## UniqueAuthors4      0.20077    0.01628   12.33  < 2e-16 ***
## UniqueAuthors5      0.33313    0.01459   22.83  < 2e-16 ***
## Year1997           -0.00167    0.02087   -0.08  0.93630
## Year1998           -0.03669    0.05330   -0.69  0.49125
## Year1999            0.20794    0.05133    4.05  5.1e-05 ***
```

```
## Year2000            0.16805    0.05374    3.13  0.00177 **
## Year2001           -0.02255    0.02136   -1.06  0.29106
## Year2002            0.12938    0.04494    2.88  0.00400 **
## Year2003           -0.12632    0.02238   -5.64  1.7e-08 ***
## Year2004           -0.09869    0.02167   -4.55  5.3e-06 ***
## Year2005           -0.09196    0.02161   -4.25  2.1e-05 ***
## Year2006           -0.05396    0.02057   -2.62  0.00872 **
## Year2007           -0.07379    0.02037   -3.62  0.00029 ***
## Year2008           -0.05637    0.02075   -2.72  0.00660 **
## Year2009           -0.04732    0.02074   -2.28  0.02255 *
## Year2010           -0.05764    0.02065   -2.79  0.00527 **
## Year2011           -0.08606    0.02070   -4.16  3.2e-05 ***
## Year2012           -0.10625    0.02126   -5.00  5.9e-07 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.488
## Multiple R-squared:  0.0563, Adjusted R-squared:  0.0551
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
##  1430 weights are ~= 1. The remaining 15584 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0718  0.8670  0.9510  0.9010  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          5.88e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.030  1           1.015
## LastAuthorFemale  1.026  1           1.013
## Year              1.030 16           1.001
```



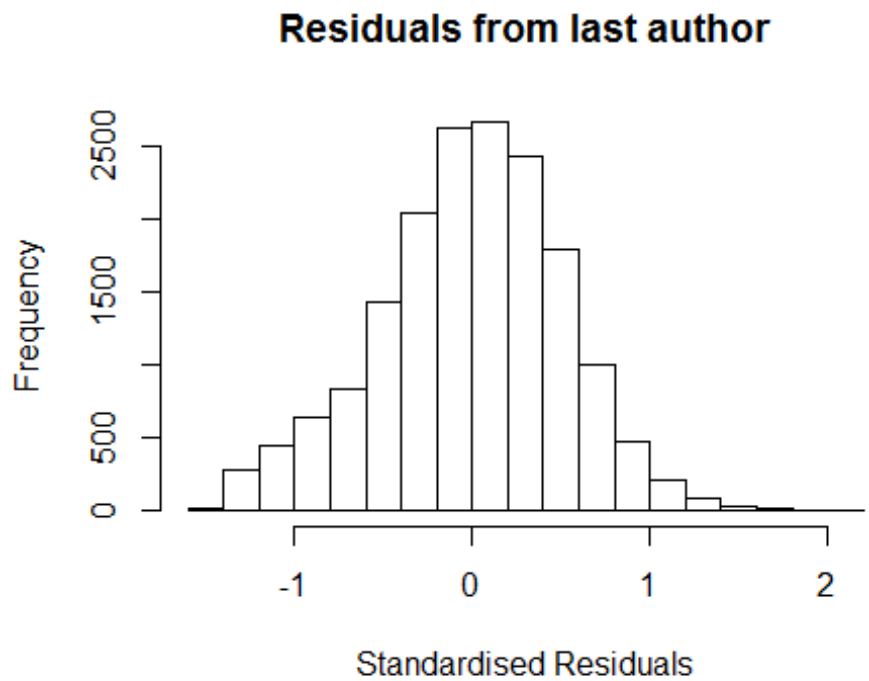

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.4411 -0.3295  0.0134  0.3395  2.0913
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)          1.25770    0.01608   78.23  < 2e-16 ***
## FirstAuthorFemale1   0.01241    0.00827    1.50  0.13341
## LastAuthorFemale1   -0.02911    0.00902   -3.23  0.00126 **
## Year1997            -0.00916    0.02144   -0.43  0.66922
## Year1998            -0.05381    0.05260   -1.02  0.30629
## Year1999             0.17095    0.04795    3.57  0.00036 ***
## Year2000             0.15776    0.05097    3.09  0.00197 **
## Year2001            -0.00161    0.02189   -0.07  0.94134
## Year2002             0.13091    0.04170    3.14  0.00170 **
## Year2003            -0.11117    0.02320   -4.79  1.7e-06 ***
## Year2004            -0.07979    0.02231   -3.58  0.00035 ***
## Year2005            -0.07618    0.02225   -3.42  0.00062 ***
```

```
## Year2006           -0.04399    0.02103   -2.09  0.03645 *
## Year2007           -0.06011    0.02067   -2.91  0.00365 **
## Year2008           -0.04206    0.02132   -1.97  0.04847 *
## Year2009           -0.03448    0.02120   -1.63  0.10386
## Year2010           -0.04286    0.02109   -2.03  0.04214 *
## Year2011           -0.07069    0.02118   -3.34  0.00085 ***
## Year2012           -0.08715    0.02184   -3.99  6.6e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.498
## Multiple R-squared:  0.0077, Adjusted R-squared:  0.00665
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  1468 weights are ~= 1. The remaining 15546 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0382  0.8680  0.9500  0.9000  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          5.88e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.019  1           1.009
## Year              1.019 16           1.001
```




```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.4320 -0.3303  0.0132  0.3414  2.0965
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)          1.25247    0.01598   78.39  < 2e-16 ***
## FirstAuthorFemale1   0.00789    0.00824    0.96  0.33809
## Year1997            -0.00959    0.02147   -0.45  0.65522
## Year1998            -0.05163    0.05250   -0.98  0.32535
## Year1999             0.17168    0.04813    3.57  0.00036 ***
## Year2000             0.15890    0.05110    3.11  0.00188 **
## Year2001            -0.00156    0.02188   -0.07  0.94334
## Year2002             0.12938    0.04193    3.09  0.00204 **
## Year2003            -0.11185    0.02321   -4.82  1.5e-06 ***
## Year2004            -0.08089    0.02232   -3.62  0.00029 ***
## Year2005            -0.07678    0.02226   -3.45  0.00056 ***
## Year2006            -0.04538    0.02102   -2.16  0.03091 *
```

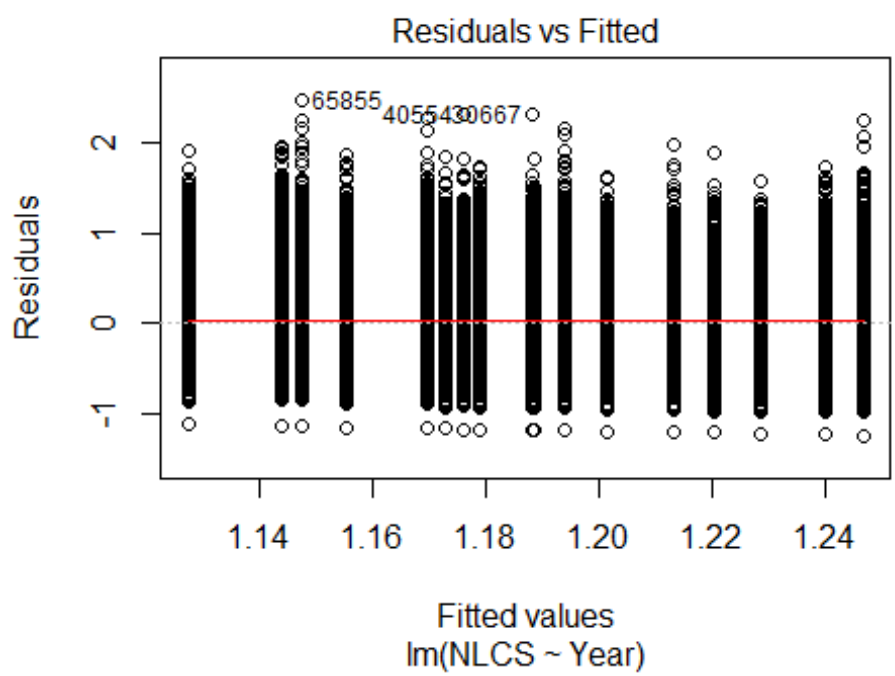
```
## Year2007           -0.06168    0.02068   -2.98  0.00286 **
## Year2008           -0.04361    0.02133   -2.04  0.04095 *
## Year2009           -0.03638    0.02121   -1.72  0.08634 .
## Year2010           -0.04540    0.02108   -2.15  0.03126 *
## Year2011           -0.07291    0.02117   -3.44  0.00057 ***
## Year2012           -0.08976    0.02183   -4.11  3.9e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.498
## Multiple R-squared:  0.00707,    Adjusted R-squared:  0.00608
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  1458 weights are ~= 1. The remaining 15556 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0371  0.8680  0.9500  0.9010  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          5.88e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##                   GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.014  1           1.007
## Year             1.014 16           1.000
```



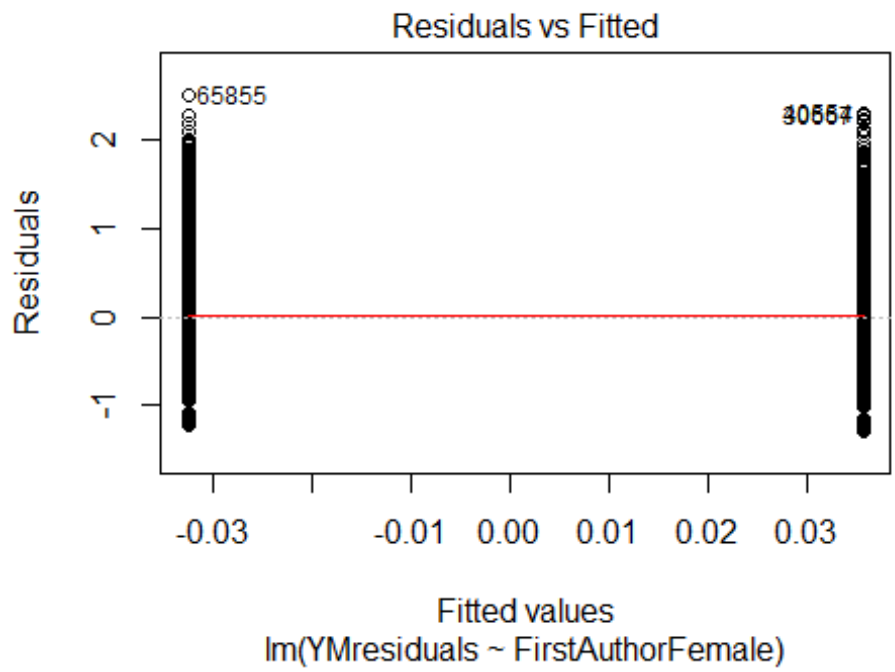

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.4314 -0.3298  0.0145  0.3388  2.0880
##
## Coefficients:
##                   Estimate Std. Error t value Pr(>|t|)
## (Intercept)        1.26102    0.01589   79.36  < 2e-16 ***
## LastAuthorFemale1 -0.02686    0.00898   -2.99  0.00279 **
## Year1997          -0.00889    0.02145   -0.41  0.67854
## Year1998          -0.05415    0.05269   -1.03  0.30409
## Year1999           0.17041    0.04791    3.56  0.00038 ***
## Year2000           0.15795    0.05093    3.10  0.00193 **
## Year2001          -0.00115    0.02188   -0.05  0.95801
## Year2002           0.13118    0.04171    3.15  0.00166 **
## Year2003          -0.11051    0.02319   -4.77  1.9e-06 ***
## Year2004          -0.07910    0.02230   -3.55  0.00039 ***
## Year2005          -0.07564    0.02225   -3.40  0.00068 ***
## Year2006          -0.04324    0.02101   -2.06  0.03964 *
```

```
## Year2007          -0.05897    0.02066   -2.85  0.00432 **
## Year2008          -0.04058    0.02128   -1.91  0.05658 .
## Year2009          -0.03332    0.02119   -1.57  0.11576
## Year2010          -0.04128    0.02105   -1.96  0.04990 *
## Year2011          -0.06881    0.02112   -3.26  0.00113 **
## Year2012          -0.08544    0.02178   -3.92  8.8e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.498
## Multiple R-squared:  0.00757,    Adjusted R-squared:  0.00658
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  1459 weights are ~= 1. The remaining 15555 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0392  0.8680  0.9500  0.9000  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          5.88e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Sample size for the above analysis:  17014"
## [1] ""
## [1] ""
## [1] "####################################"
## [1] "Analysis of AJSC 2735"
## [1] "####################################"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 3722 3483 3469 3190 3614 3727 3298 2798 2804 3022 3974 3846 4435 4886 4773
## 2011 2012
## 5131 5192
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 2304 2045 1973 2105 2337 2255 2645 2257 2331 2475 3286 3138 3621 4018 3954
## 2011 2012
```

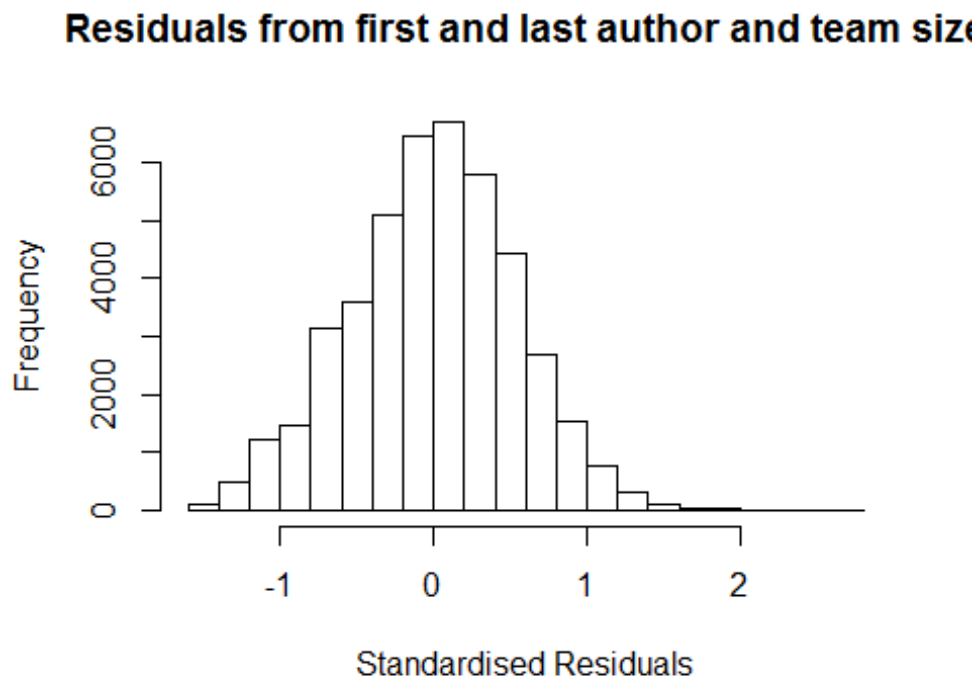
```
## 4211 4277
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 2068 1834 1789 1866 2123 2025 2378 2028 2087 2200 2938 2804 3239 3574 3513
## 2011 2012
## 3726 3826
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
##  Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 230, df = 16, p-value <2e-16
```




```
##
##  Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 1.5, df = 1, p-value = 0.2
```

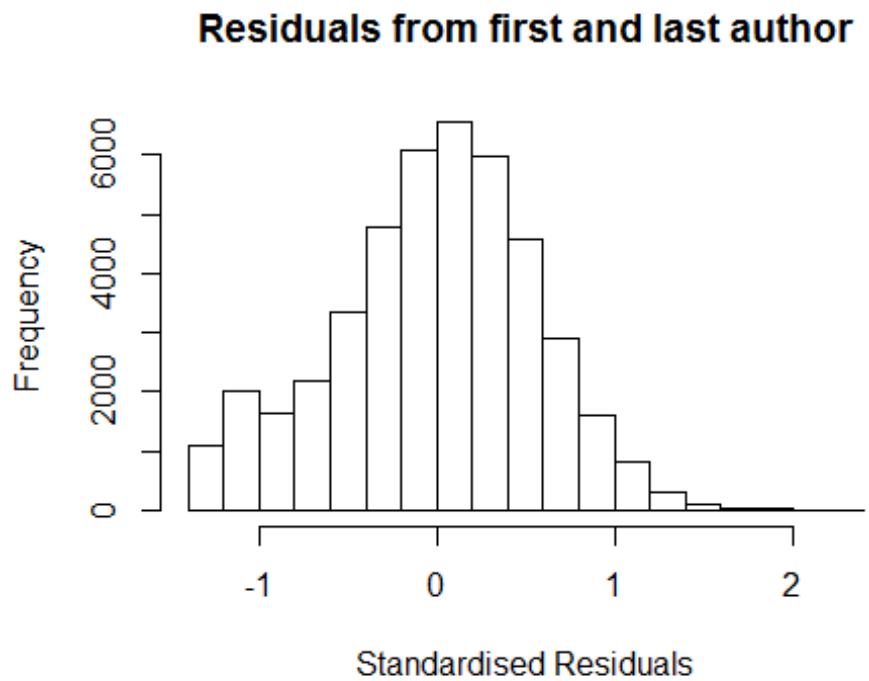



```
## [1] "Female first author team size 2018 geometric mean: 4.31990376857884"
## [1] "Male first author team size 2018 geometric mean: 4.37584172263845"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 2700000, p-value = 0.3
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 4.09872835713461"
## [1] "Male last author team size 2018 geometric mean: 4.56142478803463"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 2600000, p-value = 2e-10
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##                     GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.070  1           1.034
## LastAuthorFemale  1.068  1           1.034
## UniqueAuthors     1.056  4           1.007
## Year              1.065 16           1.002
```



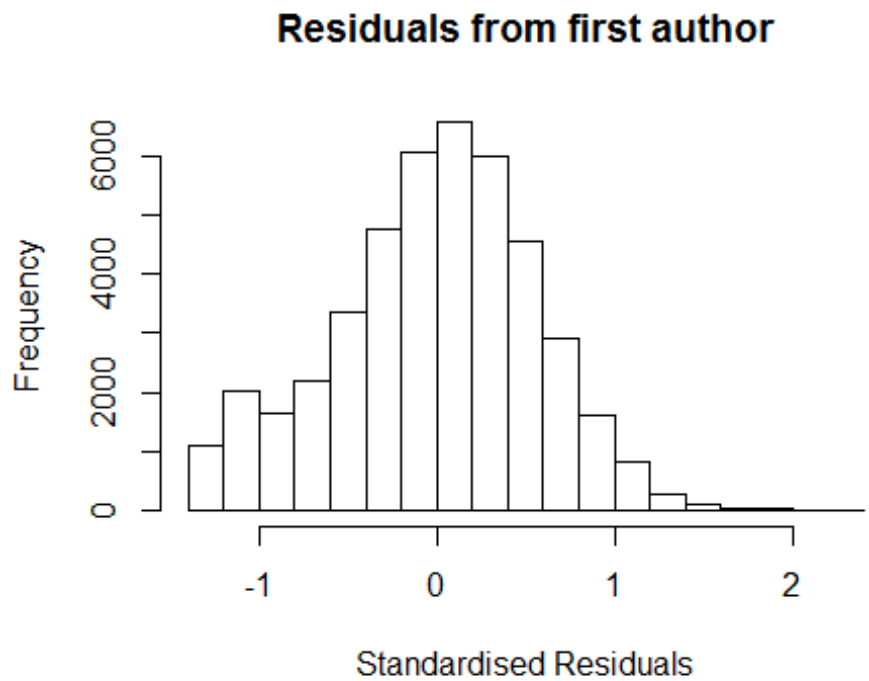

```
## [1] "List of  3 outliers with residuals above  2.5"
##          ScopusId  NLCS Year OneField Fields residuals
## 688    0029800129 3.486 1996     2735      1     2.613
## 8481   0031683919 3.358 1998     2735      1     2.574
## 40554 38449095758 3.502 2007     2735      1     2.715
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
UniqueAuthors +
##     Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##     k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.5678 -0.3541  0.0133  0.3602  2.7152
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         0.87271    0.01555   56.14  < 2e-16 ***
## FirstAuthorFemale1  0.05256    0.00547    9.61  < 2e-16 ***
## LastAuthorFemale1   0.02181    0.00569    3.83  0.00013 ***
## UniqueAuthors2      0.29466    0.01171   25.16  < 2e-16 ***
## UniqueAuthors3      0.40505    0.01107   36.58  < 2e-16 ***
## UniqueAuthors4      0.47814    0.01119   42.71  < 2e-16 ***
## UniqueAuthors5      0.62072    0.01018   60.98  < 2e-16 ***
## Year1997           -0.04119    0.01890   -2.18  0.02935 *
```

```
## Year1998           -0.08830    0.01888   -4.68  2.9e-06 ***
## Year1999           -0.07929    0.01794   -4.42  9.9e-06 ***
## Year2000           -0.05169    0.01736   -2.98  0.00291 **
## Year2001           -0.08249    0.01771   -4.66  3.2e-06 ***
## Year2002           -0.11954    0.01747   -6.84  7.9e-12 ***
## Year2003           -0.15922    0.01721   -9.25  < 2e-16 ***
## Year2004           -0.14237    0.01752   -8.12  4.6e-16 ***
## Year2005           -0.12567    0.01711   -7.34  2.1e-13 ***
## Year2006           -0.16197    0.01647   -9.84  < 2e-16 ***
## Year2007           -0.16031    0.01663   -9.64  < 2e-16 ***
## Year2008           -0.21515    0.01635  -13.16  < 2e-16 ***
## Year2009           -0.18402    0.01626  -11.31  < 2e-16 ***
## Year2010           -0.19434    0.01624  -11.97  < 2e-16 ***
## Year2011           -0.20430    0.01628  -12.55  < 2e-16 ***
## Year2012           -0.23033    0.01648  -13.97  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.531
## Multiple R-squared:  0.125,  Adjusted R-squared:  0.125
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  3 observations c(319,4238,23397)
##   are outliers with |weight| = 0 ( < 2.3e-06);
##  3703 weights are ~= 1. The remaining 40312 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0176  0.8640  0.9500  0.9050  0.9860  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          2.27e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.052  1           1.025
## LastAuthorFemale  1.046  1           1.023
## Year              1.020 16           1.001
```



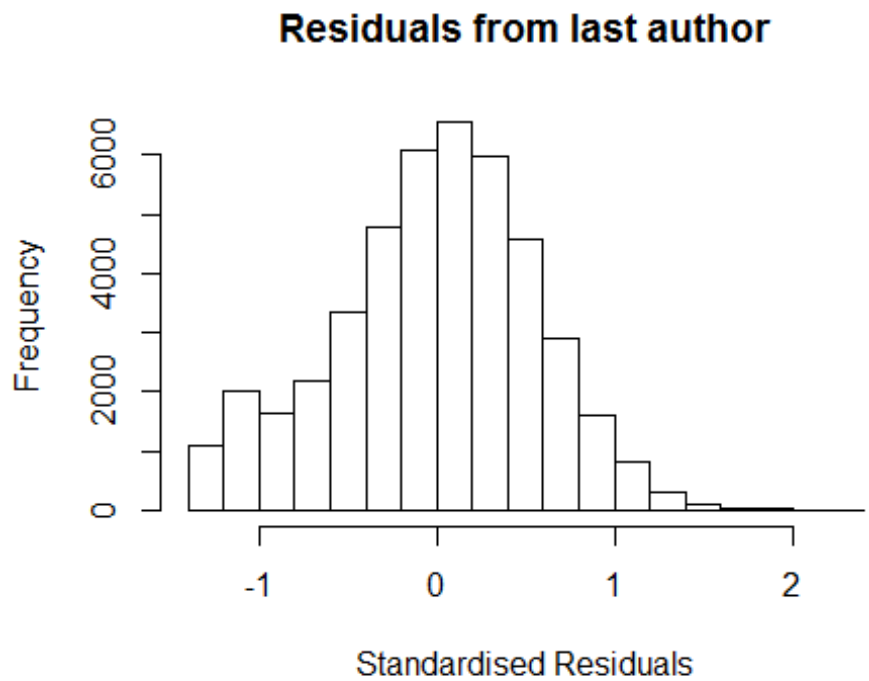

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.3079 -0.3672  0.0281  0.3769  2.2775
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)          1.23295    0.01435   85.92  < 2e-16 ***
## FirstAuthorFemale1  0.07494    0.00577   12.99  < 2e-16 ***
## LastAuthorFemale1  -0.00506    0.00600   -0.84  0.39927
## Year1997           -0.00707    0.02008   -0.35  0.72465
## Year1998           -0.06237    0.02032   -3.07  0.00214 **
## Year1999           -0.03043    0.01890   -1.61  0.10734
## Year2000           -0.01896    0.01842   -1.03  0.30349
## Year2001           -0.04738    0.01878   -2.52  0.01163 *
## Year2002           -0.06831    0.01867   -3.66  0.00025 ***
## Year2003           -0.08580    0.01834   -4.68  2.9e-06 ***
## Year2004           -0.06849    0.01867   -3.67  0.00024 ***
## Year2005           -0.03552    0.01811   -1.96  0.04987 *
```

```
## Year2006           -0.08023    0.01764   -4.55  5.4e-06 ***
## Year2007           -0.07828    0.01788   -4.38  1.2e-05 ***
## Year2008           -0.13351    0.01761   -7.58  3.5e-14 ***
## Year2009           -0.09487    0.01755   -5.41  6.5e-08 ***
## Year2010           -0.08742    0.01747   -5.00  5.6e-07 ***
## Year2011           -0.10900    0.01744   -6.25  4.2e-10 ***
## Year2012           -0.11710    0.01762   -6.65  3.1e-11 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.549
## Multiple R-squared:  0.00761,    Adjusted R-squared:  0.0072
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  3722 weights are ~= 1. The remaining 40296 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0462  0.8600  0.9490  0.9000  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          2.27e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.014  1           1.007
## Year              1.014 16           1.000
```




```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.3056 -0.3670  0.0279  0.3770  2.2750
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)          1.23178    0.01429   86.21  < 2e-16 ***
## FirstAuthorFemale1   0.07378    0.00567   13.00  < 2e-16 ***
## Year1997            -0.00705    0.02008   -0.35  0.72574
## Year1998            -0.06233    0.02032   -3.07  0.00216 **
## Year1999            -0.03047    0.01890   -1.61  0.10699
## Year2000            -0.01883    0.01842   -1.02  0.30685
## Year2001            -0.04742    0.01878   -2.52  0.01158 *
## Year2002            -0.06839    0.01867   -3.66  0.00025 ***
## Year2003            -0.08599    0.01834   -4.69  2.7e-06 ***
## Year2004            -0.06857    0.01867   -3.67  0.00024 ***
## Year2005            -0.03562    0.01811   -1.97  0.04922 *
## Year2006            -0.08041    0.01764   -4.56  5.2e-06 ***
```

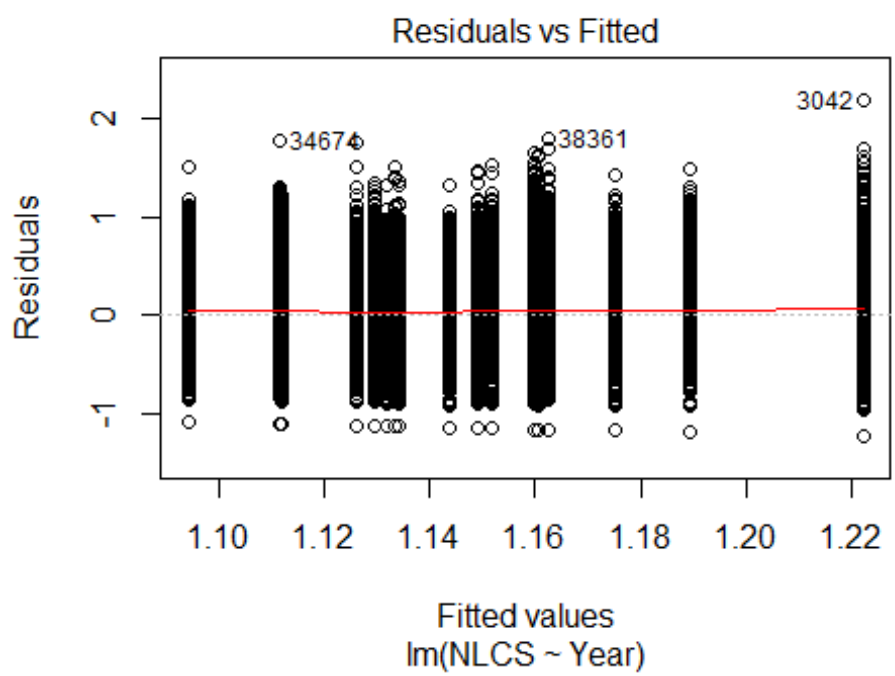
```
## Year2007           -0.07854    0.01787   -4.39  1.1e-05 ***
## Year2008           -0.13363    0.01761   -7.59  3.3e-14 ***
## Year2009           -0.09500    0.01755   -5.41  6.2e-08 ***
## Year2010           -0.08762    0.01746   -5.02  5.3e-07 ***
## Year2011           -0.10927    0.01744   -6.27  3.7e-10 ***
## Year2012           -0.11743    0.01761   -6.67  2.6e-11 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.549
## Multiple R-squared:  0.0076, Adjusted R-squared:  0.00722
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  3729 weights are ~= 1. The remaining 40289 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0469  0.8600  0.9490  0.9000  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          2.27e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##                   GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.009  1           1.004
## Year             1.009 16           1.000
```



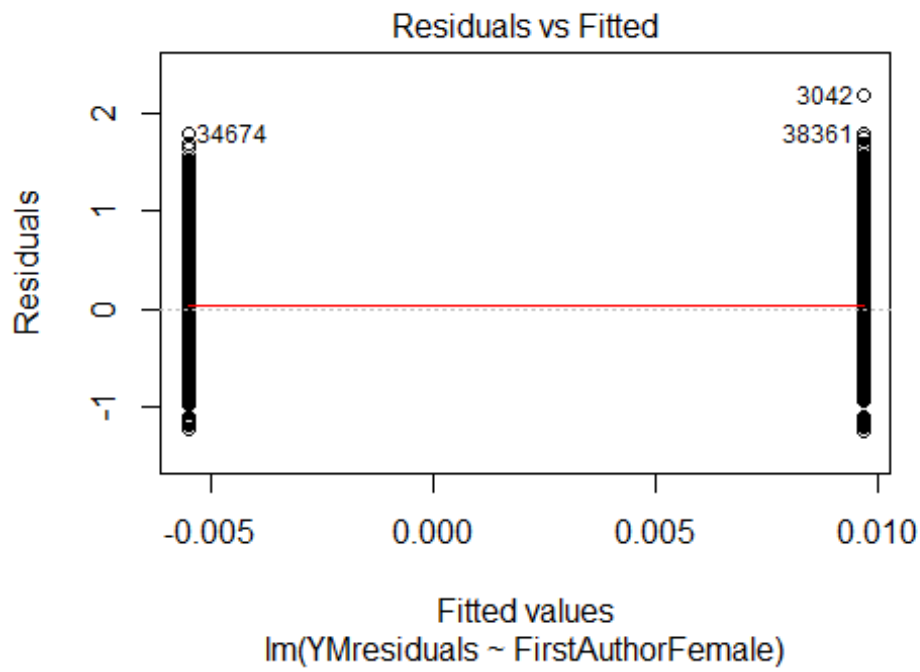

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.2694 -0.3697  0.0293  0.3785  2.3043
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.25610    0.01425   88.16  < 2e-16 ***
## LastAuthorFemale1   0.01330    0.00589    2.26  0.02406 *
## Year1997           -0.00797    0.02007   -0.40  0.69138
## Year1998           -0.06183    0.02038   -3.03  0.00241 **
## Year1999           -0.03046    0.01887   -1.61  0.10642
## Year2000           -0.01588    0.01844   -0.86  0.38912
## Year2001           -0.04379    0.01877   -2.33  0.01965 *
## Year2002           -0.06287    0.01870   -3.36  0.00077 ***
## Year2003           -0.08140    0.01837   -4.43  9.3e-06 ***
## Year2004           -0.06158    0.01870   -3.29  0.00099 ***
## Year2005           -0.02931    0.01814   -1.62  0.10612
## Year2006           -0.07213    0.01766   -4.08  4.4e-05 ***
```

```
## Year2007          -0.07168    0.01791   -4.00  6.3e-05 ***
## Year2008          -0.12536    0.01762   -7.12  1.1e-12 ***
## Year2009          -0.08645    0.01756   -4.92  8.5e-07 ***
## Year2010          -0.07821    0.01744   -4.48  7.3e-06 ***
## Year2011          -0.10005    0.01744   -5.74  9.8e-09 ***
## Year2012          -0.10702    0.01762   -6.07  1.3e-09 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.55
## Multiple R-squared:  0.00363,    Adjusted R-squared:  0.00324
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  3629 weights are ~= 1. The remaining 40389 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0397  0.8610  0.9500  0.9000  0.9860  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          2.27e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Sample size for the above analysis:  44018"
## [1] ""
## [1] ""
## [1] "####################################"
## [1] "Analysis of AJSC 2736"
## [1] "####################################"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 1657 1656 1427 1756 1726 1985 1893 1563 1625 1805 1920 2209 2286 2613 2477
## 2011 2012
## 2986 2963
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 1140 1077  943 1105 1040  908 1356 1186 1225 1356 1360 1585 1602 1847 1688
## 2011 2012
```

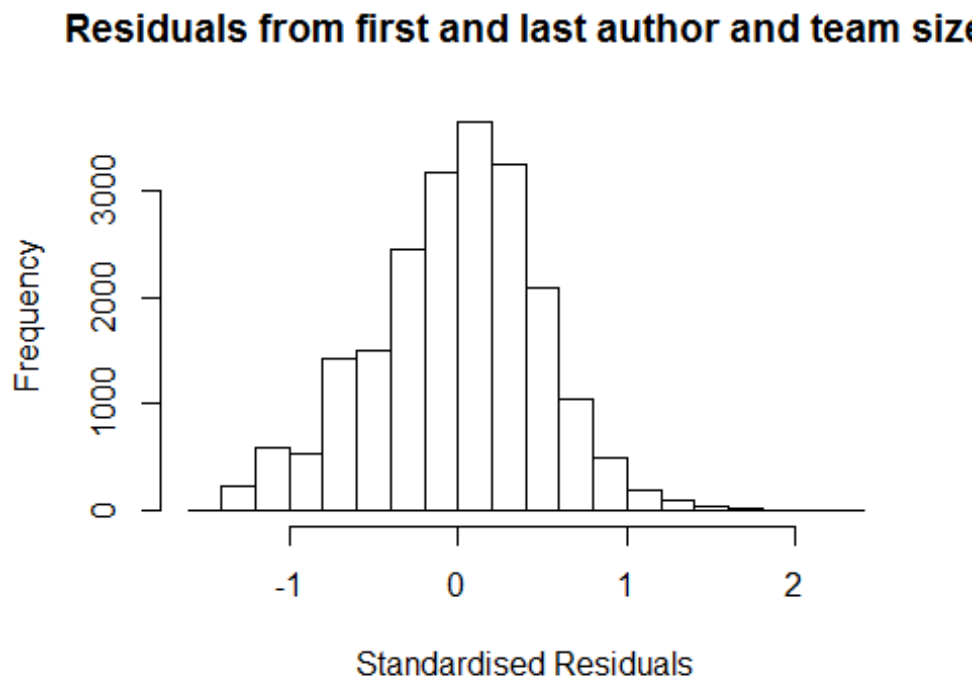
```
## 2171 2088
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 1013  963  846  987  933  804 1186 1030 1089 1176 1197 1380 1411 1606 1470
## 2011 2012
## 1879 1809
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
##  Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 430, df = 16, p-value <2e-16
```




```
##
##  Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 6.5, df = 1, p-value = 0.01
```

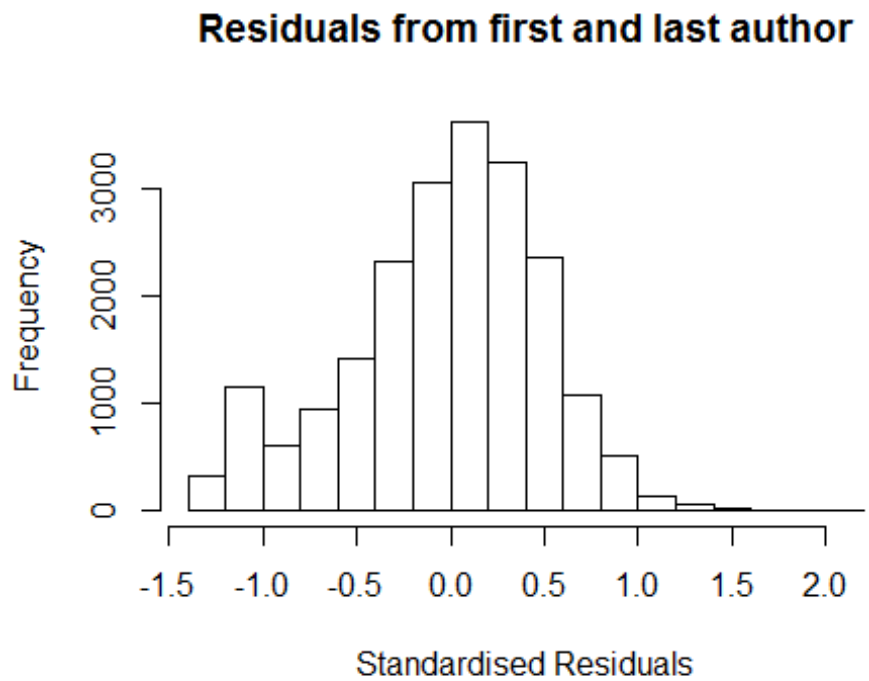



```
## [1] "Female first author team size 2018 geometric mean: 4.70708190403082"
## [1] "Male first author team size 2018 geometric mean: 4.41571431361922"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 510000, p-value = 0.06
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 4.47812035808054"
## [1] "Male last author team size 2018 geometric mean: 4.59450135520087"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 440000, p-value = 0.4
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.027  1           1.014
## LastAuthorFemale  1.021  1           1.010
## UniqueAuthors     1.040  4           1.005
## Year              1.056 16           1.002
```



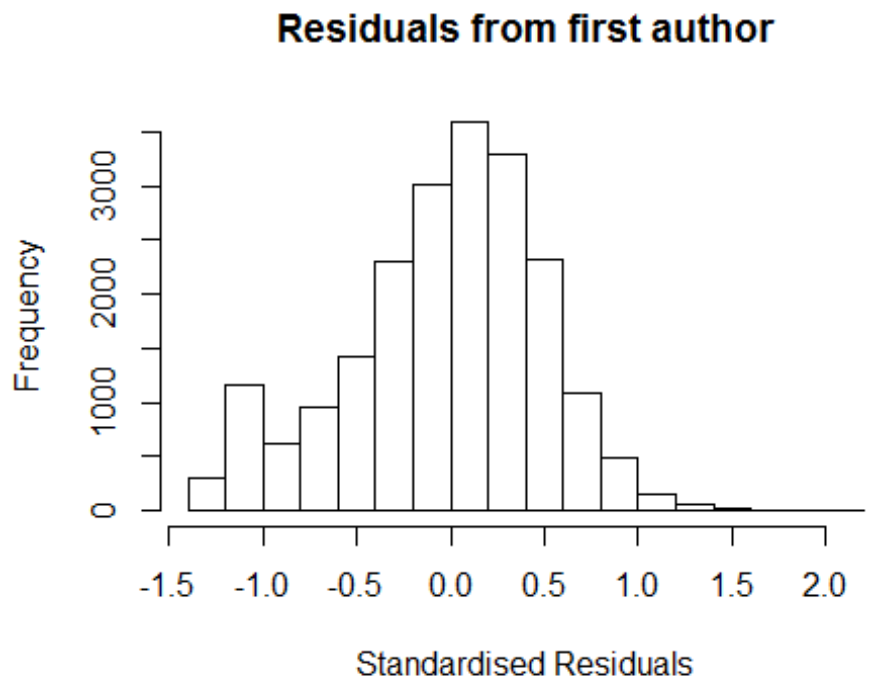

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
UniqueAuthors +
##     Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##     k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.4330 -0.3181  0.0265  0.3196  2.2205
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         0.76351    0.02667   28.62  < 2e-16 ***
## FirstAuthorFemale1  0.00699    0.00720    0.97  0.33130
## LastAuthorFemale1  -0.03262    0.00805   -4.05  5.1e-05 ***
## UniqueAuthors2      0.33457    0.01809   18.49  < 2e-16 ***
## UniqueAuthors3      0.42689    0.01734   24.62  < 2e-16 ***
## UniqueAuthors4      0.48921    0.01722   28.40  < 2e-16 ***
## UniqueAuthors5      0.57800    0.01542   37.48  < 2e-16 ***
## Year1997            0.08446    0.03067    2.75  0.00590 **
## Year1998            0.04111    0.02923    1.41  0.15966
## Year1999            0.04200    0.02847    1.48  0.14012
```

```
## Year2000           -0.05216    0.02775   -1.88  0.06018 .
## Year2001           -0.00881    0.02869   -0.31  0.75875
## Year2002           -0.09084    0.02671   -3.40  0.00067 ***
## Year2003           -0.05161    0.02607   -1.98  0.04776 *
## Year2004           -0.06649    0.02613   -2.54  0.01095 *
## Year2005           -0.06418    0.02571   -2.50  0.01256 *
## Year2006           -0.06275    0.02591   -2.42  0.01544 *
## Year2007           -0.04730    0.02560   -1.85  0.06472 .
## Year2008           -0.02840    0.02619   -1.08  0.27818
## Year2009           -0.03405    0.02535   -1.34  0.17920
## Year2010           -0.07050    0.02557   -2.76  0.00583 **
## Year2011           -0.06573    0.02550   -2.58  0.00995 **
## Year2012           -0.03897    0.02540   -1.53  0.12502
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.462
## Multiple R-squared:  0.128,  Adjusted R-squared:  0.127
## Convergence in 16 IRWLS iterations
##
## Robustness weights:
##  observation 1597 is an outlier with |weight| = 0 ( < 4.8e-06);
##  1738 weights are ~= 1. The remaining 19040 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0086  0.8560  0.9490  0.8960  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          4.81e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.016  1           1.008
## LastAuthorFemale  1.010  1           1.005
## Year              1.017 16           1.001
```



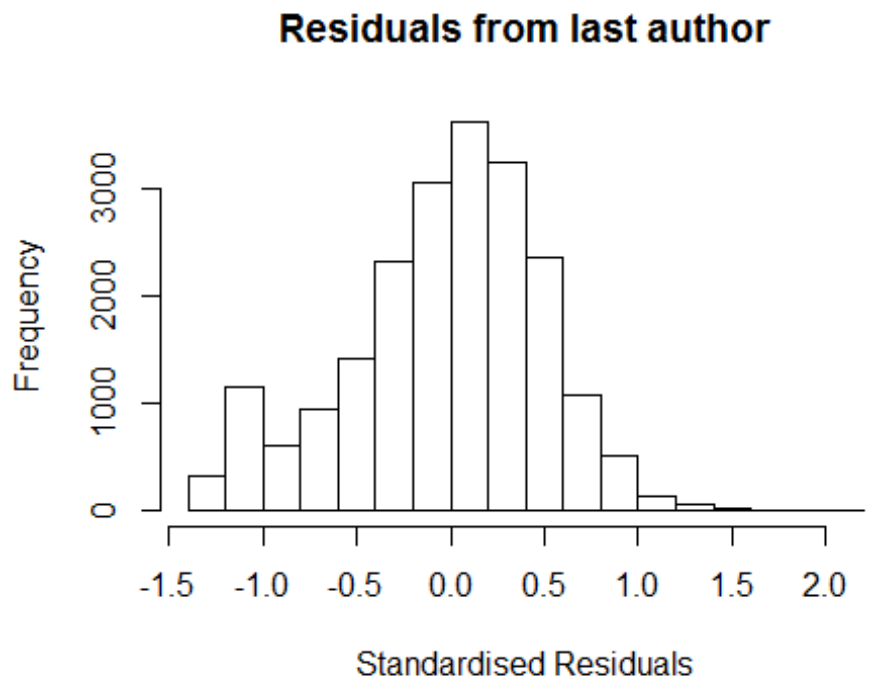

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.2945 -0.3250  0.0332  0.3308  2.1155
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)          1.19131    0.02446   48.70  < 2e-16 ***
## FirstAuthorFemale1   0.02158    0.00753    2.87   0.0042 **
## LastAuthorFemale1   -0.05553    0.00855   -6.50  8.3e-11 ***
## Year1997             0.08160    0.03281    2.49   0.0129 *
## Year1998             0.03485    0.03164    1.10   0.2708
## Year1999             0.03902    0.03040    1.28   0.1993
## Year2000            -0.05212    0.02978   -1.75   0.0801 .
## Year2001            -0.02812    0.03057   -0.92   0.3577
## Year2002            -0.06755    0.02869   -2.35   0.0186 *
## Year2003            -0.01579    0.02805   -0.56   0.5734
## Year2004            -0.02740    0.02792   -0.98   0.3264
## Year2005            -0.02304    0.02775   -0.83   0.4066
```

```
## Year2006           -0.03246    0.02804   -1.16   0.2472
## Year2007           -0.00980    0.02748   -0.36   0.7215
## Year2008            0.00390    0.02814    0.14   0.8897
## Year2009            0.00907    0.02735    0.33   0.7402
## Year2010           -0.02463    0.02768   -0.89   0.3734
## Year2011           -0.04804    0.02760   -1.74   0.0817 .
## Year2012            0.00368    0.02737    0.13   0.8932
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.474
## Multiple R-squared:  0.00675,    Adjusted R-squared:  0.00589
## Convergence in 16 IRWLS iterations
##
## Robustness weights:
##  1708 weights are ~= 1. The remaining 19071 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0082  0.8590  0.9480  0.8900  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          4.81e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.012  1           1.006
## Year              1.012 16           1.000
```




```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.2765 -0.3273  0.0334  0.3298  2.1335
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.181704   0.024416   48.40   <2e-16 ***
## FirstAuthorFemale1  0.013860   0.007569    1.83    0.067 .
## Year1997            0.080966   0.032779    2.47    0.014 *
## Year1998            0.033898   0.031617    1.07    0.284
## Year1999            0.036825   0.030389    1.21    0.226
## Year2000           -0.054287   0.029775   -1.82    0.068 .
## Year2001           -0.029217   0.030604   -0.95    0.340
## Year2002           -0.069113   0.028660   -2.41    0.016 *
## Year2003           -0.018033   0.028018   -0.64    0.520
## Year2004           -0.027297   0.027892   -0.98    0.328
## Year2005           -0.023778   0.027734   -0.86    0.391
## Year2006           -0.034997   0.028016   -1.25    0.212
```

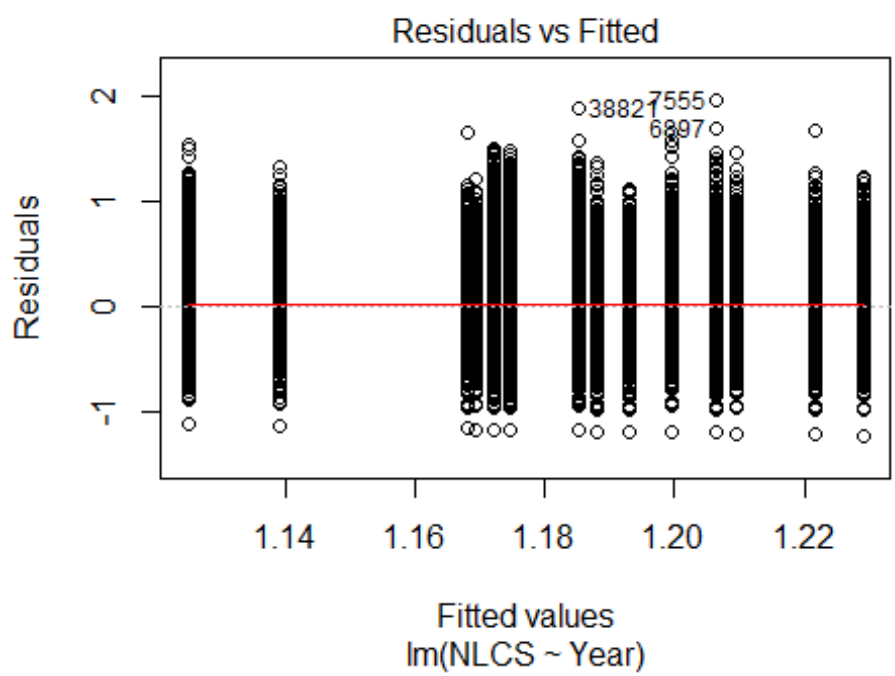
```
## Year2007           -0.012659   0.027464   -0.46    0.645
## Year2008            0.001644   0.028133    0.06    0.953
## Year2009            0.006127   0.027332    0.22    0.823
## Year2010           -0.027623   0.027645   -1.00    0.318
## Year2011           -0.053576   0.027556   -1.94    0.052 .
## Year2012            0.000286   0.027346    0.01    0.992
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.474
## Multiple R-squared:  0.00454,    Adjusted R-squared:  0.00373
## Convergence in 16 IRWLS iterations
##
## Robustness weights:
##  1734 weights are ~= 1. The remaining 19045 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0057  0.8580  0.9470  0.8890  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          4.81e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##                   GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.006  1           1.003
## Year             1.006 16           1.000
```



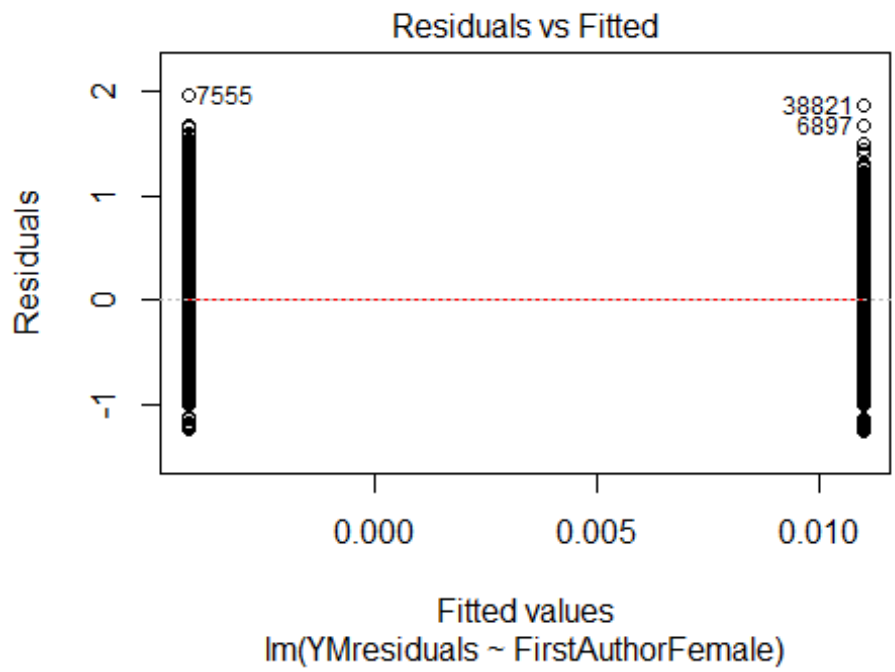

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.2791 -0.3235  0.0326  0.3303  2.1309
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.19636    0.02438   49.08  < 2e-16 ***
## LastAuthorFemale1 -0.05193    0.00854   -6.08  1.2e-09 ***
## Year1997            0.08278    0.03282    2.52    0.012 *
## Year1998            0.03577    0.03164    1.13    0.258
## Year1999            0.03979    0.03038    1.31    0.190
## Year2000           -0.05025    0.02976   -1.69    0.091 .
## Year2001           -0.02704    0.03054   -0.89    0.376
## Year2002           -0.06571    0.02866   -2.29    0.022 *
## Year2003           -0.01430    0.02802   -0.51    0.610
## Year2004           -0.02639    0.02789   -0.95    0.344
## Year2005           -0.02147    0.02772   -0.77    0.439
## Year2006           -0.03093    0.02801   -1.10    0.270
```

```
## Year2007          -0.00781    0.02744   -0.28    0.776
## Year2008           0.00600    0.02810    0.21    0.831
## Year2009           0.01207    0.02730    0.44    0.658
## Year2010          -0.02155    0.02762   -0.78    0.435
## Year2011          -0.04531    0.02755   -1.64    0.100
## Year2012           0.00715    0.02731    0.26    0.793
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.474
## Multiple R-squared:  0.0064, Adjusted R-squared:  0.00558
## Convergence in 16 IRWLS iterations
##
## Robustness weights:
##  1732 weights are ~= 1. The remaining 19047 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.006   0.858   0.948   0.890   0.985   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          4.81e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Sample size for the above analysis:  20779"
## [1] ""
## [1] ""
## [1] "####################################"
## [1] "Analysis of AJSC 2737"
## [1] "####################################"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 2419 2852 2425 2307 2437 2573 1865 1687 1692 1725 1749 1905 2110 2508 2303
## 2011 2012
## 2516 2564
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 1610 1938 1560 1507 1552 1443 1310 1168 1135 1170 1190 1286 1456 1757 1596
## 2011 2012
```

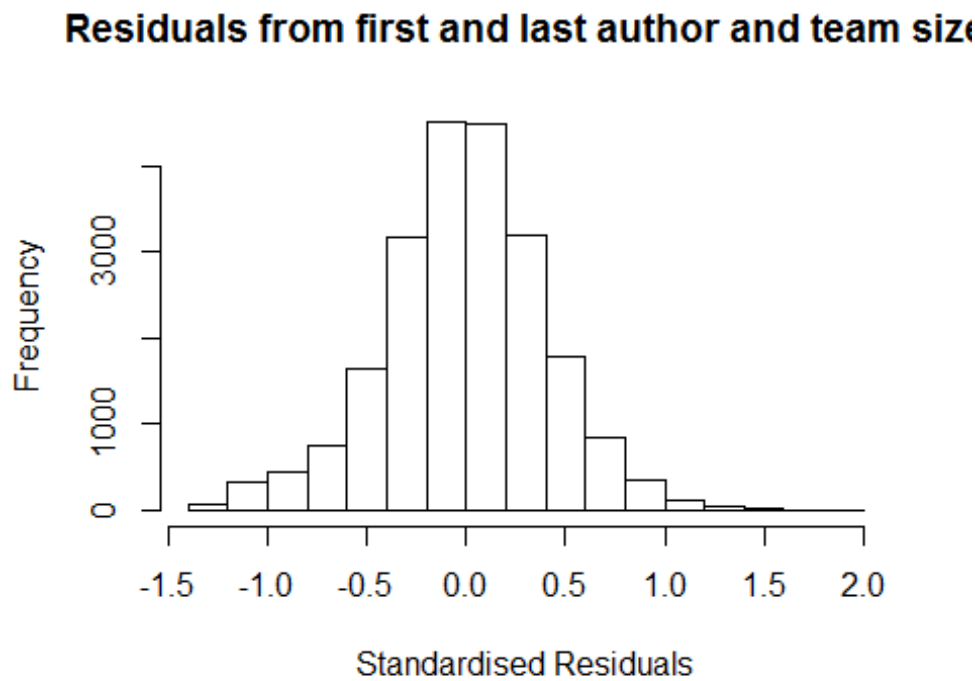
```
## 1836 1835
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 1403 1686 1376 1328 1331 1258 1116  993  978 1006 1029 1103 1250 1499 1359
## 2011 2012
## 1544 1566
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
##  Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 330, df = 16, p-value <2e-16
```




```
##
##  Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 170, df = 1, p-value <2e-16
```

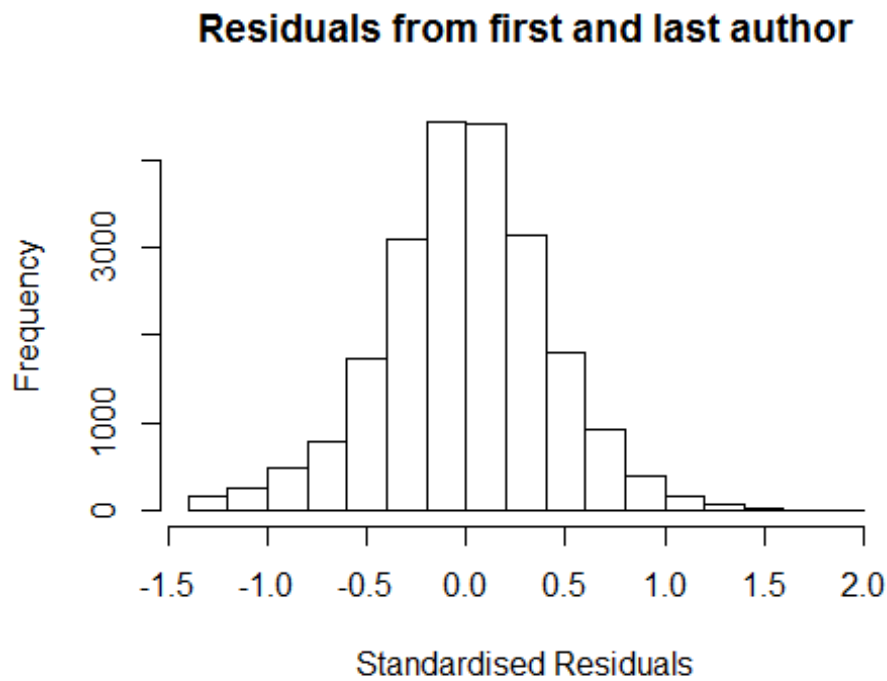



```
## [1] "Female first author team size 2018 geometric mean: 4.89445047347115"
## [1] "Male first author team size 2018 geometric mean: 5.05438777762373"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 220000, p-value = 0.1
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 4.85919975029446"
## [1] "Male last author team size 2018 geometric mean: 5.04284332613791"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 180000, p-value = 0.3
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.027  1           1.013
## LastAuthorFemale  1.023  1           1.011
## UniqueAuthors     1.062  4           1.008
## Year              1.071 16           1.002
```



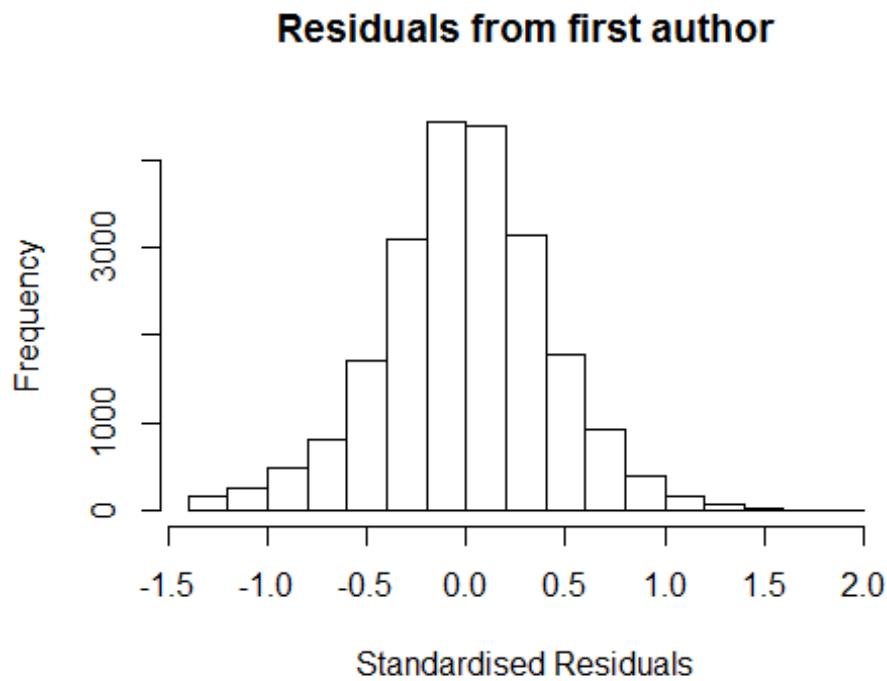

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
UniqueAuthors +
##     Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##     k.max = 1000))
##  \--> method = "MM"
## Residuals:
##       Min        1Q    Median        3Q       Max
## -1.342554 -0.250342 -0.000401  0.251353  1.825047
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.035917   0.018921   54.75  < 2e-16 ***
## FirstAuthorFemale1  0.004780   0.005781    0.83  0.40836
## LastAuthorFemale1   0.006987   0.007047    0.99  0.32148
## UniqueAuthors2      0.087721   0.016957    5.17  2.3e-07 ***
## UniqueAuthors3      0.110345   0.016542    6.67  2.6e-11 ***
## UniqueAuthors4      0.139280   0.016599    8.39  < 2e-16 ***
## UniqueAuthors5      0.299650   0.015850   18.91  < 2e-16 ***
## Year1997           -0.069920   0.016758   -4.17  3.0e-05 ***
## Year1998           -0.000614   0.016556   -0.04  0.97040
## Year1999           -0.024946   0.015939   -1.57  0.11757
```

```
## Year2000           -0.046679   0.015887   -2.94  0.00331 **
## Year2001           -0.045226   0.015940   -2.84  0.00456 **
## Year2002           -0.066483   0.016338   -4.07  4.7e-05 ***
## Year2003           -0.084914   0.016499   -5.15  2.7e-07 ***
## Year2004           -0.062836   0.016889   -3.72  0.00020 ***
## Year2005           -0.071996   0.016541   -4.35  1.4e-05 ***
## Year2006           -0.004689   0.017488   -0.27  0.78860
## Year2007           -0.024757   0.016905   -1.46  0.14307
## Year2008           -0.028516   0.016430   -1.74  0.08265 .
## Year2009           -0.038691   0.016430   -2.35  0.01854 *
## Year2010           -0.031000   0.016896   -1.83  0.06655 .
## Year2011           -0.062758   0.016564   -3.79  0.00015 ***
## Year2012           -0.062910   0.016599   -3.79  0.00015 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.373
## Multiple R-squared:  0.068,  Adjusted R-squared:  0.0671
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
##  2 observations c(3855,19506) are outliers with |weight| = 0 ( < 4.6e-06);
##  1891 weights are ~= 1. The remaining 19932 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0005  0.8610  0.9500  0.8910  0.9860  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          4.58e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.026  1           1.013
## LastAuthorFemale  1.018  1           1.009
## Year              1.017 16           1.001
```



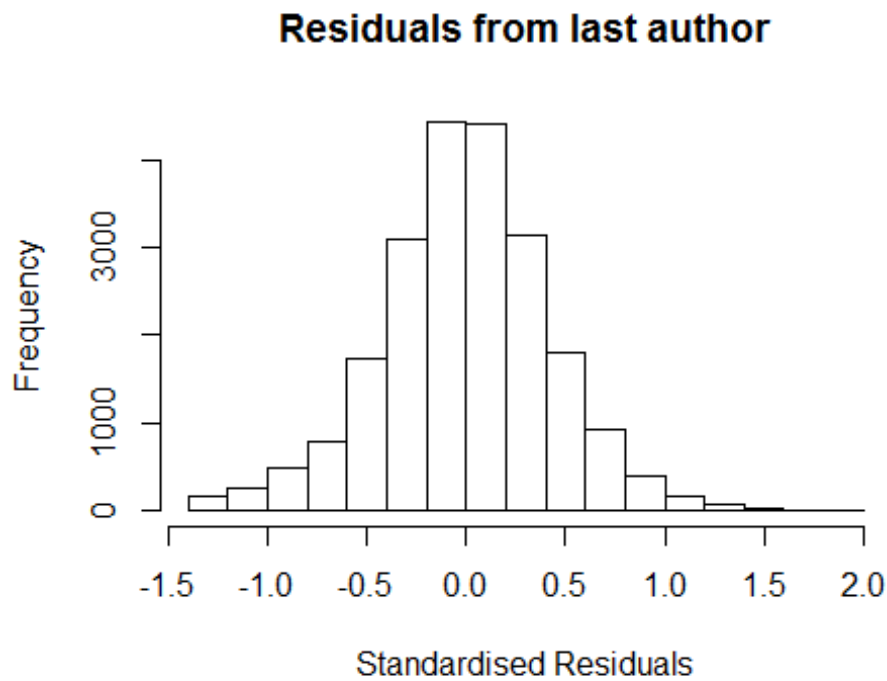

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.24669 -0.25493 -0.00123  0.25131  1.94866
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)          1.19621    0.01268   94.33  < 2e-16 ***
## FirstAuthorFemale1  0.00744    0.00593    1.25  0.20957
## LastAuthorFemale1  -0.00737    0.00725   -1.02  0.30904
## Year1997           -0.05795    0.01733   -3.34  0.00083 ***
## Year1998            0.01513    0.01706    0.89  0.37519
## Year1999            0.00102    0.01630    0.06  0.95005
## Year2000           -0.01723    0.01637   -1.05  0.29266
## Year2001           -0.00528    0.01635   -0.32  0.74661
## Year2002           -0.02328    0.01666   -1.40  0.16246
## Year2003           -0.04444    0.01686   -2.64  0.00841 **
## Year2004           -0.01973    0.01741   -1.13  0.25724
## Year2005           -0.02007    0.01694   -1.18  0.23631
```

```
## Year2006            0.04305    0.01792    2.40  0.01629 *
## Year2007            0.02257    0.01745    1.29  0.19600
## Year2008            0.02945    0.01688    1.74  0.08101 .
## Year2009            0.01005    0.01695    0.59  0.55339
## Year2010            0.02504    0.01740    1.44  0.15030
## Year2011           -0.01346    0.01689   -0.80  0.42540
## Year2012           -0.01634    0.01694   -0.96  0.33469
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.383
## Multiple R-squared:  0.00424,    Adjusted R-squared:  0.00342
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  2 observations c(3855,19506) are outliers with |weight| = 0 ( < 4.6e-06);
##  1895 weights are ~= 1. The remaining 19928 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0156  0.8610  0.9510  0.8900  0.9860  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          4.58e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.012  1           1.006
## Year              1.012 16           1.000
```




```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.24467 -0.25489 -0.00145  0.25171  1.94943
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)          1.19525    0.01264   94.55  < 2e-16 ***
## FirstAuthorFemale1   0.00669    0.00589    1.14  0.25603
## Year1997            -0.05796    0.01733   -3.34  0.00083 ***
## Year1998             0.01532    0.01706    0.90  0.36939
## Year1999             0.00101    0.01630    0.06  0.95074
## Year2000            -0.01735    0.01637   -1.06  0.28931
## Year2001            -0.00536    0.01635   -0.33  0.74297
## Year2002            -0.02316    0.01667   -1.39  0.16458
## Year2003            -0.04460    0.01687   -2.64  0.00820 **
## Year2004            -0.01983    0.01742   -1.14  0.25507
## Year2005            -0.02005    0.01695   -1.18  0.23682
## Year2006             0.04273    0.01791    2.39  0.01708 *
```

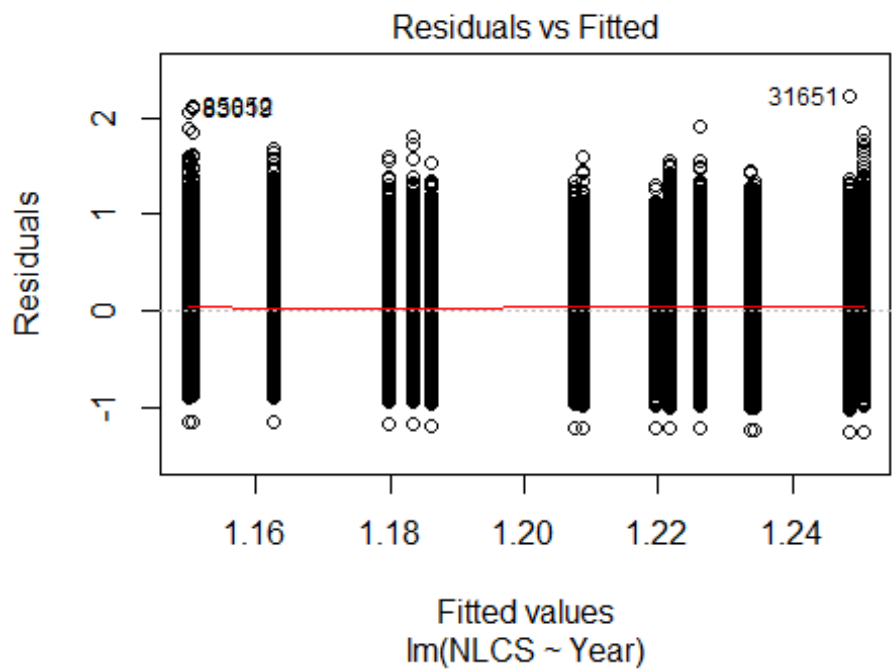
```
## Year2007            0.02245    0.01745    1.29  0.19829
## Year2008            0.02923    0.01688    1.73  0.08327 .
## Year2009            0.00987    0.01696    0.58  0.56052
## Year2010            0.02481    0.01740    1.43  0.15412
## Year2011           -0.01381    0.01689   -0.82  0.41373
## Year2012           -0.01647    0.01694   -0.97  0.33098
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.383
## Multiple R-squared:  0.00419,    Adjusted R-squared:  0.00342
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  2 observations c(3855,19506) are outliers with |weight| = 0 ( < 4.6e-06);
##  1887 weights are ~= 1. The remaining 19936 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0152  0.8610  0.9510  0.8900  0.9860  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          4.58e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##                   GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.005  1           1.003
## Year             1.005 16           1.000
```



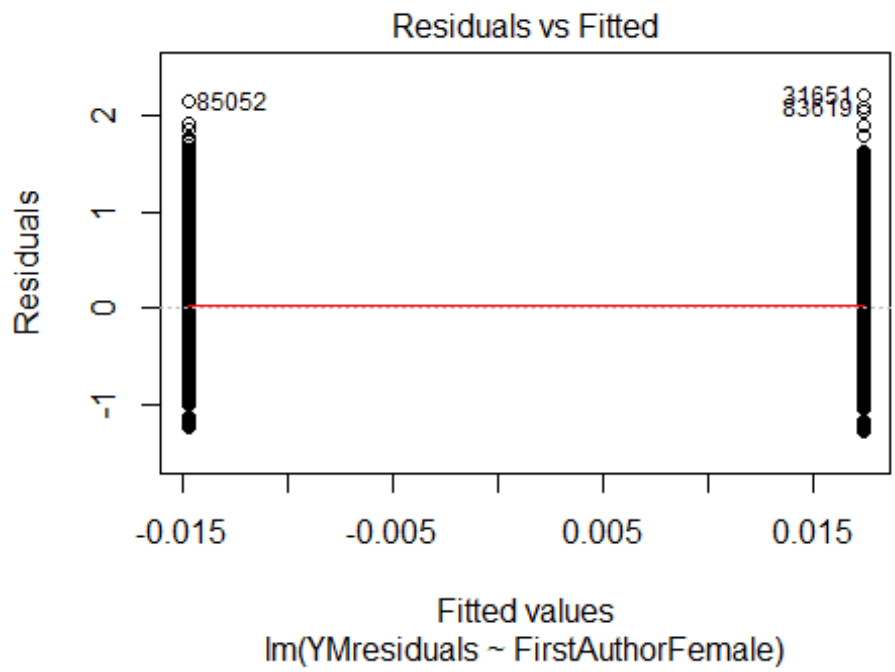

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##       Min        1Q    Median        3Q       Max
## -1.241623 -0.255531 -0.000623  0.251677  1.946879
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.19772    0.01261   95.02  < 2e-16 ***
## LastAuthorFemale1 -0.00628    0.00720   -0.87  0.38311
## Year1997          -0.05797    0.01732   -3.35  0.00082 ***
## Year1998           0.01540    0.01705    0.90  0.36662
## Year1999           0.00133    0.01629    0.08  0.93491
## Year2000          -0.01667    0.01636   -1.02  0.30813
## Year2001          -0.00479    0.01633   -0.29  0.76927
## Year2002          -0.02287    0.01666   -1.37  0.16968
## Year2003          -0.04409    0.01686   -2.61  0.00893 **
## Year2004          -0.01919    0.01740   -1.10  0.27017
## Year2005          -0.01940    0.01693   -1.15  0.25179
## Year2006           0.04390    0.01789    2.45  0.01414 *
```

```
## Year2007           0.02306    0.01744    1.32  0.18620
## Year2008           0.03009    0.01686    1.78  0.07435 .
## Year2009           0.01083    0.01693    0.64  0.52213
## Year2010           0.02581    0.01739    1.48  0.13788
## Year2011          -0.01268    0.01687   -0.75  0.45211
## Year2012          -0.01564    0.01692   -0.92  0.35539
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.383
## Multiple R-squared:  0.00417,    Adjusted R-squared:  0.0034
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  2 observations c(3855,19506) are outliers with |weight| = 0 ( < 4.6e-06);
##  1919 weights are ~= 1. The remaining 19904 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0142  0.8610  0.9510  0.8900  0.9860  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          4.58e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Sample size for the above analysis:  21825"
## [1] ""
## [1] ""
## [1] "####################################"
## [1] "Analysis of AJSC 2738"
## [1] "####################################"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 4018 3755 3803 3550 4223 4489 4218 3385 3770 4044 4946 5200 5641 6084 6206
## 2011 2012
## 6857 6874
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 3273 2948 2900 2752 2921 2678 3497 2906 3163 3478 4256 4414 4782 5134 5225
```

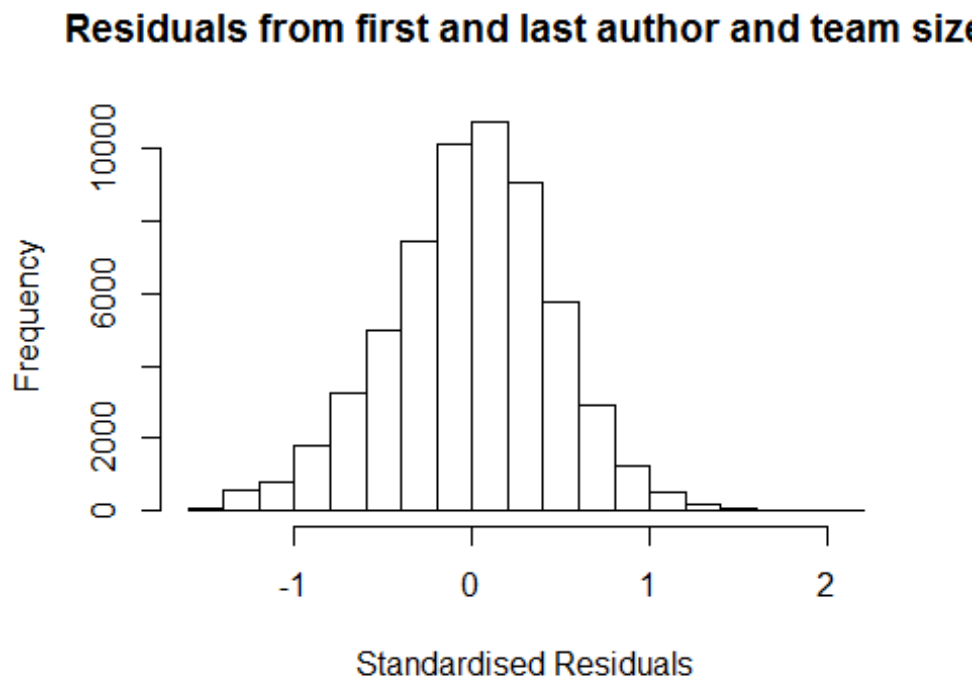
```
## 2011 2012
## 5810 5807
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 3014 2689 2622 2484 2652 2419 3124 2588 2824 3084 3864 3956 4298 4660 4697
## 2011 2012
## 5247 5235
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
##  Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 470, df = 16, p-value <2e-16
```




```
##
##  Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 180, df = 1, p-value <2e-16
```

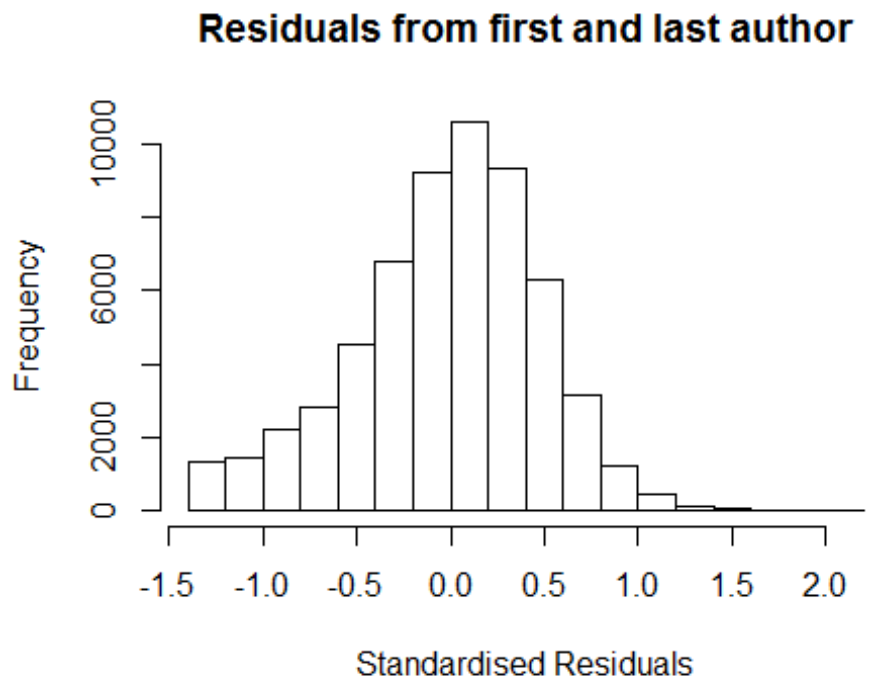



```
## [1] "Female first author team size 2018 geometric mean: 4.16262586304201"
## [1] "Male first author team size 2018 geometric mean: 3.78821442953702"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 4400000, p-value = 4e-05
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 3.9295463401178"
## [1] "Male last author team size 2018 geometric mean: 4.06751112435417"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 4e+06, p-value = 0.001
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.034  1           1.017
## LastAuthorFemale  1.032  1           1.016
## UniqueAuthors     1.045  4           1.006
## Year              1.047 16           1.001
```



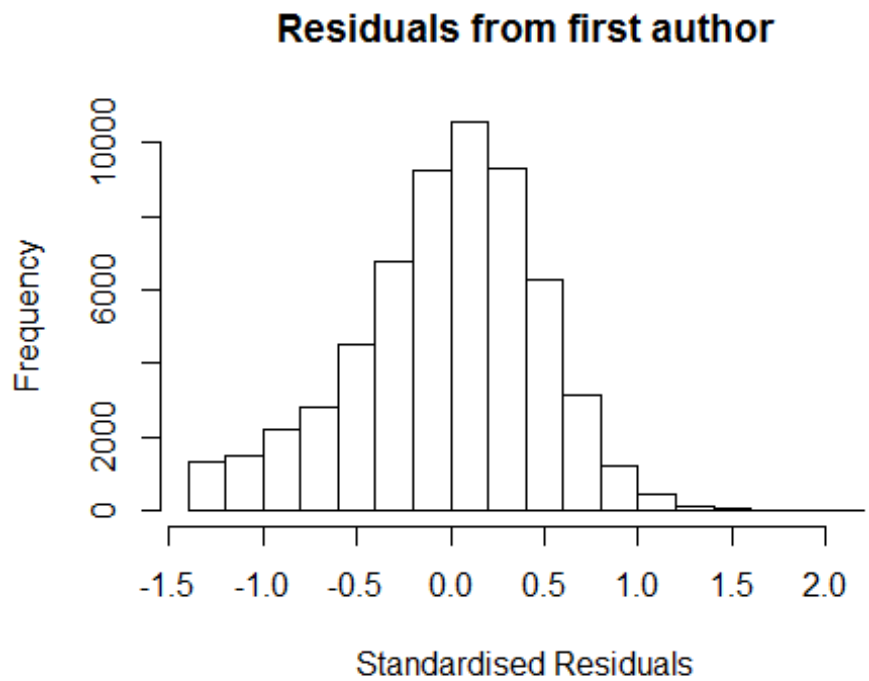

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
UniqueAuthors +
##     Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##     k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.5000 -0.2985  0.0131  0.2989  2.1248
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         0.89325    0.01198   74.56  < 2e-16 ***
## FirstAuthorFemale1  0.01876    0.00388    4.84  1.3e-06 ***
## LastAuthorFemale1  -0.00228    0.00399   -0.57   0.5669
## UniqueAuthors2      0.35775    0.00841   42.54  < 2e-16 ***
## UniqueAuthors3      0.43569    0.00808   53.91  < 2e-16 ***
## UniqueAuthors4      0.49490    0.00810   61.07  < 2e-16 ***
## UniqueAuthors5      0.58800    0.00733   80.23  < 2e-16 ***
## Year1997           -0.02618    0.01450   -1.81   0.0710 .
## Year1998           -0.04253    0.01455   -2.92   0.0035 **
## Year1999           -0.05508    0.01400   -3.93  8.3e-05 ***
```

```
## Year2000           -0.05844    0.01363   -4.29  1.8e-05 ***
## Year2001           -0.05820    0.01388   -4.19  2.7e-05 ***
## Year2002           -0.05264    0.01318   -3.99  6.5e-05 ***
## Year2003           -0.09197    0.01336   -6.89  5.8e-12 ***
## Year2004           -0.10955    0.01293   -8.47  < 2e-16 ***
## Year2005           -0.12217    0.01263   -9.68  < 2e-16 ***
## Year2006           -0.14408    0.01238  -11.64  < 2e-16 ***
## Year2007           -0.14071    0.01229  -11.45  < 2e-16 ***
## Year2008           -0.17386    0.01217  -14.28  < 2e-16 ***
## Year2009           -0.14760    0.01215  -12.14  < 2e-16 ***
## Year2010           -0.17750    0.01224  -14.51  < 2e-16 ***
## Year2011           -0.17851    0.01219  -14.64  < 2e-16 ***
## Year2012           -0.17995    0.01220  -14.75  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.441
## Multiple R-squared:  0.158,  Adjusted R-squared:  0.157
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
##  2 observations c(38368,53914) are outliers with |weight| = 0 ( < 1.7e-
06);
##  5119 weights are ~= 1. The remaining 54336 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0016  0.8640  0.9500  0.8990  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.68e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.016  1           1.008
## LastAuthorFemale  1.012  1           1.006
## Year              1.018 16           1.001
```



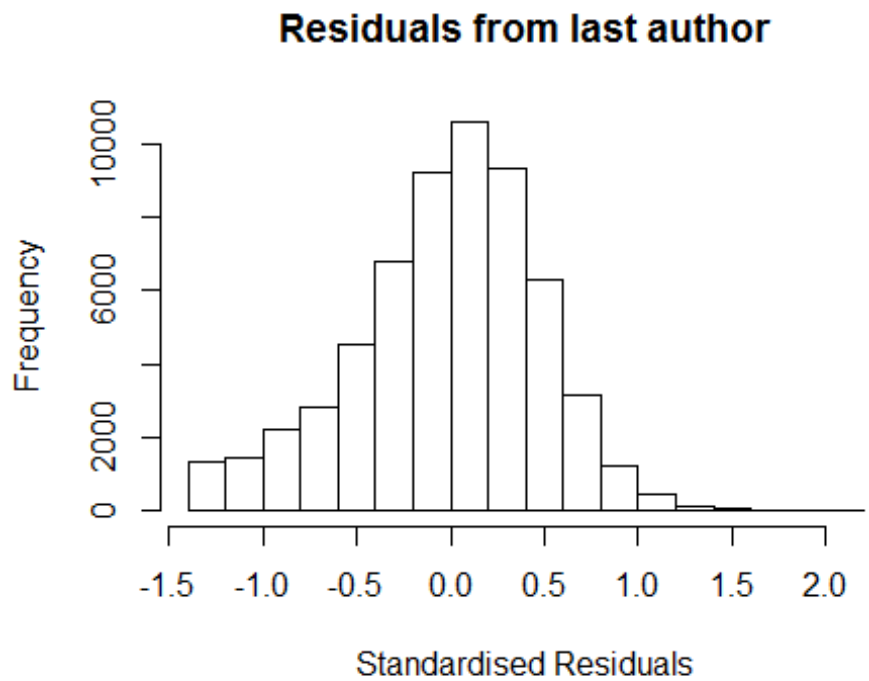

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.3057 -0.3174  0.0271  0.3139  2.1653
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)          1.26436    0.01100  114.97  < 2e-16 ***
## FirstAuthorFemale1  0.03658    0.00412    8.88  < 2e-16 ***
## LastAuthorFemale1  -0.01928    0.00422   -4.57  4.8e-06 ***
## Year1997           -0.02559    0.01556   -1.64  0.10016
## Year1998           -0.02515    0.01561   -1.61  0.10727
## Year1999           -0.04921    0.01496   -3.29  0.00100 **
## Year2000           -0.02662    0.01446   -1.84  0.06575 .
## Year2001           -0.02050    0.01492   -1.37  0.16929
## Year2002            0.00478    0.01387    0.34  0.73045
## Year2003           -0.03350    0.01408   -2.38  0.01738 *
## Year2004           -0.03485    0.01370   -2.54  0.01095 *
## Year2005           -0.04561    0.01337   -3.41  0.00065 ***
```

```
## Year2006           -0.07308    0.01313   -5.57  2.6e-08 ***
## Year2007           -0.07060    0.01301   -5.43  5.7e-08 ***
## Year2008           -0.09705    0.01301   -7.46  8.7e-14 ***
## Year2009           -0.06803    0.01296   -5.25  1.5e-07 ***
## Year2010           -0.11091    0.01311   -8.46  < 2e-16 ***
## Year2011           -0.10107    0.01298   -7.79  7.0e-15 ***
## Year2012           -0.09421    0.01298   -7.26  3.9e-13 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.461
## Multiple R-squared:  0.00653,    Adjusted R-squared:  0.00623
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
##  observation 18384 is an outlier with |weight| = 0 ( < 1.7e-06);
##  4985 weights are ~= 1. The remaining 54471 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0007  0.8610  0.9490  0.8930  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.68e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.012  1           1.006
## Year              1.012 16           1.000
```




```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##    Min     1Q Median     3Q    Max
## -1.297 -0.317  0.028  0.315  2.174
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)          1.25901    0.01092  115.33  < 2e-16 ***
## FirstAuthorFemale1  0.03311    0.00414    8.00  1.2e-15 ***
## Year1997           -0.02545    0.01556   -1.64  0.10194
## Year1998           -0.02487    0.01561   -1.59  0.11125
## Year1999           -0.04911    0.01497   -3.28  0.00103 **
## Year2000           -0.02704    0.01446   -1.87  0.06156 .
## Year2001           -0.02052    0.01492   -1.38  0.16889
## Year2002            0.00472    0.01388    0.34  0.73385
## Year2003           -0.03383    0.01409   -2.40  0.01638 *
## Year2004           -0.03509    0.01370   -2.56  0.01043 *
## Year2005           -0.04579    0.01337   -3.42  0.00062 ***
## Year2006           -0.07350    0.01313   -5.60  2.2e-08 ***
```

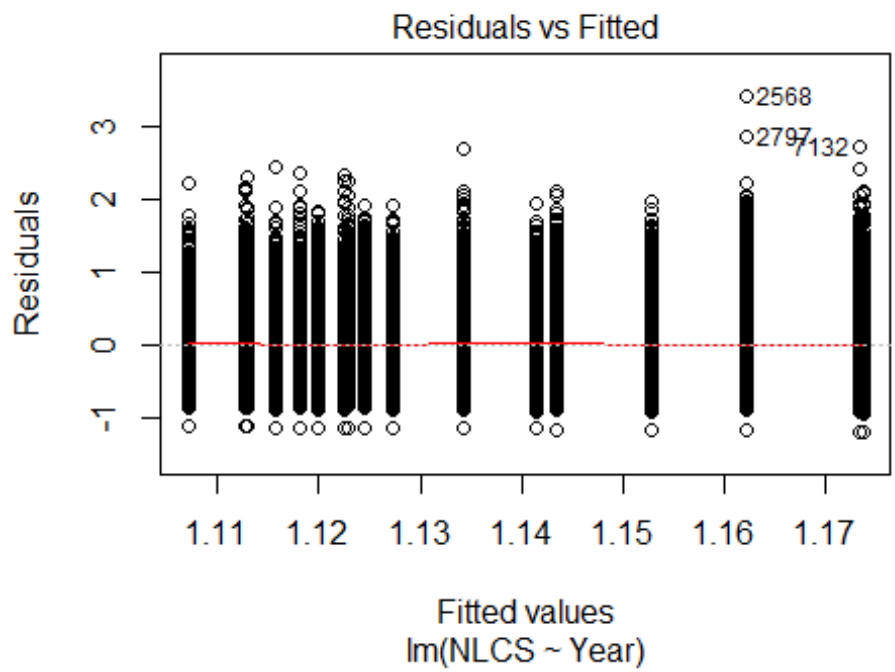
```
## Year2007           -0.07109    0.01301   -5.46  4.7e-08 ***
## Year2008           -0.09752    0.01301   -7.49  6.7e-14 ***
## Year2009           -0.06854    0.01296   -5.29  1.2e-07 ***
## Year2010           -0.11156    0.01311   -8.51  < 2e-16 ***
## Year2011           -0.10194    0.01298   -7.85  4.1e-15 ***
## Year2012           -0.09538    0.01297   -7.35  2.0e-13 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.461
## Multiple R-squared:  0.00623,    Adjusted R-squared:  0.00594
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
##  observation 18384 is an outlier with |weight| = 0 ( < 1.7e-06);
##  4970 weights are ~= 1. The remaining 54486 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0015  0.8610  0.9490  0.8930  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.68e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##                   GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.008  1           1.004
## Year             1.008 16           1.000
```



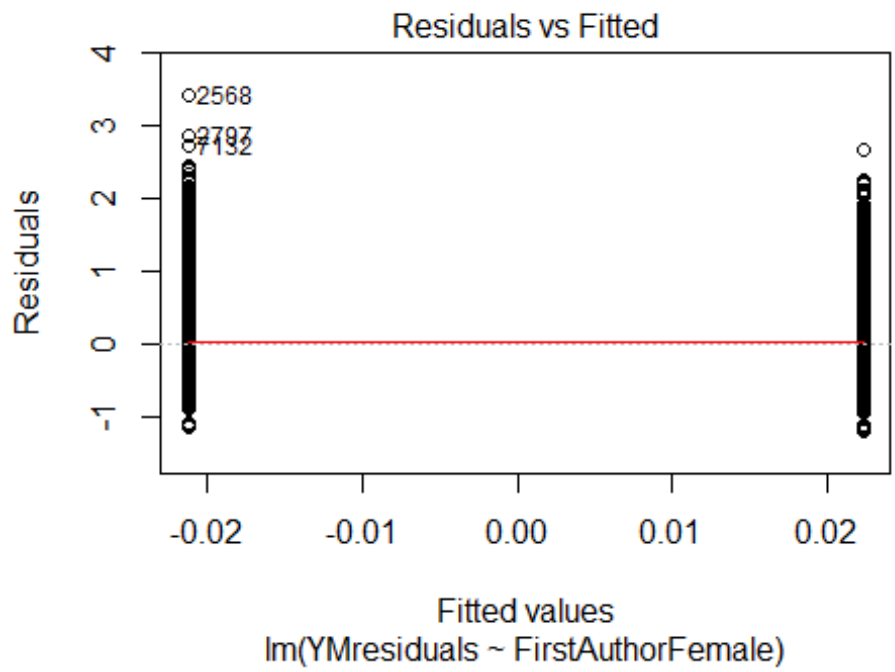

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.2826 -0.3183  0.0275  0.3145  2.1884
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.27622    0.01088  117.32  < 2e-16 ***
## LastAuthorFemale1 -0.01238    0.00423   -2.93   0.0034 **
## Year1997          -0.02558    0.01556   -1.64   0.1001
## Year1998          -0.02532    0.01561   -1.62   0.1046
## Year1999          -0.04910    0.01497   -3.28   0.0010 **
## Year2000          -0.02502    0.01447   -1.73   0.0838 .
## Year2001          -0.01918    0.01492   -1.29   0.1987
## Year2002           0.00637    0.01388    0.46   0.6462
## Year2003          -0.03125    0.01409   -2.22   0.0265 *
## Year2004          -0.03138    0.01370   -2.29   0.0219 *
## Year2005          -0.04274    0.01337   -3.20   0.0014 **
## Year2006          -0.07048    0.01313   -5.37  8.1e-08 ***
```

```
## Year2007          -0.06774    0.01301   -5.21  1.9e-07 ***
## Year2008          -0.09302    0.01300   -7.15  8.5e-13 ***
## Year2009          -0.06379    0.01296   -4.92  8.5e-07 ***
## Year2010          -0.10678    0.01310   -8.15  3.7e-16 ***
## Year2011          -0.09682    0.01297   -7.46  8.6e-14 ***
## Year2012          -0.08912    0.01297   -6.87  6.4e-12 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.461
## Multiple R-squared:  0.00522,    Adjusted R-squared:  0.00494
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
##  observation 18384 is an outlier with |weight| = 0 ( < 1.7e-06);
##  4988 weights are ~= 1. The remaining 54468 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0023  0.8610  0.9490  0.8930  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.68e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Sample size for the above analysis:  59457"
## [1] ""
## [1] ""
## [1] "####################################"
## [1] "Analysis of AJSC 2739"
## [1] "####################################"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 3953 3924 3884 3852 4566 4594 4444 3996 4191 4782 5300 5650 5880 6519 6735
## 2011 2012
## 7299 7300
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 2833 2808 2799 2959 3031 2550 3646 3294 3374 3840 4340 4636 4735 5291 5508
```

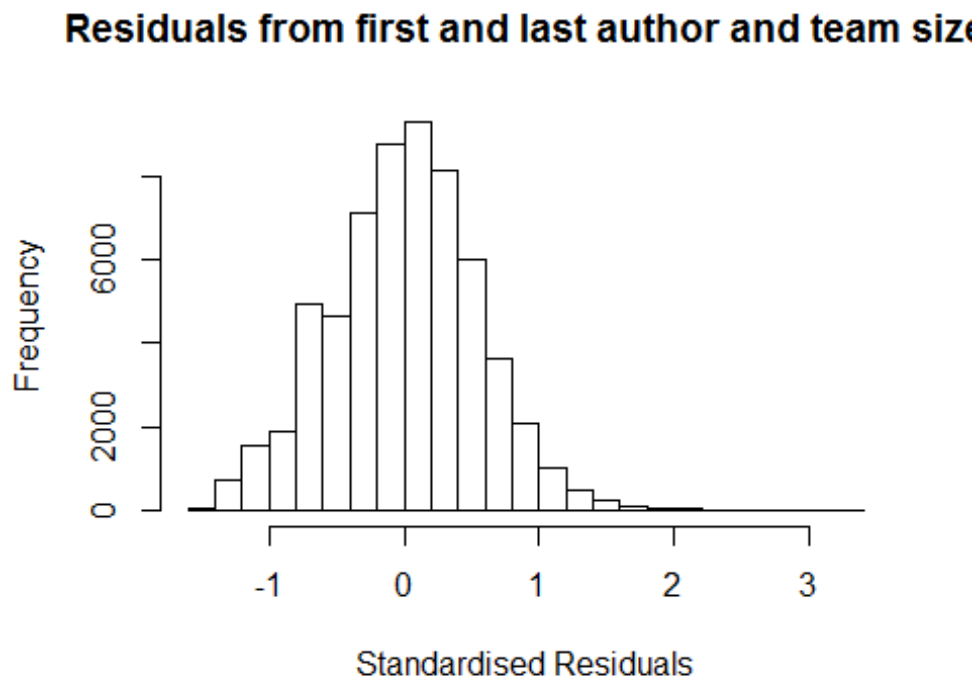
```
## 2011 2012
## 5938 5949
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 2599 2600 2552 2701 2757 2310 3281 2933 3021 3441 3845 4203 4249 4736 4938
## 2011 2012
## 5333 5316
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
##  Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 790, df = 16, p-value <2e-16
```




```
##
##  Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 75, df = 1, p-value <2e-16
```



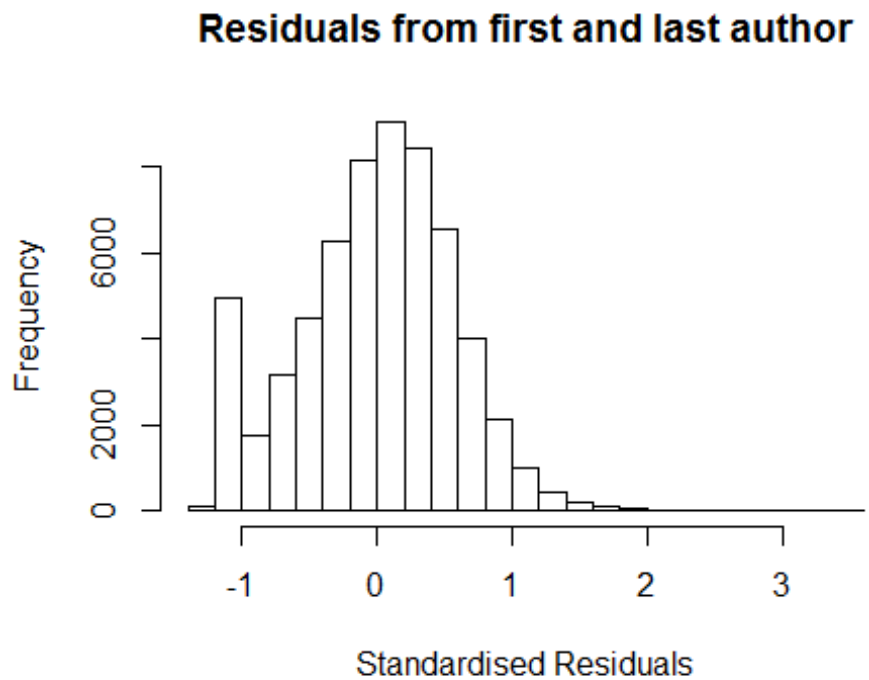

```
## [1] "Female first author team size 2018 geometric mean: 4.20605214371508"
## [1] "Male first author team size 2018 geometric mean: 3.76572246037767"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 4900000, p-value = 5e-10
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 3.98923932519689"
## [1] "Male last author team size 2018 geometric mean: 4.06487472826512"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 4600000, p-value = 0.2
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.069  1           1.034
## LastAuthorFemale  1.064  1           1.032
## UniqueAuthors     1.043  4           1.005
## Year              1.045 16           1.001
```




```
## [1] "List of  4 outliers with residuals above  2.5"
##         ScopusId  NLCS Year OneField Fields residuals
## 2568  0030096228 4.564 1996     2739      1     3.304
## 2797  0029680562 4.006 1996     2719      2     2.728
## 7132  0030881741 3.884 1997     2739      2     2.706
## 10924 0031613172 3.817 1998     2739      1     2.534
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
UniqueAuthors +
##     Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##     k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.4556 -0.3578  0.0147  0.3586  3.3041
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)          0.83389    0.01538   54.23  < 2e-16 ***
## FirstAuthorFemale1   0.02802    0.00463    6.05  1.4e-09 ***
## LastAuthorFemale1   -0.01841    0.00468   -3.93  8.4e-05 ***
## UniqueAuthors2       0.32877    0.00897   36.64  < 2e-16 ***
## UniqueAuthors3       0.44444    0.00857   51.86  < 2e-16 ***
## UniqueAuthors4       0.50377    0.00854   58.97  < 2e-16 ***
## UniqueAuthors5       0.57793    0.00755   76.54  < 2e-16 ***
```

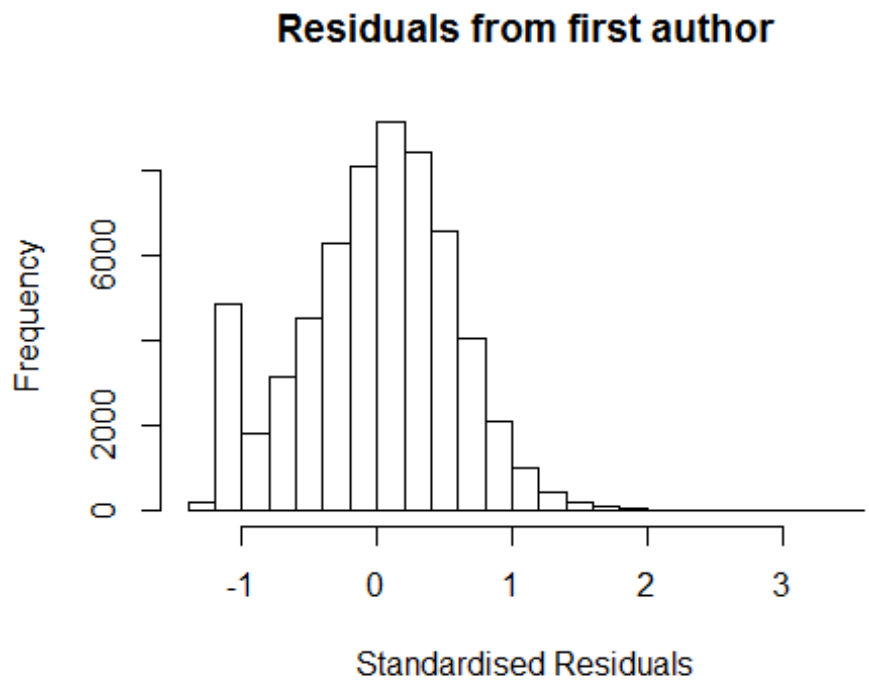
```
## Year1997            0.01579    0.01966    0.80  0.42182
## Year1998           -0.06427    0.01917   -3.35  0.00080 ***
## Year1999           -0.02047    0.01827   -1.12  0.26246
## Year2000           -0.03973    0.01805   -2.20  0.02772 *
## Year2001           -0.06065    0.01839   -3.30  0.00097 ***
## Year2002           -0.10508    0.01736   -6.05  1.4e-09 ***
## Year2003           -0.11695    0.01738   -6.73  1.7e-11 ***
## Year2004           -0.09567    0.01725   -5.55  2.9e-08 ***
## Year2005           -0.13418    0.01690   -7.94  2.0e-15 ***
## Year2006           -0.12900    0.01647   -7.83  4.9e-15 ***
## Year2007           -0.13087    0.01643   -7.97  1.7e-15 ***
## Year2008           -0.13584    0.01633   -8.32  < 2e-16 ***
## Year2009           -0.14213    0.01611   -8.82  < 2e-16 ***
## Year2010           -0.13425    0.01614   -8.32  < 2e-16 ***
## Year2011           -0.15149    0.01600   -9.47  < 2e-16 ***
## Year2012           -0.14650    0.01617   -9.06  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.53
## Multiple R-squared:  0.129,  Adjusted R-squared:  0.129
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
##  4 observations c(1473,1601,4165,6385)
##   are outliers with |weight| = 0 ( < 1.6e-06);
##  5144 weights are ~= 1. The remaining 55667 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0023  0.8550  0.9500  0.9040  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.64e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.044  1           1.022
```

```
## LastAuthorFemale  1.038  1           1.019
## Year              1.014 16           1.000
```



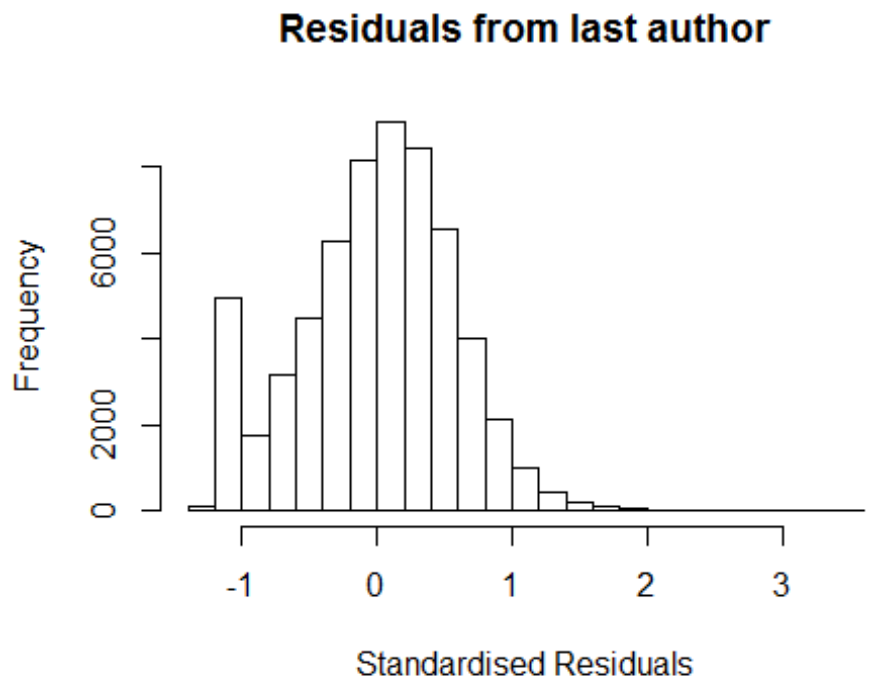

```
## [1] "List of  4 outliers with residuals above  2.5"
##         ScopusId  NLCS Year OneField Fields residuals
## 2568  0030096228 4.564 1996     2739      1     3.456
## 2797  0029680562 4.006 1996     2719      2     2.860
## 7132  0030881741 3.884 1997     2739      2     2.712
## 10924 0031613172 3.817 1998     2739      1     2.681
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.2325 -0.3723  0.0356  0.3801  3.4558
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)          1.14604    0.01575   72.74  < 2e-16 ***
## FirstAuthorFemale1   0.06058    0.00490   12.36  < 2e-16 ***
## LastAuthorFemale1   -0.03788    0.00496   -7.64  2.2e-14 ***
## Year1997             0.02587    0.02138    1.21   0.2262
## Year1998            -0.03276    0.02088   -1.57   0.1166
## Year1999             0.02067    0.01974    1.05   0.2950
```

```
## Year2000           -0.00321    0.01948   -0.16   0.8690
## Year2001           -0.00473    0.02013   -0.24   0.8142
## Year2002           -0.03616    0.01894   -1.91   0.0562 .
## Year2003           -0.03732    0.01891   -1.97   0.0485 *
## Year2004           -0.00676    0.01878   -0.36   0.7187
## Year2005           -0.04539    0.01839   -2.47   0.0136 *
## Year2006           -0.03291    0.01786   -1.84   0.0654 .
## Year2007           -0.02998    0.01780   -1.68   0.0921 .
## Year2008           -0.03546    0.01770   -2.00   0.0451 *
## Year2009           -0.04077    0.01749   -2.33   0.0197 *
## Year2010           -0.02857    0.01754   -1.63   0.1034
## Year2011           -0.04477    0.01736   -2.58   0.0099 **
## Year2012           -0.03105    0.01753   -1.77   0.0766 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.551
## Multiple R-squared:  0.00379,    Adjusted R-squared:  0.0035
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
##  4 observations c(1473,1601,4165,6385)
##   are outliers with |weight| = 0 ( < 1.6e-06);
##  5097 weights are ~= 1. The remaining 55714 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0095  0.8580  0.9490  0.9000  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.64e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.011  1           1.006
## Year              1.011 16           1.000
```




```
## [1] "List of  4 outliers with residuals above  2.5"
##         ScopusId  NLCS Year OneField Fields residuals
## 2568  0030096228 4.564 1996     2739      1     3.456
## 2797  0029680562 4.006 1996     2719      2     2.860
## 7132  0030881741 3.884 1997     2739      2     2.712
## 10924 0031613172 3.817 1998     2739      1     2.681
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##    Min     1Q Median     3Q    Max
## -1.212 -0.372  0.035  0.381  3.427
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.13672    0.01567   72.56   <2e-16 ***
## FirstAuthorFemale1  0.04988    0.00486   10.27   <2e-16 ***
## Year1997            0.02528    0.02136    1.18   0.2367
## Year1998           -0.03281    0.02088   -1.57   0.1160
## Year1999            0.02032    0.01973    1.03   0.3031
## Year2000           -0.00431    0.01948   -0.22   0.8249
## Year2001           -0.00610    0.02012   -0.30   0.7617
## Year2002           -0.03739    0.01893   -1.98   0.0482 *
## Year2003           -0.03858    0.01890   -2.04   0.0412 *
```

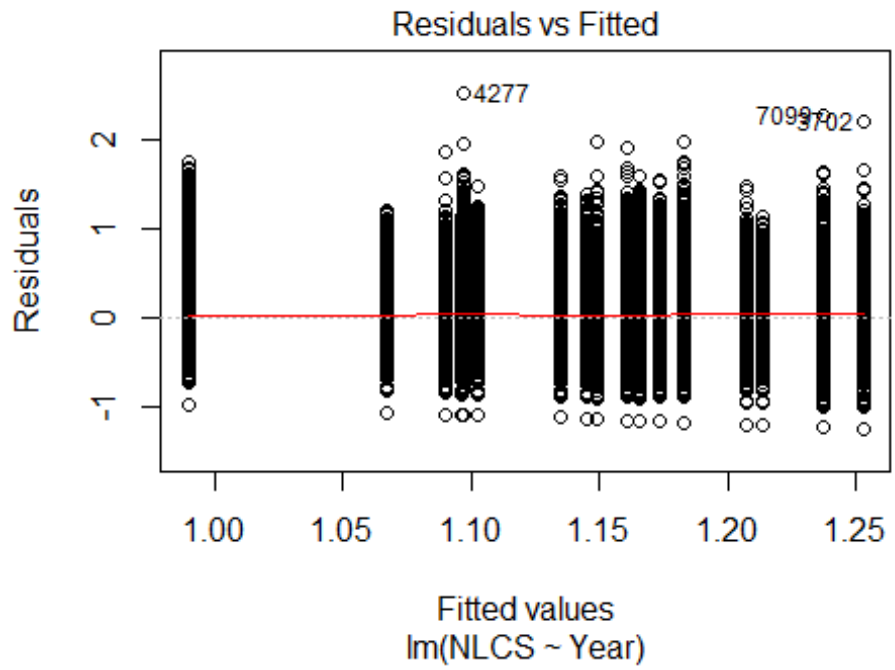
```
## Year2004           -0.00681    0.01877   -0.36   0.7167
## Year2005           -0.04612    0.01838   -2.51   0.0121 *
## Year2006           -0.03408    0.01785   -1.91   0.0561 .
## Year2007           -0.03141    0.01778   -1.77   0.0773 .
## Year2008           -0.03694    0.01769   -2.09   0.0368 *
## Year2009           -0.04230    0.01747   -2.42   0.0155 *
## Year2010           -0.03021    0.01753   -1.72   0.0848 .
## Year2011           -0.04658    0.01734   -2.69   0.0072 **
## Year2012           -0.03368    0.01752   -1.92   0.0545 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.551
## Multiple R-squared:  0.0029, Adjusted R-squared:  0.00263
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
##  4 observations c(1473,1601,4165,6385)
##   are outliers with |weight| = 0 ( < 1.6e-06);
##  5063 weights are ~= 1. The remaining 55748 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0088  0.8590  0.9490  0.9000  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.64e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##                   GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.005  1           1.003
## Year             1.005 16           1.000
```



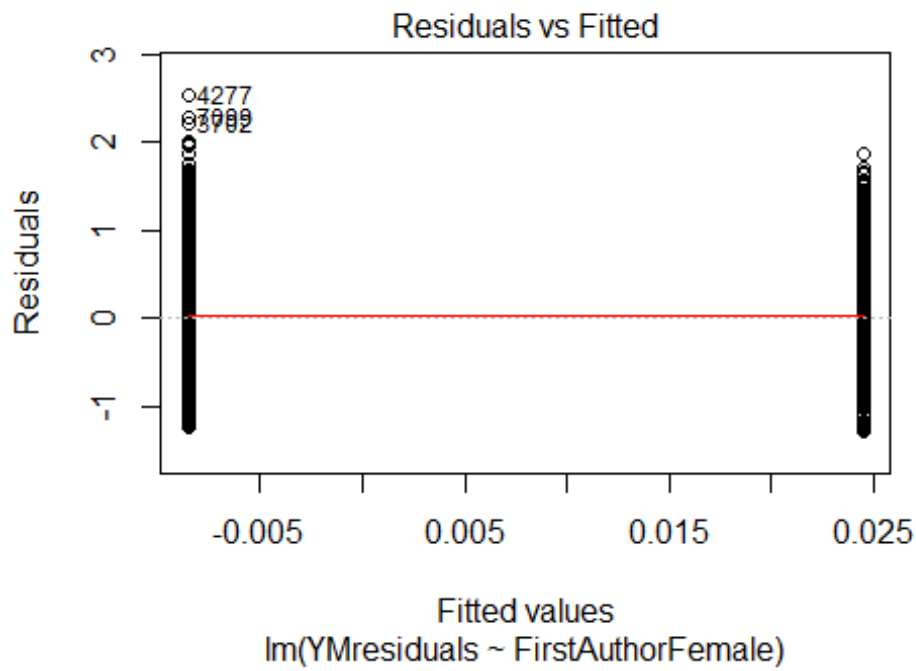

```
## [1] "List of  4 outliers with residuals above  2.5"
##         ScopusId  NLCS Year OneField Fields residuals
## 2568  0030096228 4.564 1996     2739      1     3.456
## 2797  0029680562 4.006 1996     2719      2     2.860
## 7132  0030881741 3.884 1997     2739      2     2.712
## 10924 0031613172 3.817 1998     2739      1     2.681
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.1894 -0.3741  0.0363  0.3804  3.4214
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)        1.163047   0.015704   74.06  < 2e-16 ***
## LastAuthorFemale1 -0.020400   0.004905   -4.16  3.2e-05 ***
## Year1997           0.026385   0.021434    1.23    0.218
## Year1998          -0.030381   0.020926   -1.45    0.147
## Year1999           0.023029   0.019759    1.17    0.244
## Year2000           0.001012   0.019518    0.05    0.959
## Year2001          -0.000911   0.020126   -0.05    0.964
## Year2002          -0.032913   0.018968   -1.74    0.083 .
## Year2003          -0.032954   0.018948   -1.74    0.082 .
```

```
## Year2004          -0.001458   0.018818   -0.08    0.938
## Year2005          -0.038844   0.018426   -2.11    0.035 *
## Year2006          -0.027434   0.017887   -1.53    0.125
## Year2007          -0.024370   0.017826   -1.37    0.172
## Year2008          -0.029086   0.017726   -1.64    0.101
## Year2009          -0.032970   0.017512   -1.88    0.060 .
## Year2010          -0.020303   0.017561   -1.16    0.248
## Year2011          -0.036201   0.017369   -2.08    0.037 *
## Year2012          -0.022479   0.017552   -1.28    0.200
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.551
## Multiple R-squared:  0.00134,    Adjusted R-squared:  0.00106
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
##  4 observations c(1473,1601,4165,6385)
##   are outliers with |weight| = 0 ( < 1.6e-06);
##  5096 weights are ~= 1. The remaining 55715 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.014   0.860   0.949   0.900   0.985   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.64e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Sample size for the above analysis:  60815"
## [1] ""
## [1] ""
## [1] "####################################"
## [1] "Analysis of AJSC 2740"
## [1] "####################################"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 2039 1672 1897 1653 1847 1096 1641 1268 1311 1441 1573 1837 1865 2090 2151
## 2011 2012
```

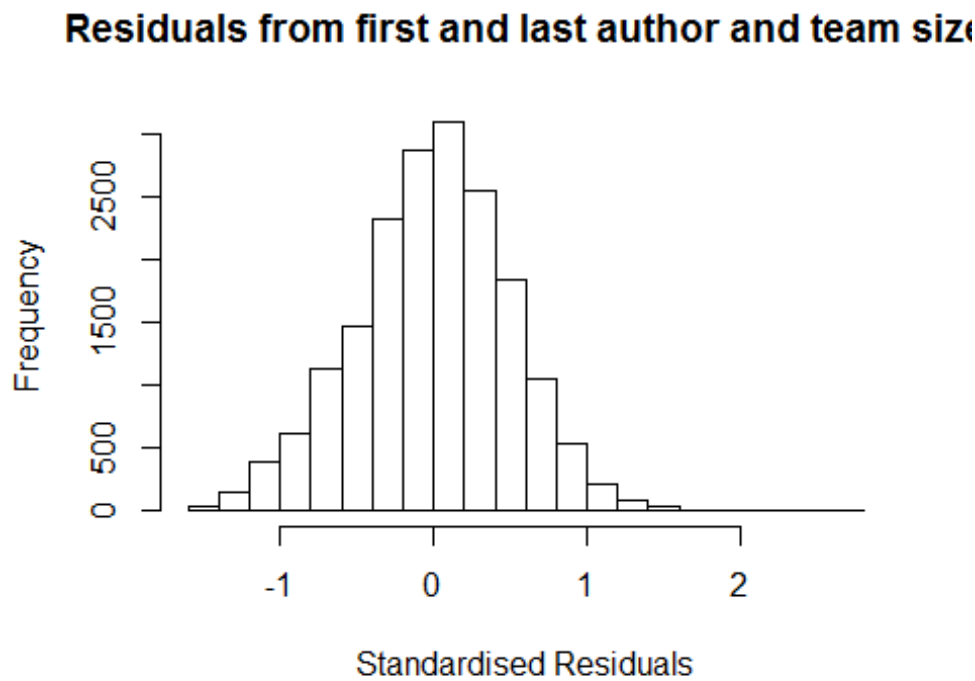
```
## 2219 2163
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 1264  997 1147 1076 1129  699 1282 1002 1011 1109 1208 1422 1441 1591 1687
## 2011 2012
## 1713 1705
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  870  889  878  921  967  611 1134  874  893  961 1036 1277 1250 1410 1438
## 2011 2012
## 1485 1427
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
##  Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 620, df = 16, p-value <2e-16
```




```
##
##  Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 49, df = 1, p-value = 3e-12
```

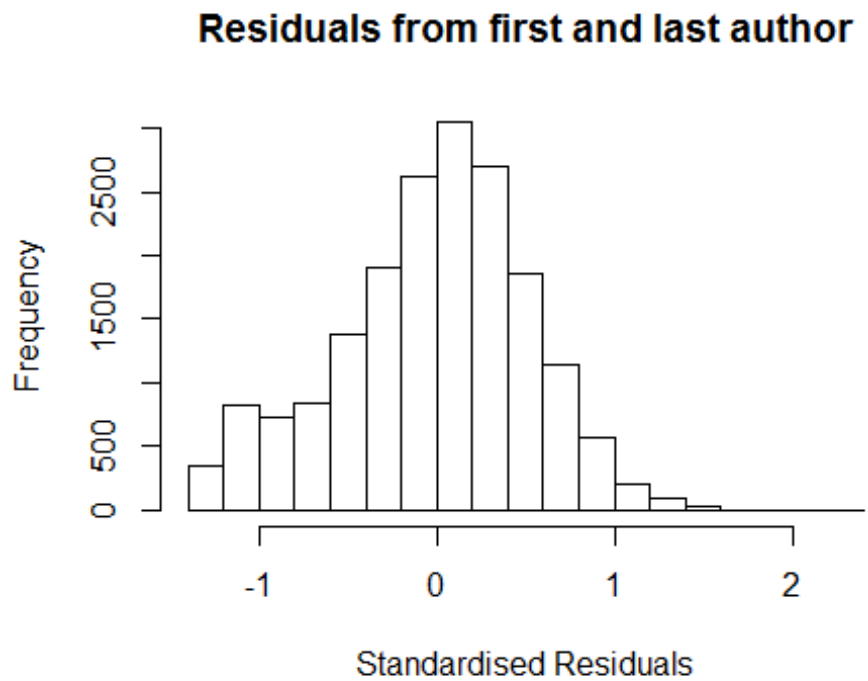



```
## [1] "Female first author team size 2018 geometric mean: 5.4650665592614"
## [1] "Male first author team size 2018 geometric mean: 5.46693074159695"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 290000, p-value = 0.3
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 5.61145414614268"
## [1] "Male last author team size 2018 geometric mean: 5.41573270932997"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 260000, p-value = 0.4
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.053  1           1.026
## LastAuthorFemale  1.039  1           1.019
## UniqueAuthors     1.041  4           1.005
## Year              1.044 16           1.001
```



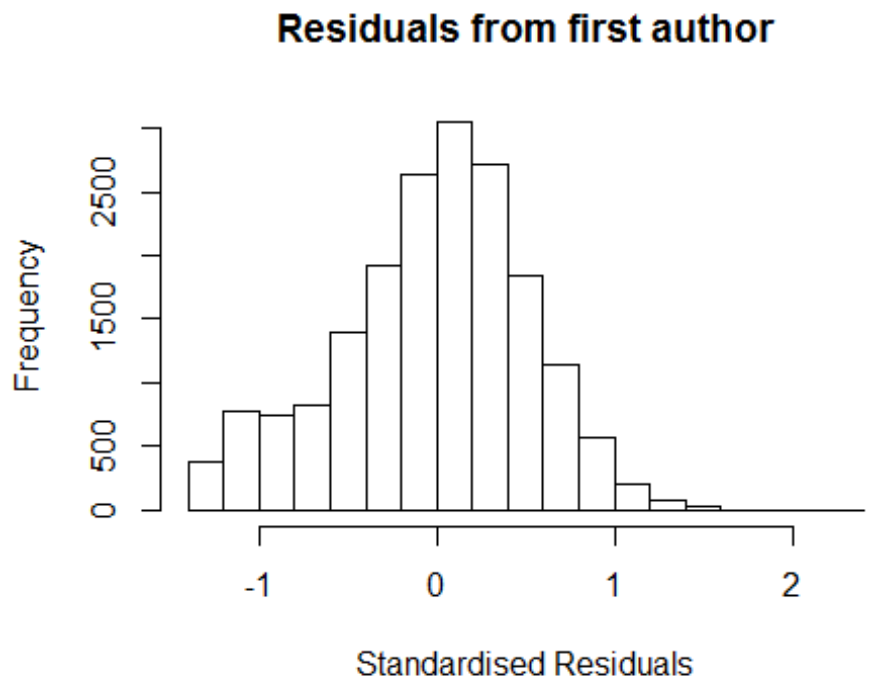

```
## [1] "List of  2 outliers with residuals above  2.5"
##        ScopusId  NLCS Year OneField Fields residuals
## 3702 0030912190 3.457 1997     2705      3     2.610
## 4277 0031761535 3.615 1998     2700      3     2.513
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
UniqueAuthors +
##     Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##     k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.5608 -0.3295  0.0139  0.3257  2.6095
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         0.85569    0.02623   32.62  < 2e-16 ***
## FirstAuthorFemale1  0.01023    0.00848    1.21   0.2277
## LastAuthorFemale1   0.00529    0.01028    0.51   0.6070
## UniqueAuthors2      0.20920    0.02088   10.02  < 2e-16 ***
## UniqueAuthors3      0.34130    0.01918   17.79  < 2e-16 ***
## UniqueAuthors4      0.41401    0.01879   22.03  < 2e-16 ***
## UniqueAuthors5      0.65302    0.01667   39.18  < 2e-16 ***
## Year1997           -0.00820    0.02807   -0.29   0.7702
## Year1998            0.03662    0.02870    1.28   0.2019
```

```
## Year1999           -0.00621    0.02837   -0.22   0.8266
## Year2000           -0.08004    0.02626   -3.05   0.0023 **
## Year2001           -0.07948    0.02867   -2.77   0.0056 **
## Year2002           -0.15757    0.02562   -6.15  7.9e-10 ***
## Year2003           -0.21055    0.02681   -7.85  4.2e-15 ***
## Year2004           -0.20308    0.02658   -7.64  2.3e-14 ***
## Year2005           -0.26240    0.02670   -9.83  < 2e-16 ***
## Year2006           -0.23651    0.02628   -9.00  < 2e-16 ***
## Year2007           -0.17604    0.02583   -6.81  9.8e-12 ***
## Year2008           -0.20015    0.02571   -7.78  7.4e-15 ***
## Year2009           -0.16649    0.02565   -6.49  8.7e-11 ***
## Year2010           -0.18562    0.02567   -7.23  4.9e-13 ***
## Year2011           -0.16252    0.02583   -6.29  3.2e-10 ***
## Year2012           -0.17761    0.02571   -6.91  5.1e-12 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.484
## Multiple R-squared:  0.176,  Adjusted R-squared:  0.175
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
##  3 observations c(1558,1826,3044)
##   are outliers with |weight| = 0 ( < 5.5e-06);
##  1501 weights are ~= 1. The remaining 16817 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.110   0.865   0.951   0.903   0.986   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          5.46e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.038  1           1.019
## LastAuthorFemale  1.031  1           1.015
## Year              1.017 16           1.001
```



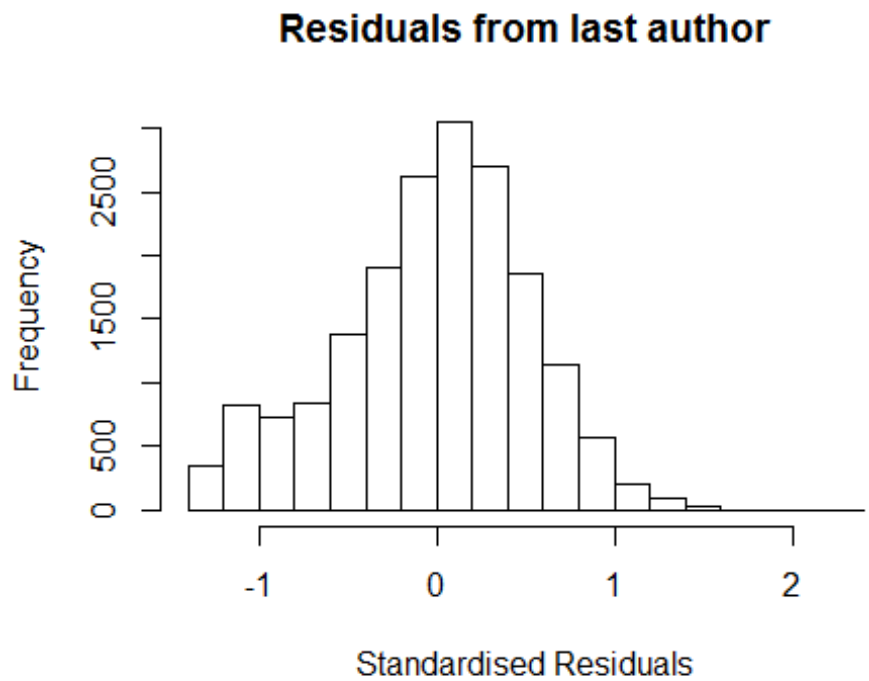

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.3411 -0.3453  0.0327  0.3429  2.3026
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)          1.27429    0.02342   54.40  < 2e-16 ***
## FirstAuthorFemale1   0.02876    0.00900    3.20  0.00140 **
## LastAuthorFemale1   -0.01782    0.01082   -1.65  0.09970 .
## Year1997            -0.00567    0.03020   -0.19  0.85095
## Year1998             0.03806    0.03074    1.24  0.21560
## Year1999             0.01502    0.03029    0.50  0.62011
## Year2000            -0.04856    0.02796   -1.74  0.08244 .
## Year2001            -0.03749    0.03026   -1.24  0.21544
## Year2002            -0.12103    0.02750   -4.40  1.1e-05 ***
## Year2003            -0.15710    0.02902   -5.41  6.2e-08 ***
## Year2004            -0.17459    0.02851   -6.12  9.3e-10 ***
## Year2005            -0.20901    0.02875   -7.27  3.7e-13 ***
```

```
## Year2006           -0.17291    0.02823   -6.12  9.3e-10 ***
## Year2007           -0.11745    0.02788   -4.21  2.5e-05 ***
## Year2008           -0.12750    0.02800   -4.55  5.3e-06 ***
## Year2009           -0.08792    0.02827   -3.11  0.00188 **
## Year2010           -0.10344    0.02824   -3.66  0.00025 ***
## Year2011           -0.07570    0.02788   -2.72  0.00662 **
## Year2012           -0.08793    0.02784   -3.16  0.00159 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.506
## Multiple R-squared:  0.0148, Adjusted R-squared:  0.0139
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
##  1520 weights are ~= 1. The remaining 16801 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0032  0.8540  0.9480  0.8950  0.9860  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          5.46e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.013  1           1.006
## Year              1.013 16           1.000
```




```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##    Min     1Q Median     3Q    Max
## -1.336 -0.347  0.033  0.342  2.305
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)          1.27232    0.02340   54.38  < 2e-16 ***
## FirstAuthorFemale1   0.02586    0.00890    2.91  0.00367 **
## Year1997            -0.00569    0.03022   -0.19  0.85076
## Year1998             0.03806    0.03076    1.24  0.21606
## Year1999             0.01481    0.03032    0.49  0.62534
## Year2000            -0.04880    0.02797   -1.74  0.08107 .
## Year2001            -0.03732    0.03027   -1.23  0.21766
## Year2002            -0.12166    0.02751   -4.42  9.8e-06 ***
## Year2003            -0.15749    0.02903   -5.43  5.8e-08 ***
## Year2004            -0.17497    0.02852   -6.14  8.7e-10 ***
## Year2005            -0.20928    0.02876   -7.28  3.6e-13 ***
## Year2006            -0.17332    0.02825   -6.14  8.7e-10 ***
```

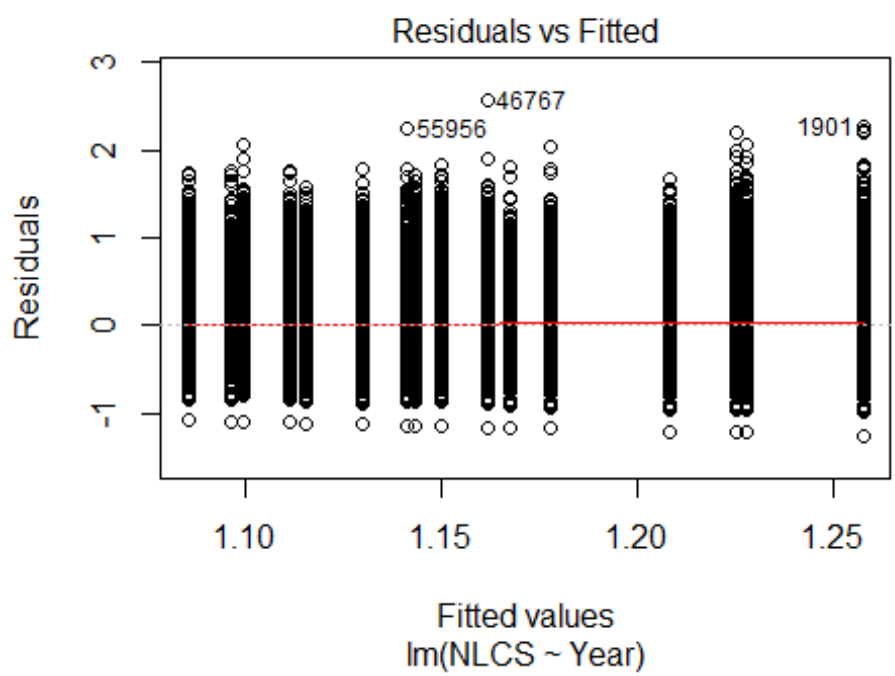
```
## Year2007           -0.11790    0.02790   -4.23  2.4e-05 ***
## Year2008           -0.12800    0.02801   -4.57  4.9e-06 ***
## Year2009           -0.08844    0.02829   -3.13  0.00177 **
## Year2010           -0.10387    0.02826   -3.67  0.00024 ***
## Year2011           -0.07619    0.02790   -2.73  0.00632 **
## Year2012           -0.08854    0.02785   -3.18  0.00148 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.506
## Multiple R-squared:  0.0147, Adjusted R-squared:  0.0138
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
##  1519 weights are ~= 1. The remaining 16802 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0031  0.8540  0.9480  0.8950  0.9860  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          5.46e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##                   GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.005  1           1.003
## Year             1.005 16           1.000
```



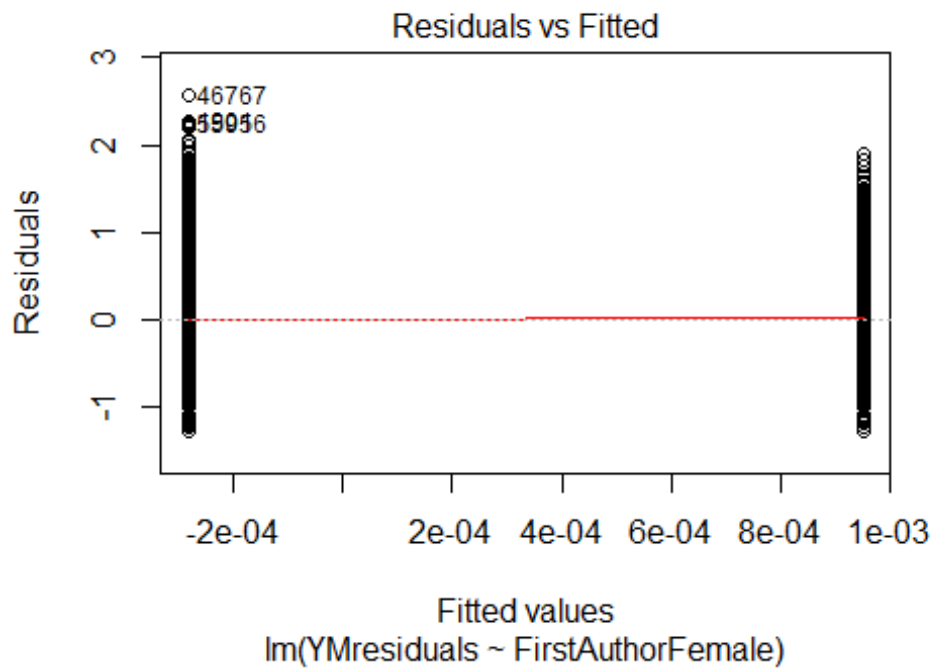

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.3174 -0.3484  0.0338  0.3430  2.2976
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.27937    0.02333   54.84  < 2e-16 ***
## LastAuthorFemale1 -0.01142    0.01067   -1.07  0.28458
## Year1997          -0.00578    0.03017   -0.19  0.84811
## Year1998           0.03803    0.03072    1.24  0.21581
## Year1999           0.01563    0.03028    0.52  0.60579
## Year2000          -0.04801    0.02793   -1.72  0.08573 .
## Year2001          -0.03677    0.03021   -1.22  0.22362
## Year2002          -0.11993    0.02748   -4.36  1.3e-05 ***
## Year2003          -0.15579    0.02899   -5.37  7.8e-08 ***
## Year2004          -0.17301    0.02848   -6.07  1.3e-09 ***
## Year2005          -0.20712    0.02871   -7.21  5.7e-13 ***
## Year2006          -0.17065    0.02820   -6.05  1.5e-09 ***
```

```
## Year2007          -0.11567    0.02785   -4.15  3.3e-05 ***
## Year2008          -0.12488    0.02794   -4.47  7.9e-06 ***
## Year2009          -0.08612    0.02825   -3.05  0.00231 **
## Year2010          -0.10099    0.02821   -3.58  0.00034 ***
## Year2011          -0.07252    0.02782   -2.61  0.00915 **
## Year2012          -0.08475    0.02780   -3.05  0.00230 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.506
## Multiple R-squared:  0.0143, Adjusted R-squared:  0.0134
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
##  1538 weights are ~= 1. The remaining 16783 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0037  0.8540  0.9480  0.8950  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          5.46e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Sample size for the above analysis:  18321"
## [1] ""
## [1] ""
## [1] "####################################"
## [1] "Analysis of AJSC 2741"
## [1] "####################################"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 4263 4399 3976 3858 4125 4000 3883 3101 3176 4279 3946 5161 4509 4582 4392
## 2011 2012
## 4779 4962
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 2285 2474 2216 2256 2305 1592 2789 2198 2259 2345 2736 2890 3153 3200 3133
## 2011 2012
```

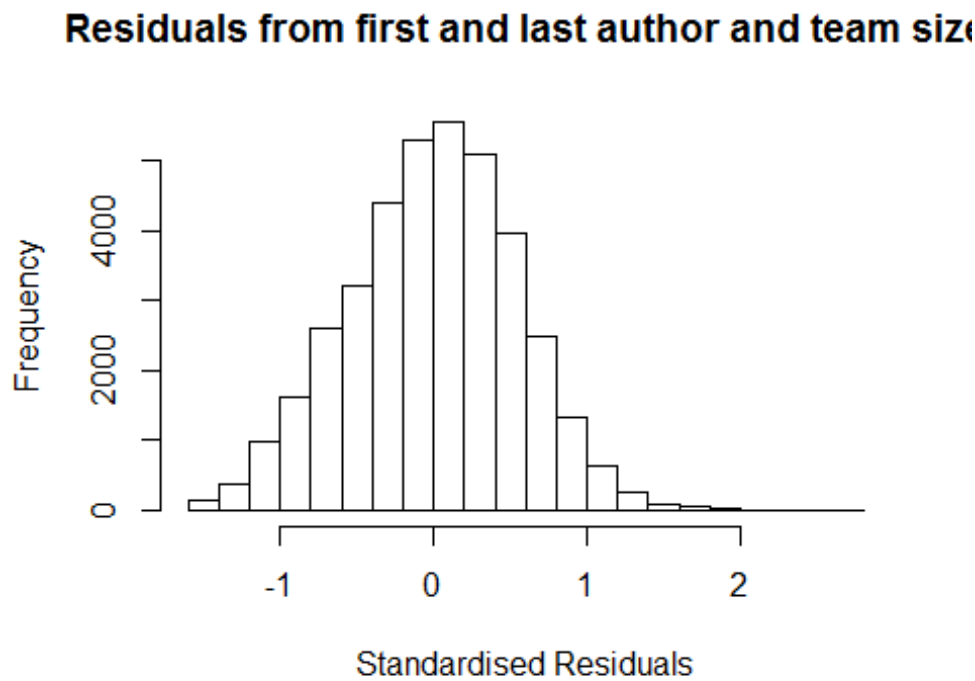
```
## 3403 3511
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 2023 2189 1939 1986 2015 1336 2396 1847 1931 2033 2260 2378 2652 2691 2639
## 2011 2012
## 2860 2950
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
##  Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 330, df = 16, p-value <2e-16
```




```
##
##  Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 0.55, df = 1, p-value = 0.5
```

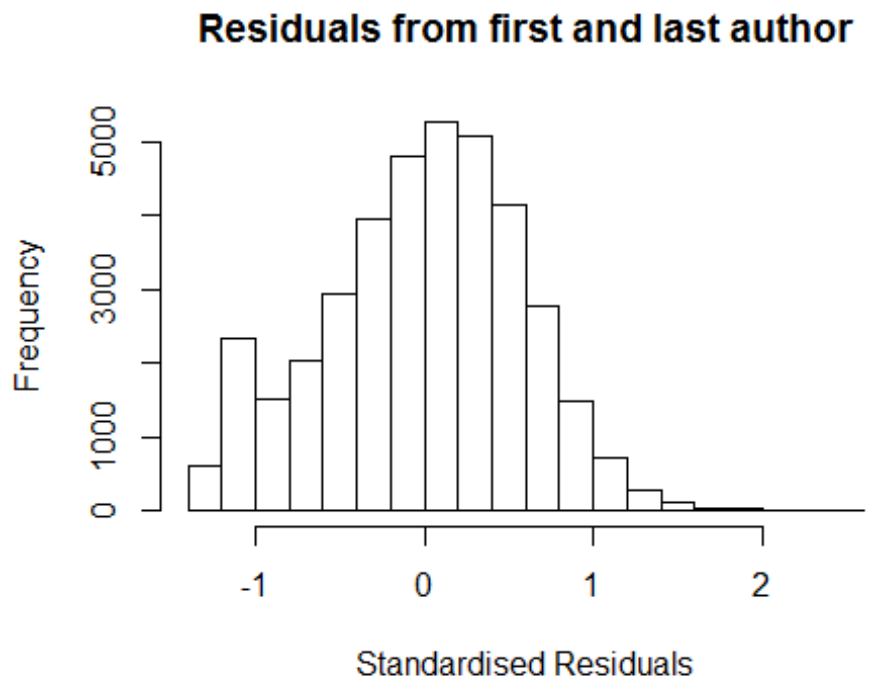



```
## [1] "Female first author team size 2018 geometric mean: 4.86020412026787"
## [1] "Male first author team size 2018 geometric mean: 4.90523564351987"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 1300000, p-value = 1
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 4.63164413755934"
## [1] "Male last author team size 2018 geometric mean: 4.97197967693065"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 1e+06, p-value = 0.05
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.030  1           1.015
## LastAuthorFemale  1.033  1           1.016
## UniqueAuthors     1.023  4           1.003
## Year              1.027 16           1.001
```



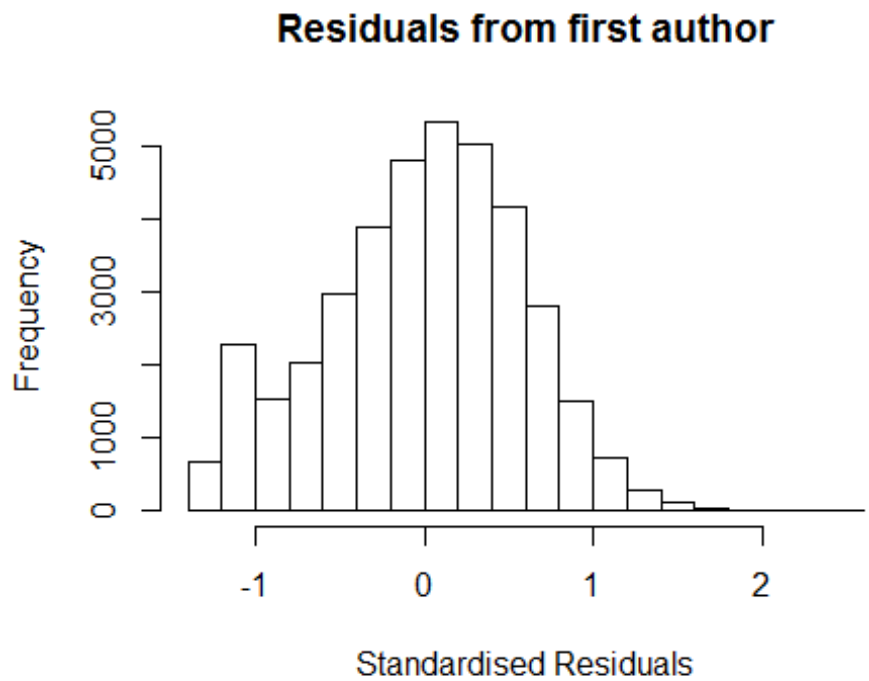

```
## [1] "List of  1 outliers with residuals above  2.5"
##          ScopusId NLCS Year OneField Fields residuals
## 46767 36849088522 3.73 2007     2741      1     2.678
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
UniqueAuthors +
##     Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##     k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.4996 -0.3668  0.0174  0.3703  2.6776
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         0.81450    0.01759   46.31  < 2e-16 ***
## FirstAuthorFemale1 -0.00678    0.00686   -0.99  0.32328
## LastAuthorFemale1  -0.03463    0.00800   -4.33  1.5e-05 ***
## UniqueAuthors2      0.25880    0.01380   18.75  < 2e-16 ***
## UniqueAuthors3      0.36299    0.01303   27.86  < 2e-16 ***
## UniqueAuthors4      0.49272    0.01285   38.35  < 2e-16 ***
## UniqueAuthors5      0.68511    0.01138   60.18  < 2e-16 ***
## Year1997           -0.03309    0.02127   -1.56  0.11981
## Year1998           -0.02949    0.01964   -1.50  0.13336
## Year1999           -0.04765    0.01897   -2.51  0.01202 *
```

```
## Year2000           -0.07260    0.01906   -3.81  0.00014 ***
## Year2001           -0.14427    0.02121   -6.80  1.0e-11 ***
## Year2002           -0.18295    0.01827  -10.02  < 2e-16 ***
## Year2003           -0.17997    0.01903   -9.46  < 2e-16 ***
## Year2004           -0.18728    0.01869  -10.02  < 2e-16 ***
## Year2005           -0.12702    0.01837   -6.91  4.8e-12 ***
## Year2006           -0.16641    0.01834   -9.07  < 2e-16 ***
## Year2007           -0.12506    0.01810   -6.91  4.9e-12 ***
## Year2008           -0.14465    0.01758   -8.23  < 2e-16 ***
## Year2009           -0.16259    0.01783   -9.12  < 2e-16 ***
## Year2010           -0.18072    0.01783  -10.13  < 2e-16 ***
## Year2011           -0.19918    0.01744  -11.42  < 2e-16 ***
## Year2012           -0.22309    0.01750  -12.75  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.541
## Multiple R-squared:  0.152,  Adjusted R-squared:  0.152
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
##  observation 22011 is an outlier with |weight| = 0 ( < 2.6e-06);
##  3078 weights are ~= 1. The remaining 35046 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0062  0.8710  0.9490  0.9060  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          2.62e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.012  1           1.006
## LastAuthorFemale  1.013  1           1.006
## Year              1.009 16           1.000
```



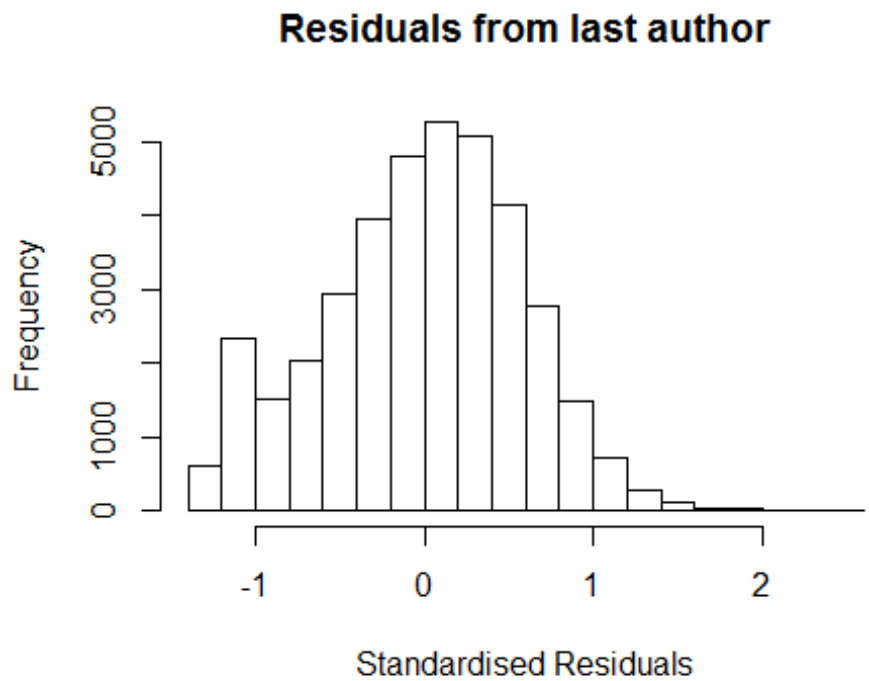

```
## [1] "List of  1 outliers with residuals above  2.5"
##          ScopusId NLCS Year OneField Fields residuals
## 46767 36849088522 3.73 2007     2741      1     2.562
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.2904 -0.3913  0.0342  0.4024  2.5616
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)          1.27476    0.01539   82.81  < 2e-16 ***
## FirstAuthorFemale1  0.01567    0.00747    2.10  0.03580 *
## LastAuthorFemale1  -0.06612    0.00891   -7.42  1.2e-13 ***
## Year1997           -0.03127    0.02295   -1.36  0.17295
## Year1998           -0.03323    0.02108   -1.58  0.11498
## Year1999           -0.04961    0.02019   -2.46  0.01401 *
## Year2000           -0.07758    0.02039   -3.80  0.00014 ***
## Year2001           -0.12941    0.02218   -5.83  5.4e-09 ***
## Year2002           -0.18818    0.01993   -9.44  < 2e-16 ***
## Year2003           -0.15969    0.02061   -7.75  9.6e-15 ***
## Year2004           -0.16574    0.02023   -8.19  2.6e-16 ***
## Year2005           -0.09926    0.01977   -5.02  5.2e-07 ***
```

```
## Year2006           -0.14681    0.01966   -7.47  8.3e-14 ***
## Year2007           -0.10633    0.01949   -5.45  4.9e-08 ***
## Year2008           -0.11349    0.01889   -6.01  1.9e-09 ***
## Year2009           -0.11351    0.01908   -5.95  2.7e-09 ***
## Year2010           -0.12789    0.01927   -6.64  3.2e-11 ***
## Year2011           -0.15434    0.01885   -8.19  2.8e-16 ***
## Year2012           -0.16317    0.01891   -8.63  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.58
## Multiple R-squared:  0.00925,    Adjusted R-squared:  0.00878
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
##  3171 weights are ~= 1. The remaining 34954 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0125  0.8640  0.9480  0.9060  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          2.62e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.005  1           1.002
## Year              1.005 16           1.000
```




```
## [1] "List of  1 outliers with residuals above  2.5"
##          ScopusId NLCS Year OneField Fields residuals
## 46767 36849088522 3.73 2007     2741      1     2.562
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.2720 -0.3951  0.0361  0.4029  2.5704
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)          1.26620    0.01533   82.61  < 2e-16 ***
## FirstAuthorFemale1  0.00583    0.00751    0.78  0.43735
## Year1997           -0.03120    0.02293   -1.36  0.17366
## Year1998           -0.03291    0.02109   -1.56  0.11869
## Year1999           -0.04794    0.02018   -2.38  0.01750 *
## Year2000           -0.07654    0.02037   -3.76  0.00017 ***
## Year2001           -0.12735    0.02217   -5.74  9.4e-09 ***
## Year2002           -0.18655    0.01994   -9.36  < 2e-16 ***
## Year2003           -0.15903    0.02062   -7.71  1.3e-14 ***
## Year2004           -0.16491    0.02021   -8.16  3.5e-16 ***
## Year2005           -0.09835    0.01976   -4.98  6.5e-07 ***
## Year2006           -0.14672    0.01965   -7.47  8.3e-14 ***
```

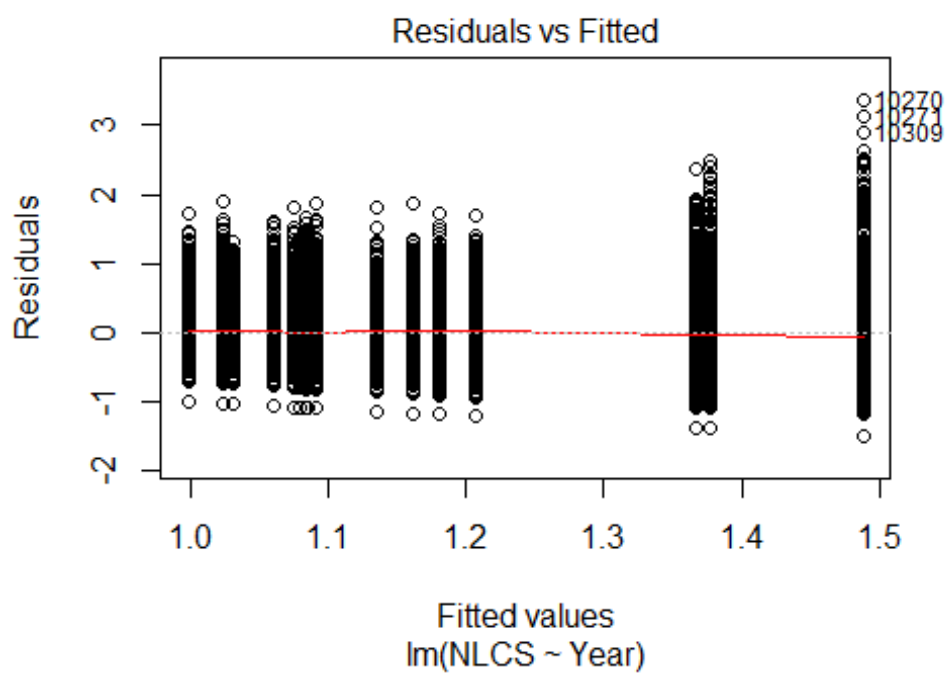
```
## Year2007           -0.10662    0.01949   -5.47  4.5e-08 ***
## Year2008           -0.11291    0.01888   -5.98  2.2e-09 ***
## Year2009           -0.11465    0.01907   -6.01  1.9e-09 ***
## Year2010           -0.12799    0.01925   -6.65  3.0e-11 ***
## Year2011           -0.15608    0.01884   -8.28  < 2e-16 ***
## Year2012           -0.16473    0.01889   -8.72  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.58
## Multiple R-squared:  0.00759,    Adjusted R-squared:  0.00714
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
##  3155 weights are ~= 1. The remaining 34970 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0112  0.8640  0.9490  0.9050  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          2.62e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##                   GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.006  1           1.003
## Year             1.006 16           1.000
```



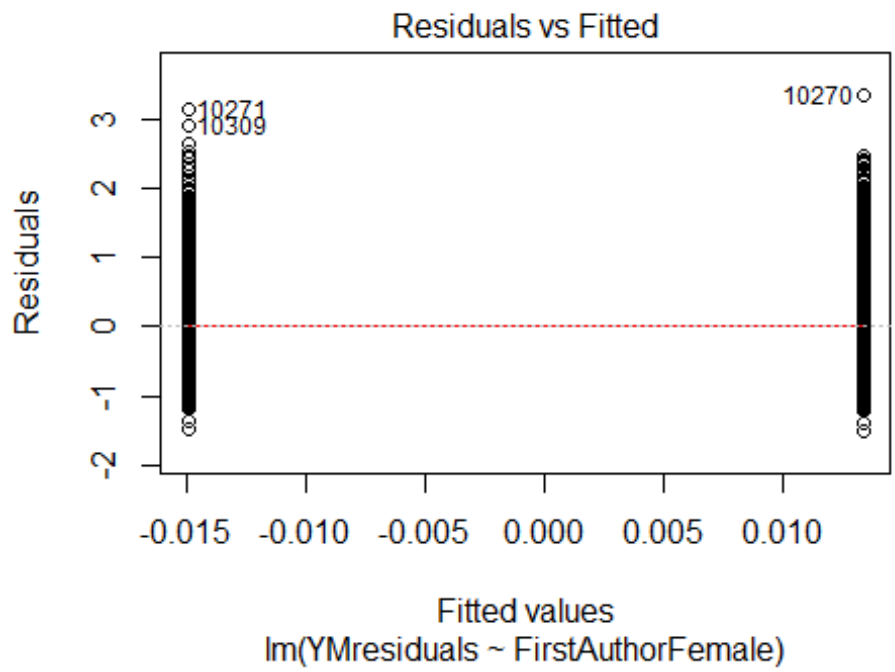

```
## [1] "List of  1 outliers with residuals above  2.5"
##          ScopusId NLCS Year OneField Fields residuals
## 46767 36849088522 3.73 2007     2741      1     2.562
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.2778 -0.3927  0.0347  0.4023  2.5584
##
## Coefficients:
##                   Estimate Std. Error t value Pr(>|t|)
## (Intercept)        1.27777    0.01530   83.54  < 2e-16 ***
## LastAuthorFemale1 -0.06303    0.00889   -7.09  1.4e-12 ***
## Year1997          -0.03150    0.02295   -1.37  0.16990
## Year1998          -0.03335    0.02109   -1.58  0.11376
## Year1999          -0.04969    0.02019   -2.46  0.01386 *
## Year2000          -0.07788    0.02039   -3.82  0.00013 ***
## Year2001          -0.12946    0.02217   -5.84  5.3e-09 ***
## Year2002          -0.18820    0.01993   -9.44  < 2e-16 ***
## Year2003          -0.16001    0.02061   -7.76  8.6e-15 ***
## Year2004          -0.16574    0.02022   -8.20  2.6e-16 ***
## Year2005          -0.09908    0.01977   -5.01  5.4e-07 ***
## Year2006          -0.14700    0.01966   -7.48  7.7e-14 ***
```

```
## Year2007          -0.10622    0.01949   -5.45  5.1e-08 ***
## Year2008          -0.11300    0.01889   -5.98  2.2e-09 ***
## Year2009          -0.11310    0.01909   -5.93  3.1e-09 ***
## Year2010          -0.12758    0.01927   -6.62  3.6e-11 ***
## Year2011          -0.15399    0.01885   -8.17  3.2e-16 ***
## Year2012          -0.16264    0.01891   -8.60  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.58
## Multiple R-squared:  0.00915,    Adjusted R-squared:  0.00871
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
##  3166 weights are ~= 1. The remaining 34959 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0132  0.8640  0.9490  0.9060  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          2.62e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Sample size for the above analysis:  38125"
## [1] ""
## [1] ""
## [1] "####################################"
## [1] "Analysis of AJSC 2742"
## [1] "####################################"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  995 1074  998  891  982 1155  987  922  845 1108 1095 1214 1410 1596 1674
## 2011 2012
## 1627 1675
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  751  759  702  674  616  655  832  804  718  931  922 1042 1215 1344 1376
## 2011 2012
```

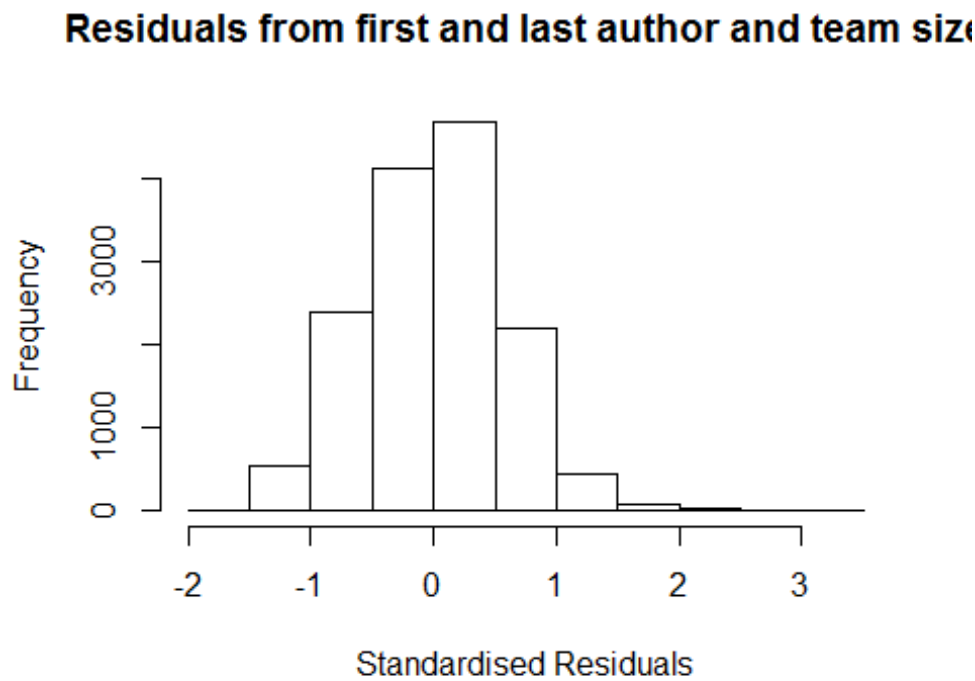
```
## 1357 1391
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  687  686  616  604  567  584  742  729  649  841  827  941 1087 1193 1238
## 2011 2012
## 1226 1253
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
##  Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 410, df = 16, p-value <2e-16
```




```
##
##  Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 5.4, df = 1, p-value = 0.02
```



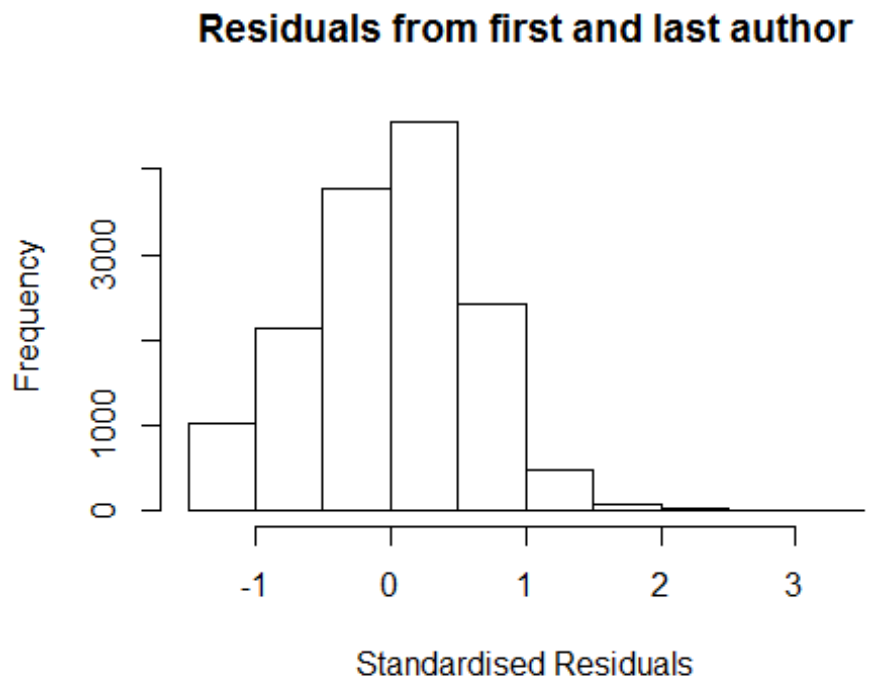

```
## [1] "Female first author team size 2018 geometric mean: 3.83474772858157"
## [1] "Male first author team size 2018 geometric mean: 4.04294182159961"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 140000, p-value = 0.3
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 3.58458994097702"
## [1] "Male last author team size 2018 geometric mean: 4.20948580911036"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 120000, p-value = 1e-04
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##                     GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.096  1           1.047
## LastAuthorFemale  1.111  1           1.054
## UniqueAuthors     1.087  4           1.011
## Year              1.074 16           1.002
```




```
## [1] "List of  5 outliers with residuals above  2.5"
##          ScopusId  NLCS Year OneField Fields residuals
## 9434   2342589377 3.867 2004     2742      1     2.608
## 10270 23644445382 4.848 2005     2742      1     3.393
## 10271 23644457022 4.605 2005     2742      1     3.136
## 10309 23944445391 4.398 2005     2742      1     2.782
## 10734 16244371358 3.961 2005     2742      1     2.699
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
UniqueAuthors +
##     Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##     k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.6159 -0.3913  0.0168  0.3888  3.3929
##
## Coefficients:
##                      Estimate Std. Error t value Pr(>|t|)
## (Intercept)          0.752389   0.027071   27.79  < 2e-16 ***
## FirstAuthorFemale1   0.013652   0.010128    1.35  0.17772
## LastAuthorFemale1   -0.062607   0.010398   -6.02  1.8e-09 ***
## UniqueAuthors2       0.333164   0.016668   19.99  < 2e-16 ***
## UniqueAuthors3       0.464036   0.016248   28.56  < 2e-16 ***
## UniqueAuthors4       0.555558   0.017041   32.60  < 2e-16 ***
```

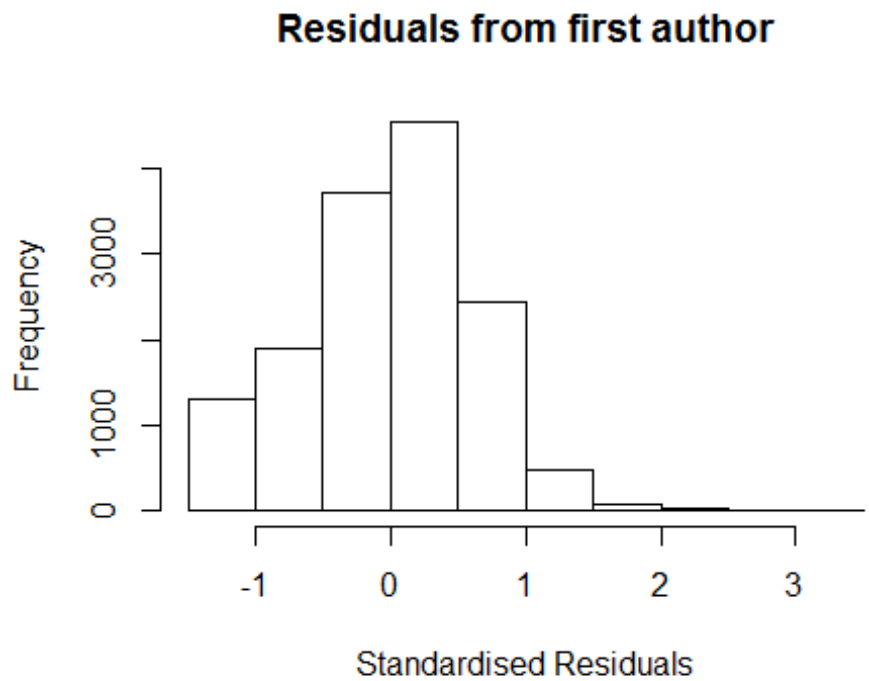
```
## UniqueAuthors5      0.673772   0.015084   44.67  < 2e-16 ***
## Year1997           -0.043172   0.034865   -1.24  0.21563
## Year1998           -0.028509   0.036028   -0.79  0.42878
## Year1999           -0.001811   0.035485   -0.05  0.95929
## Year2000            0.000922   0.033126    0.03  0.97780
## Year2001            0.034414   0.033782    1.02  0.30836
## Year2002            0.079150   0.032761    2.42  0.01571 *
## Year2003            0.033885   0.033776    1.00  0.31577
## Year2004            0.173658   0.037361    4.65  3.4e-06 ***
## Year2005            0.238691   0.035713    6.68  2.4e-11 ***
## Year2006            0.175559   0.034069    5.15  2.6e-07 ***
## Year2007           -0.083071   0.029834   -2.78  0.00537 **
## Year2008           -0.106452   0.029856   -3.57  0.00036 ***
## Year2009           -0.154180   0.029741   -5.18  2.2e-07 ***
## Year2010           -0.154130   0.029184   -5.28  1.3e-07 ***
## Year2011           -0.136692   0.029118   -4.69  2.7e-06 ***
## Year2012           -0.163888   0.029805   -5.50  3.9e-08 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.569
## Multiple R-squared:  0.183,  Adjusted R-squared:  0.182
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  4 observations c(6158,6159,6186,6475)
##   are outliers with |weight| = 0 ( < 6.9e-06);
##  1096 weights are ~= 1. The remaining 13370 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0018  0.8760  0.9490  0.9090  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          6.91e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.069  1           1.034
```

```
## LastAuthorFemale  1.060  1           1.030
## Year              1.011 16           1.000
```



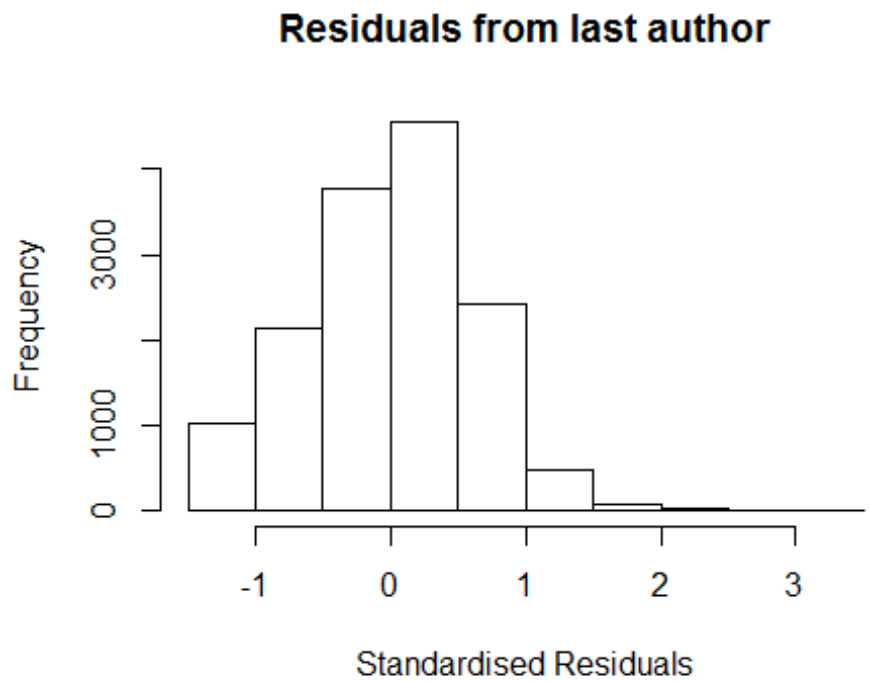

```
## [1] "List of  8 outliers with residuals above  2.5"
##          ScopusId  NLCS Year OneField Fields residuals
## 8990  12344273284 3.769 2004     2742      1     2.516
## 9860  29944433045 3.989 2005     2742      1     2.513
## 10025 29244478299 4.133 2005     2742      1     2.657
## 10165 25144493160 3.924 2005     2742      1     2.575
## 10270 23644445382 4.848 2005     2742      1     3.381
## 10271 23644457022 4.605 2005     2742      1     3.129
## 10309 23944445391 4.398 2005     2742      1     3.040
## 10734 16244371358 3.961 2005     2742      1     2.612
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.4757 -0.4313  0.0358  0.4227  3.3810
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.11537    0.02957   37.71  < 2e-16 ***
## FirstAuthorFemale1  0.00869    0.01090    0.80   0.4252
```

```
## LastAuthorFemale1  -0.11806    0.01110  -10.63  < 2e-16 ***
## Year1997           -0.04405    0.03951   -1.11   0.2649
## Year1998            0.01151    0.03968    0.29   0.7717
## Year1999           -0.00200    0.03925   -0.05   0.9593
## Year2000            0.04219    0.03688    1.14   0.2526
## Year2001            0.08397    0.03781    2.22   0.0264 *
## Year2002            0.14106    0.03628    3.89   0.0001 ***
## Year2003            0.11005    0.03809    2.89   0.0039 **
## Year2004            0.25549    0.04158    6.15  8.2e-10 ***
## Year2005            0.35167    0.03979    8.84  < 2e-16 ***
## Year2006            0.30183    0.03820    7.90  2.9e-15 ***
## Year2007            0.01162    0.03435    0.34   0.7353
## Year2008            0.00422    0.03391    0.12   0.9009
## Year2009           -0.05021    0.03407   -1.47   0.1405
## Year2010           -0.06062    0.03387   -1.79   0.0735 .
## Year2011           -0.04036    0.03340   -1.21   0.2269
## Year2012           -0.04536    0.03403   -1.33   0.1826
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.621
## Multiple R-squared:  0.0465, Adjusted R-squared:  0.0453
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
##  3 observations c(6158,6159,6186)
##   are outliers with |weight| = 0 ( < 6.9e-06);
##  1139 weights are ~= 1. The remaining 13328 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0272  0.8730  0.9500  0.9110  0.9840  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          6.91e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##                   GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.01  1           1.005
## Year              1.01 16           1.000
```




```
## [1] "List of  8 outliers with residuals above  2.5"
##          ScopusId  NLCS Year OneField Fields residuals
## 8990  12344273284 3.769 2004     2742      1     2.516
## 9860  29944433045 3.989 2005     2742      1     2.513
## 10025 29244478299 4.133 2005     2742      1     2.657
## 10165 25144493160 3.924 2005     2742      1     2.575
## 10270 23644445382 4.848 2005     2742      1     3.381
## 10271 23644457022 4.605 2005     2742      1     3.129
## 10309 23944445391 4.398 2005     2742      1     3.040
## 10734 16244371358 3.961 2005     2742      1     2.612
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.4342 -0.4328  0.0379  0.4282  3.4138
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.08651    0.02953   36.80  < 2e-16 ***
## FirstAuthorFemale1 -0.02637    0.01069   -2.47  0.01362 *
## Year1997           -0.04484    0.03965   -1.13  0.25816
## Year1998            0.01116    0.03982    0.28  0.77934
## Year1999           -0.00118    0.03931   -0.03  0.97612
```

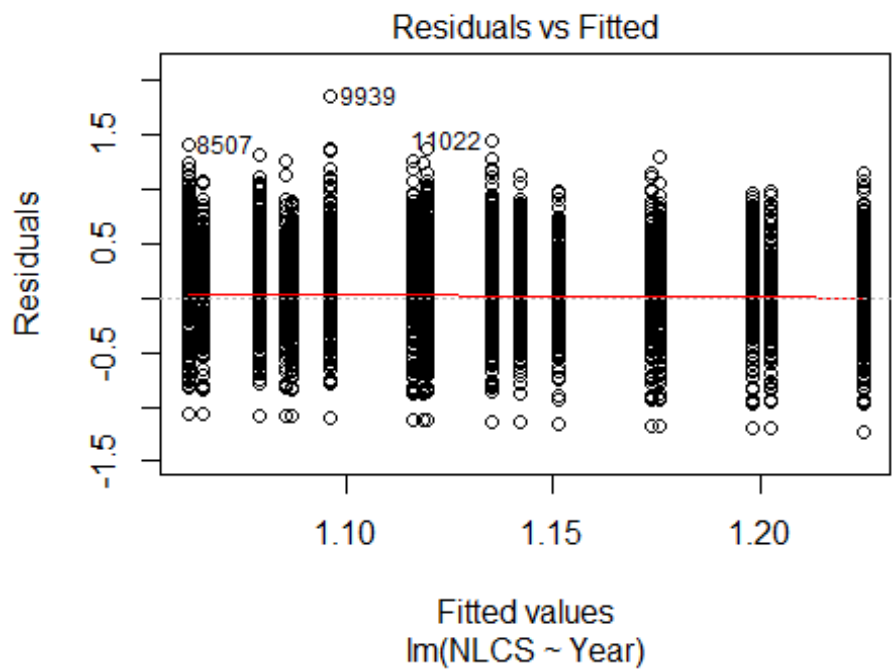
```
## Year2000            0.04170    0.03703    1.13  0.26005
## Year2001            0.08003    0.03796    2.11  0.03502 *
## Year2002            0.13651    0.03643    3.75  0.00018 ***
## Year2003            0.10817    0.03826    2.83  0.00470 **
## Year2004            0.25282    0.04183    6.04  1.5e-09 ***
## Year2005            0.34766    0.03993    8.71  < 2e-16 ***
## Year2006            0.30038    0.03841    7.82  5.6e-15 ***
## Year2007            0.00974    0.03456    0.28  0.77813
## Year2008            0.00180    0.03402    0.05  0.95776
## Year2009           -0.05356    0.03425   -1.56  0.11791
## Year2010           -0.06550    0.03405   -1.92  0.05443 .
## Year2011           -0.04368    0.03363   -1.30  0.19401
## Year2012           -0.04666    0.03418   -1.37  0.17223
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.624
## Multiple R-squared:  0.0389, Adjusted R-squared:  0.0378
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
##  3 observations c(6158,6159,6186)
##   are outliers with |weight| = 0 ( < 6.9e-06);
##  1120 weights are ~= 1. The remaining 13347 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.017   0.875   0.949   0.911   0.985   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          6.91e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##                   GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.002  1           1.001
## Year             1.002 16           1.000
```




```
## [1] "List of  8 outliers with residuals above  2.5"
##          ScopusId  NLCS Year OneField Fields residuals
## 8990  12344273284 3.769 2004     2742      1     2.516
## 9860  29944433045 3.989 2005     2742      1     2.513
## 10025 29244478299 4.133 2005     2742      1     2.657
## 10165 25144493160 3.924 2005     2742      1     2.575
## 10270 23644445382 4.848 2005     2742      1     3.381
## 10271 23644457022 4.605 2005     2742      1     3.129
## 10309 23944445391 4.398 2005     2742      1     3.040
## 10734 16244371358 3.961 2005     2742      1     2.612
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.4699 -0.4311  0.0363  0.4228  3.3781
##
## Coefficients:
##                   Estimate Std. Error t value Pr(>|t|)
## (Intercept)        1.11793    0.02934   38.10  < 2e-16 ***
## LastAuthorFemale1 -0.11540    0.01081  -10.67  < 2e-16 ***
## Year1997          -0.04415    0.03953   -1.12   0.2641
## Year1998           0.01150    0.03968    0.29   0.7720
## Year1999          -0.00212    0.03926   -0.05   0.9569
```

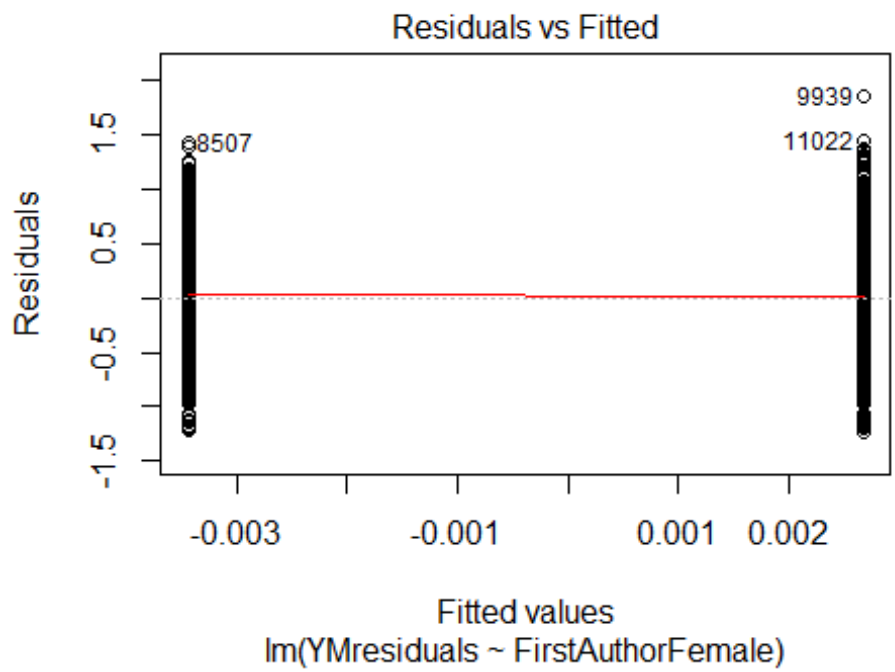
```
## Year2000           0.04254    0.03689    1.15   0.2488
## Year2001           0.08406    0.03782    2.22   0.0262 *
## Year2002           0.14134    0.03629    3.89  9.9e-05 ***
## Year2003           0.11029    0.03811    2.89   0.0038 **
## Year2004           0.25643    0.04157    6.17  7.1e-10 ***
## Year2005           0.35200    0.03981    8.84  < 2e-16 ***
## Year2006           0.30256    0.03819    7.92  2.5e-15 ***
## Year2007           0.01247    0.03435    0.36   0.7166
## Year2008           0.00512    0.03390    0.15   0.8800
## Year2009          -0.04936    0.03405   -1.45   0.1472
## Year2010          -0.05985    0.03386   -1.77   0.0772 .
## Year2011          -0.03947    0.03341   -1.18   0.2374
## Year2012          -0.04448    0.03403   -1.31   0.1912
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.621
## Multiple R-squared:  0.0464, Adjusted R-squared:  0.0453
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
##  3 observations c(6158,6159,6186)
##   are outliers with |weight| = 0 ( < 6.9e-06);
##  1136 weights are ~= 1. The remaining 13331 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.026   0.873   0.950   0.911   0.984   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          6.91e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Sample size for the above analysis:  14470"
## [1] ""
## [1] ""
## [1] "####################################"
## [1] "Analysis of AJSC 2743"
## [1] "####################################"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
```

```
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  636  606  459  476  485  533  430  435  461  514  496  598  613  805  824
## 2011 2012
##  842  669
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  285  387  289  321  239  185  276  277  311  364  336  420  449  578  568
## 2011 2012
##  593  494
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  248  350  250  284  213  174  249  234  278  303  297  362  412  516  495
## 2011 2012
##  532  438
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
##  Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 82, df = 16, p-value = 6e-11
```



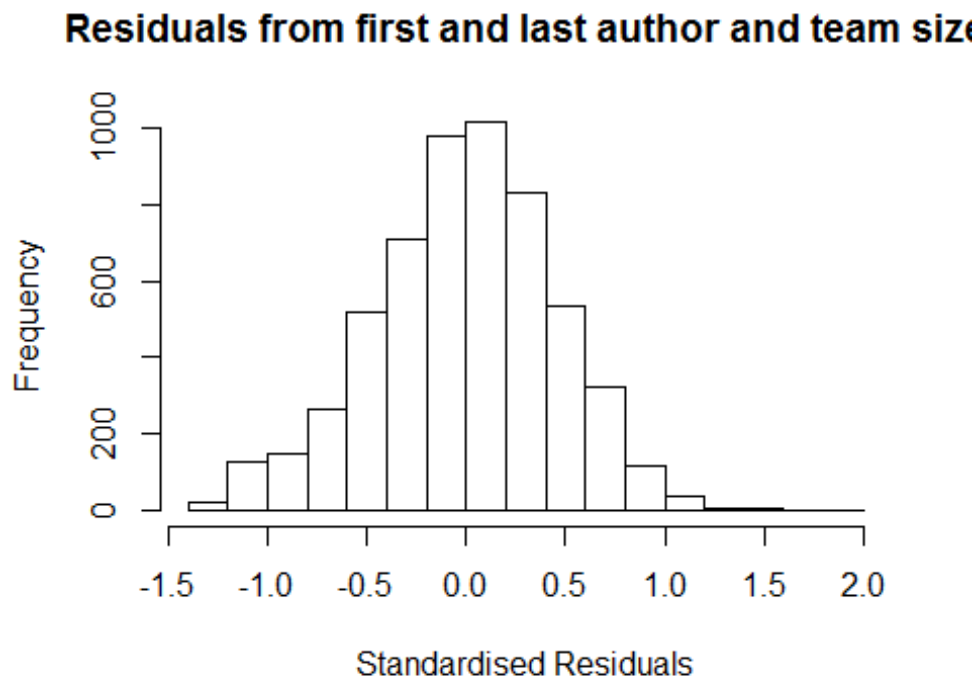

```
##
##  Bartlett test of homogeneity of variances
##
```

```
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 2.4, df = 1, p-value = 0.1
```



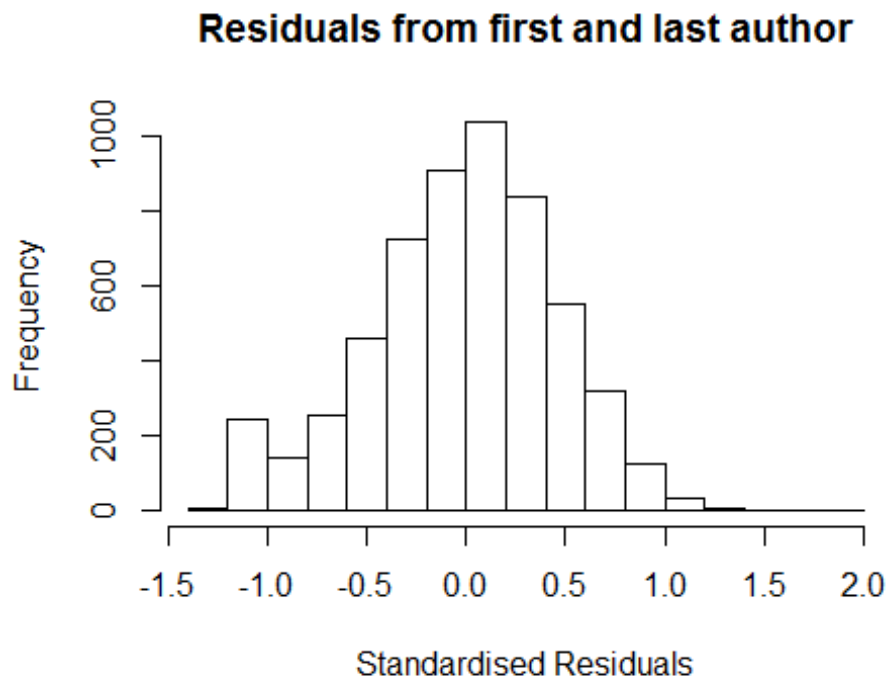

```
## [1] "Female first author team size 2018 geometric mean: 4.98198562647724"
## [1] "Male first author team size 2018 geometric mean: 4.24821362082528"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 25000, p-value = 0.04
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 4.56647327957624"
## [1] "Male last author team size 2018 geometric mean: 4.86560102244315"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 23000, p-value = 0.07
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.106  1           1.052
## LastAuthorFemale  1.088  1           1.043
```

```
## UniqueAuthors     1.109  4           1.013
## Year              1.179 16           1.005
```



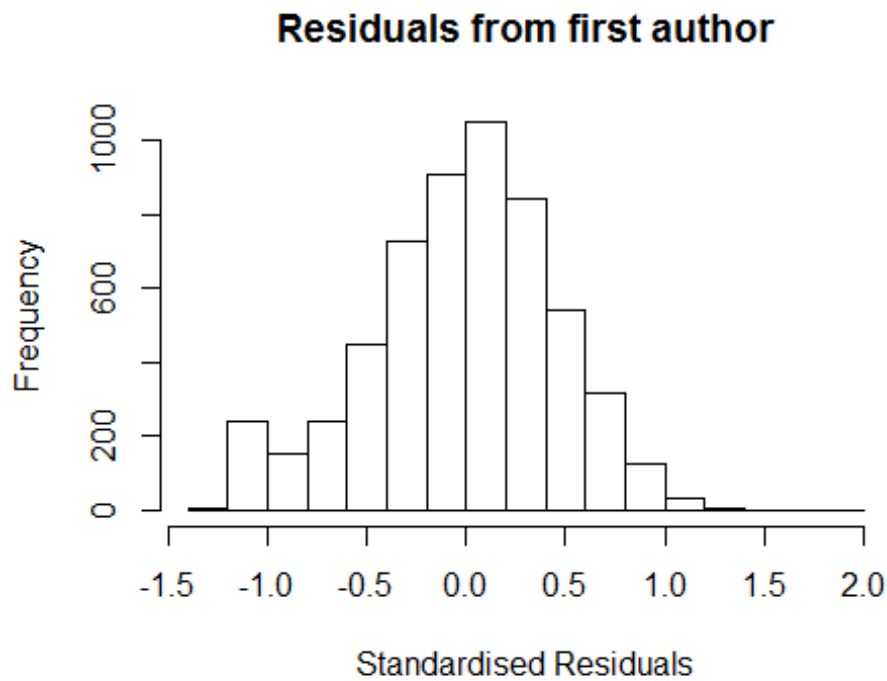

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
UniqueAuthors +
##     Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##     k.max = 1000))
##  \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.28488 -0.30098  0.00948  0.29287  1.95041
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         0.82861    0.03519   23.55  < 2e-16 ***
## FirstAuthorFemale1 -0.03261    0.01303   -2.50  0.01235 *
## LastAuthorFemale1  -0.01463    0.01384   -1.06  0.29028
## UniqueAuthors2      0.21090    0.03200    6.59  4.8e-11 ***
## UniqueAuthors3      0.27496    0.02914    9.44  < 2e-16 ***
## UniqueAuthors4      0.28128    0.02948    9.54  < 2e-16 ***
## UniqueAuthors5      0.41661    0.02709   15.38  < 2e-16 ***
```

```
## Year1997            0.12980    0.03562    3.64  0.00027 ***
## Year1998            0.05868    0.03805    1.54  0.12312
## Year1999            0.06301    0.03753    1.68  0.09322 .
## Year2000            0.11708    0.03983    2.94  0.00330 **
## Year2001            0.08135    0.04429    1.84  0.06631 .
## Year2002           -0.01438    0.03792   -0.38  0.70457
## Year2003            0.05727    0.03722    1.54  0.12397
## Year2004           -0.02037    0.03582   -0.57  0.56959
## Year2005            0.02104    0.03615    0.58  0.56050
## Year2006            0.04213    0.03848    1.09  0.27369
## Year2007           -0.00412    0.03889   -0.11  0.91557
## Year2008            0.03120    0.03650    0.85  0.39269
## Year2009           -0.03736    0.03561   -1.05  0.29416
## Year2010            0.00852    0.03496    0.24  0.80734
## Year2011           -0.02428    0.03474   -0.70  0.48466
## Year2012            0.03966    0.03681    1.08  0.28136
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.441
## Multiple R-squared:  0.0797, Adjusted R-squared:  0.0761
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
##  509 weights are ~= 1. The remaining 5126 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.012   0.865   0.949   0.902   0.985   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.77e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.083  1           1.041
## LastAuthorFemale  1.068  1           1.033
## Year              1.084 16           1.003
```



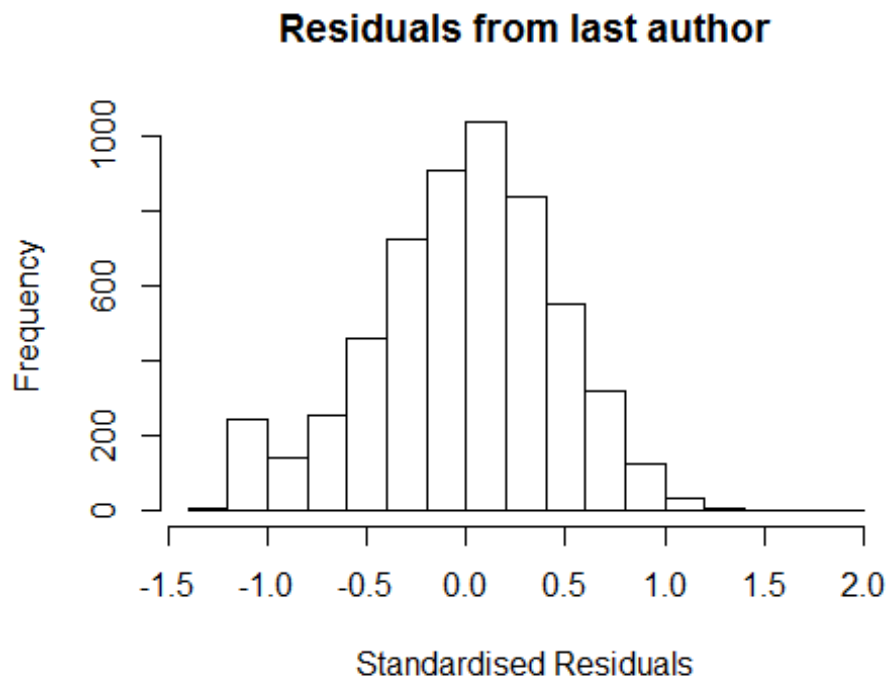

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.2408 -0.3072  0.0154  0.3075  1.8494
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)          1.09240    0.02786   39.22  < 2e-16 ***
## FirstAuthorFemale1 -0.00758    0.01328   -0.57  0.56844
## LastAuthorFemale1  -0.02846    0.01423   -2.00  0.04555 *
## Year1997            0.13672    0.03642    3.75  0.00018 ***
## Year1998            0.07805    0.03898    2.00  0.04529 *
## Year1999            0.09117    0.03767    2.42  0.01554 *
## Year2000            0.14839    0.04097    3.62  0.00029 ***
## Year2001            0.11615    0.04488    2.59  0.00967 **
## Year2002            0.00577    0.03817    0.15  0.87989
## Year2003            0.08092    0.03882    2.08  0.03715 *
## Year2004            0.01219    0.03778    0.32  0.74693
## Year2005            0.07282    0.03754    1.94  0.05247 .
```

```
## Year2006            0.06109    0.03971    1.54  0.12399
## Year2007            0.01512    0.04031    0.38  0.70766
## Year2008            0.06104    0.03740    1.63  0.10271
## Year2009           -0.00425    0.03663   -0.12  0.90774
## Year2010            0.06007    0.03568    1.68  0.09228 .
## Year2011            0.03763    0.03576    1.05  0.29273
## Year2012            0.08524    0.03761    2.27  0.02348 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.45
## Multiple R-squared:  0.00965,    Adjusted R-squared:  0.00648
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  498 weights are ~= 1. The remaining 5137 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0525  0.8640  0.9480  0.8990  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.77e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.053  1           1.026
## Year              1.053 16           1.002
```




```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##    Min     1Q Median     3Q    Max
## -1.233 -0.307  0.018  0.310  1.828
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)          1.08848    0.02791   39.00  < 2e-16 ***
## FirstAuthorFemale1 -0.01320    0.01311   -1.01  0.31418
## Year1997             0.13569    0.03649    3.72  0.00020 ***
## Year1998             0.07742    0.03910    1.98  0.04774 *
## Year1999             0.09056    0.03771    2.40  0.01635 *
## Year2000             0.14463    0.04099    3.53  0.00042 ***
## Year2001             0.11596    0.04492    2.58  0.00985 **
## Year2002             0.00506    0.03823    0.13  0.89467
## Year2003             0.08078    0.03889    2.08  0.03783 *
## Year2004             0.01160    0.03782    0.31  0.75911
## Year2005             0.07059    0.03758    1.88  0.06039 .
## Year2006             0.05775    0.03964    1.46  0.14519
```

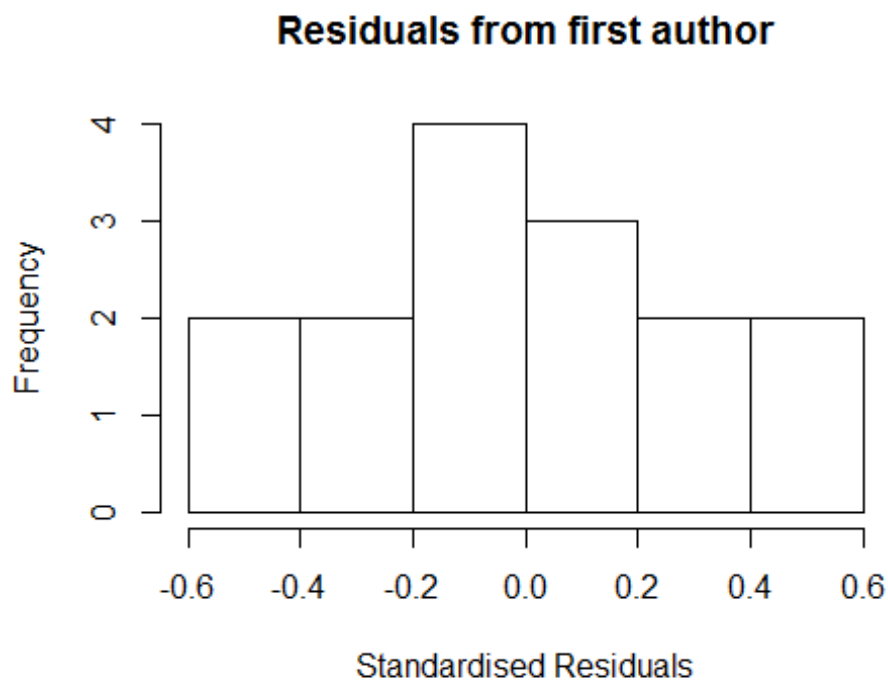
```
## Year2007            0.01149    0.04020    0.29  0.77504
## Year2008            0.05724    0.03732    1.53  0.12511
## Year2009           -0.00761    0.03661   -0.21  0.83541
## Year2010            0.05681    0.03564    1.59  0.11099
## Year2011            0.03414    0.03582    0.95  0.34064
## Year2012            0.08050    0.03747    2.15  0.03174 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.45
## Multiple R-squared:  0.00889,    Adjusted R-squared:  0.00589
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  488 weights are ~= 1. The remaining 5147 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0609  0.8640  0.9480  0.8990  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.77e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##                   GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.038  1           1.019
## Year             1.038 16           1.001
```



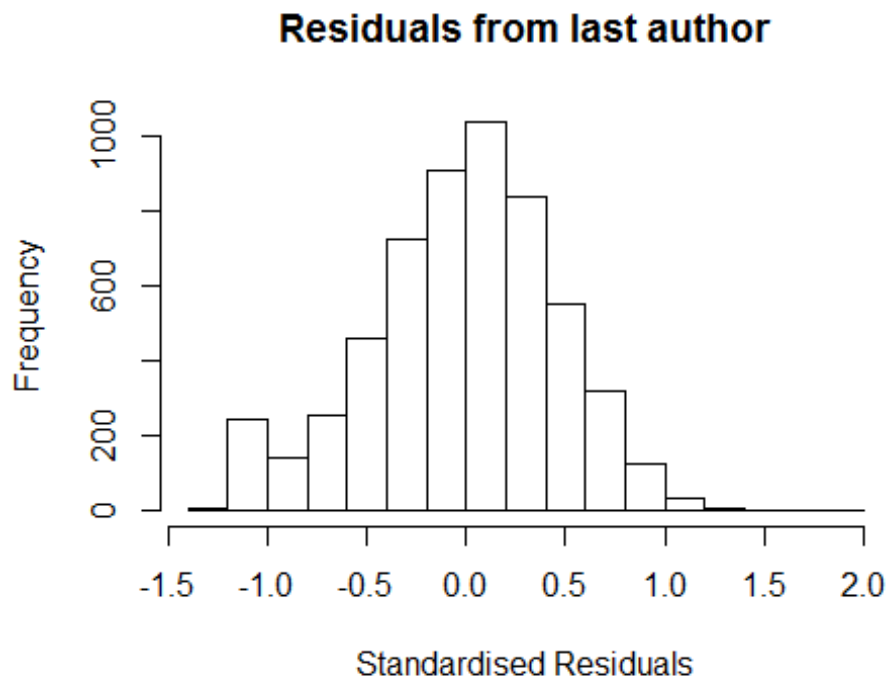

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.2383 -0.3069  0.0172  0.3082  1.8549
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.09051    0.02756   39.57  < 2e-16 ***
## LastAuthorFemale1 -0.03017    0.01404   -2.15  0.03176 *
## Year1997            0.13631    0.03640    3.75  0.00018 ***
## Year1998            0.07770    0.03894    2.00  0.04606 *
## Year1999            0.09089    0.03762    2.42  0.01572 *
## Year2000            0.14783    0.04096    3.61  0.00031 ***
## Year2001            0.11610    0.04488    2.59  0.00970 **
## Year2002            0.00584    0.03816    0.15  0.87843
## Year2003            0.08042    0.03878    2.07  0.03818 *
## Year2004            0.01152    0.03775    0.31  0.76019
## Year2005            0.07209    0.03751    1.92  0.05467 .
## Year2006            0.06039    0.03969    1.52  0.12820
```

```
## Year2007           0.01368    0.04027    0.34  0.73414
## Year2008           0.05943    0.03721    1.60  0.11034
## Year2009          -0.00544    0.03659   -0.15  0.88191
## Year2010           0.05851    0.03554    1.65  0.09981 .
## Year2011           0.03575    0.03551    1.01  0.31409
## Year2012           0.08360    0.03749    2.23  0.02580 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.45
## Multiple R-squared:  0.00958,    Adjusted R-squared:  0.00659
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  505 weights are ~= 1. The remaining 5130 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0503  0.8630  0.9480  0.8990  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.77e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Sample size for the above analysis:  5635"
## [1] ""
## [1] ""
## [1] "####################################"
## [1] "Analysis of AJSC 2744"
## [1] "####################################"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1997 2004 2005 2006 2007 2008
##    3    2    2    3    5    3
##
## 1997 2004 2005 2006 2007 2008
##    1    2    2    3    5    3
##
## 1997 2004 2005 2006 2007 2008
##    1    2    2    3    5    2
```

```
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
## [1] "Female first author team size 2018 geometric mean: NaN"
## [1] "Male first author team size 2018 geometric mean: NaN"
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
## [1] "Regression 3: First author gender, Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 5.021  1           2.241
## Year              5.021  5           1.175

## [1] "Regression 4: Last author gender, Year as factors"

## Warning in lf.cov(init, x = x): .vcov.avar1: negative diag(<vcov>) fixed
## up; consider 'cov=".vcov.w."' instead
```

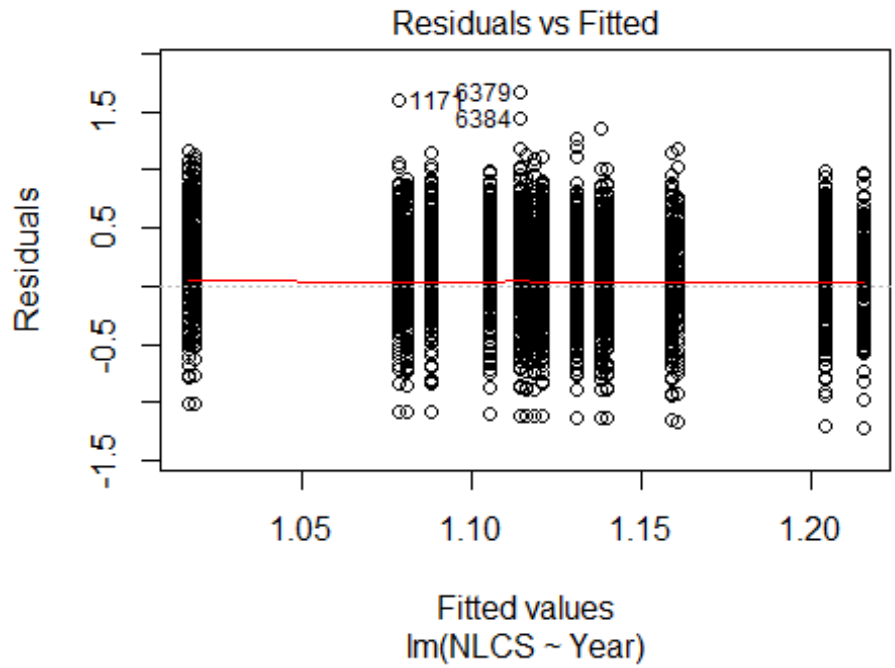



```
##                   GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 3.665  1           1.914
## Year             3.665  5           1.139
```

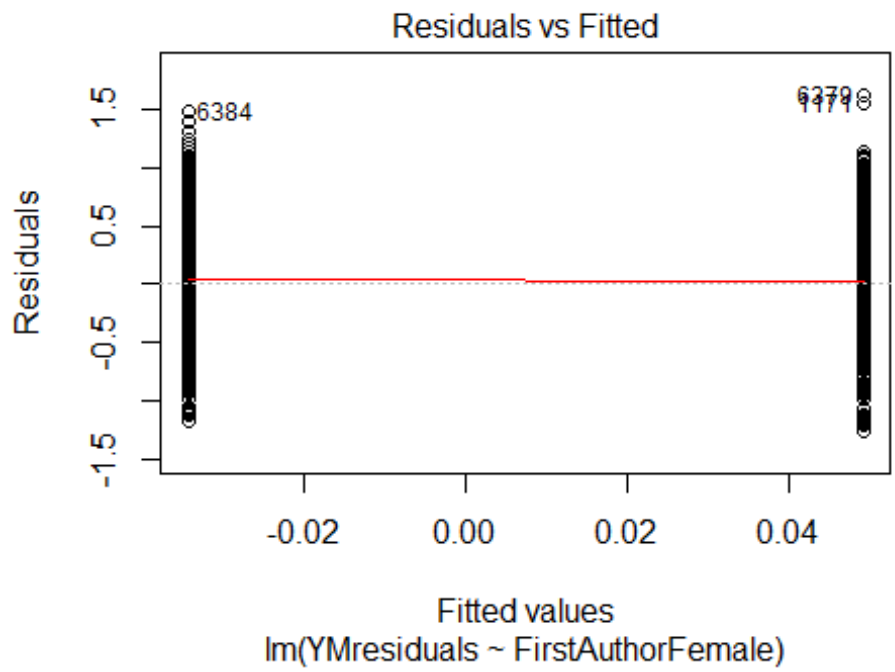



```
## [1] "Sample size for the above analysis:  15"
## [1] ""
## [1] ""
## [1] "####################################"
## [1] "Analysis of AJSC 2745"
## [1] "####################################"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  443  411  366  362  340  425  410  364  351  392  419  472  518  532  541
## 2011 2012
##  579  564
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  266  248  222  214  179  171  254  247  247  286  264  322  347  338  344
## 2011 2012
##  396  386
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  243  231  196  182  159  156  232  214  215  236  227  279  302  293  300
## 2011 2012
##  345  334
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
##  Bartlett test of homogeneity of variances
```

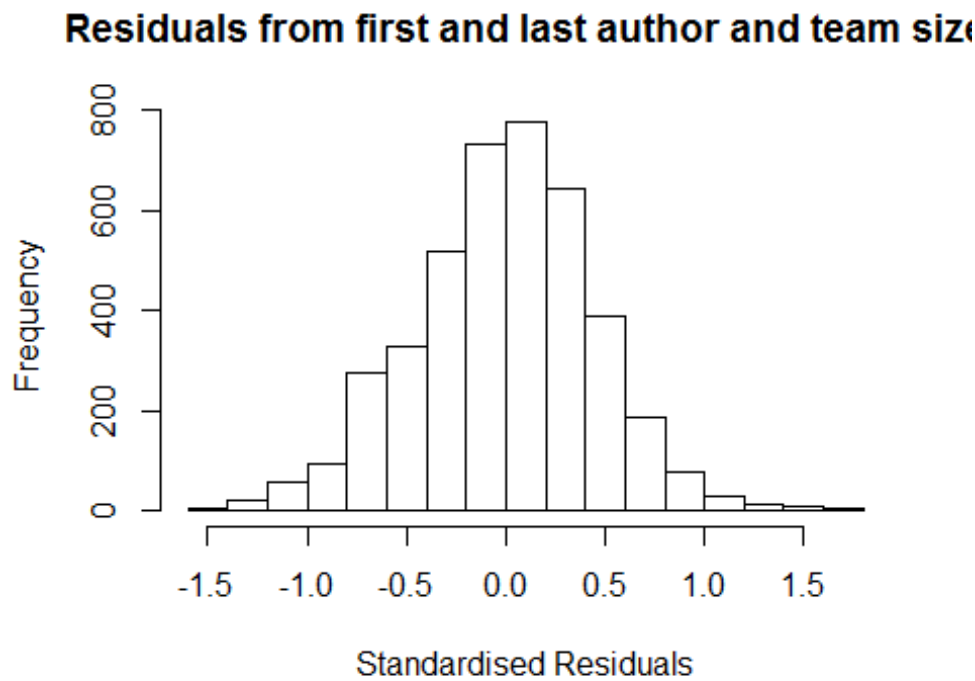
```
##
## data:  NLCS by Year
## Bartlett's K-squared = 83, df = 16, p-value = 4e-11
```




```
##
##  Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 28, df = 1, p-value = 1e-07
```

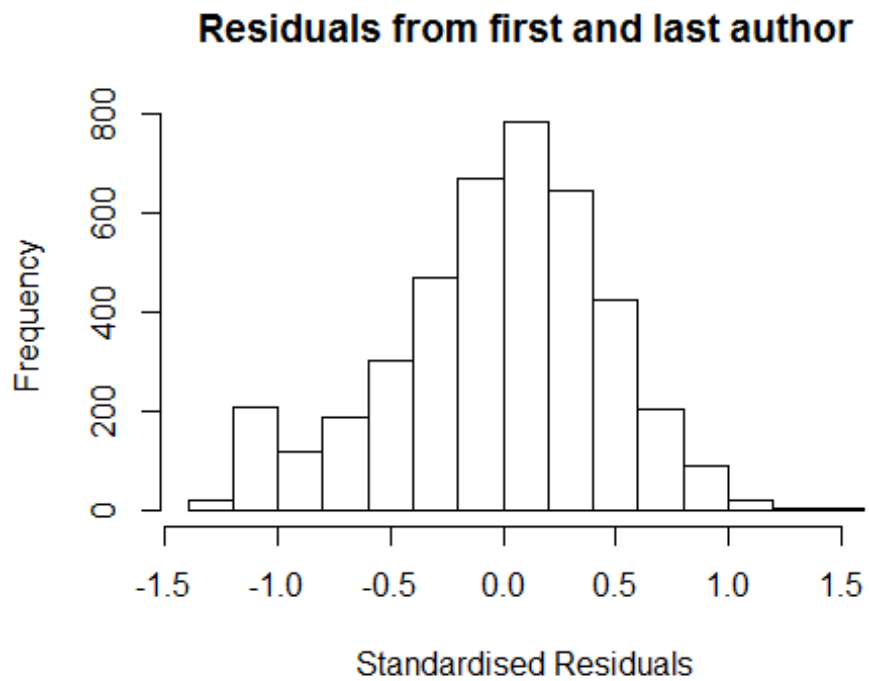



```
## [1] "Female first author team size 2018 geometric mean: 5.36649966962671"
## [1] "Male first author team size 2018 geometric mean: 5.24952895336164"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 11000, p-value = 0.6
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 5.20576738758418"
## [1] "Male last author team size 2018 geometric mean: 5.40165726352305"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 10000, p-value = 0.6
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##                     GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.057  1           1.028
## LastAuthorFemale  1.042  1           1.021
## UniqueAuthors     1.119  4           1.014
## Year              1.165 16           1.005
```



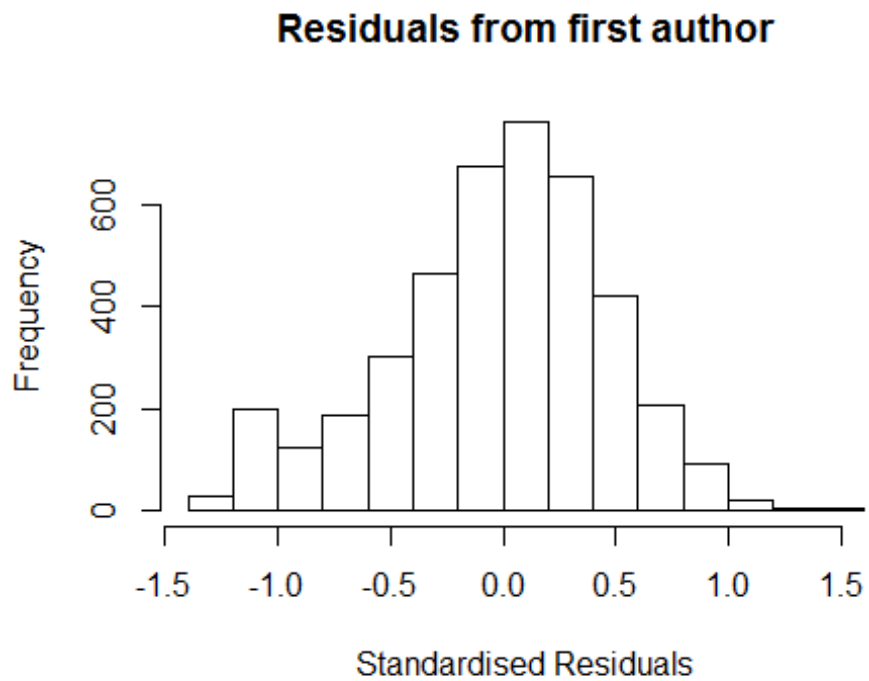

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
UniqueAuthors +
##     Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##     k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.4096 -0.2888  0.0106  0.2916  1.6719
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         0.64424    0.04671   13.79  < 2e-16 ***
## FirstAuthorFemale1  0.03028    0.01418    2.14   0.0328 *
## LastAuthorFemale1   0.02495    0.01517    1.64   0.1002
## UniqueAuthors2      0.27351    0.03895    7.02  2.5e-12 ***
## UniqueAuthors3      0.33850    0.03735    9.06  < 2e-16 ***
## UniqueAuthors4      0.47388    0.03655   12.96  < 2e-16 ***
## UniqueAuthors5      0.59590    0.03336   17.86  < 2e-16 ***
## Year1997           -0.01043    0.05260   -0.20   0.8428
## Year1998            0.02393    0.05270    0.45   0.6498
## Year1999            0.05160    0.05298    0.97   0.3301
```

```
## Year2000            0.16949    0.05213    3.25   0.0012 **
## Year2001            0.12352    0.05209    2.37   0.0178 *
## Year2002            0.05744    0.04603    1.25   0.2122
## Year2003            0.00930    0.04780    0.19   0.8457
## Year2004            0.07424    0.04608    1.61   0.1072
## Year2005            0.02472    0.04546    0.54   0.5867
## Year2006            0.05442    0.04660    1.17   0.2430
## Year2007            0.01782    0.04497    0.40   0.6919
## Year2008            0.01049    0.04602    0.23   0.8198
## Year2009            0.02467    0.04474    0.55   0.5814
## Year2010           -0.04390    0.04491   -0.98   0.3284
## Year2011            0.09393    0.04363    2.15   0.0314 *
## Year2012            0.00443    0.04614    0.10   0.9235
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.423
## Multiple R-squared:  0.178,  Adjusted R-squared:  0.174
## Convergence in 15 IRWLS iterations
##
## Robustness weights:
##  345 weights are ~= 1. The remaining 3799 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.083   0.860   0.949   0.899   0.985   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          2.41e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.039  1           1.019
## LastAuthorFemale  1.041  1           1.020
## Year              1.064 16           1.002
```



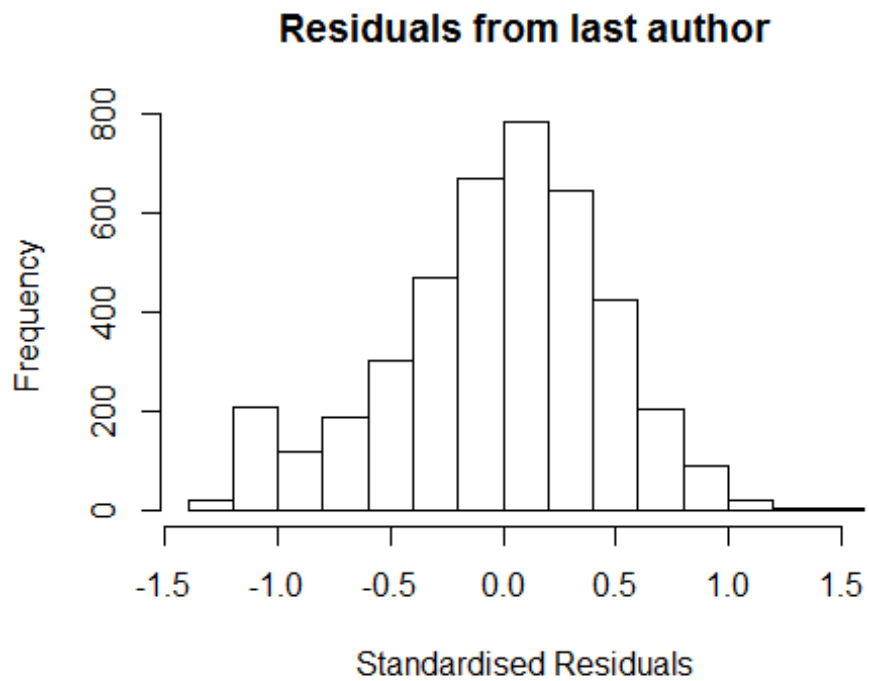

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.2557 -0.2954  0.0246  0.3028  1.5954
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)        1.001445   0.042821   23.39  < 2e-16 ***
## FirstAuthorFemale1 0.075370   0.015221    4.95  7.7e-07 ***
## LastAuthorFemale1  0.037482   0.016248    2.31  0.02111 *
## Year1997           0.000423   0.061453    0.01  0.99451
## Year1998           0.077832   0.058095    1.34  0.18041
## Year1999           0.086421   0.057565    1.50  0.13336
## Year2000           0.198520   0.055359    3.59  0.00034 ***
## Year2001           0.141401   0.056415    2.51  0.01223 *
## Year2002           0.119432   0.051393    2.32  0.02018 *
## Year2003           0.054860   0.051921    1.06  0.29076
## Year2004           0.128409   0.050442    2.55  0.01094 *
## Year2005           0.078982   0.050891    1.55  0.12074
```

```
## Year2006           0.112011   0.051499    2.18  0.02968 *
## Year2007           0.099229   0.050244    1.97  0.04834 *
## Year2008           0.109824   0.050961    2.16  0.03121 *
## Year2009           0.116102   0.050586    2.30  0.02178 *
## Year2010           0.070068   0.050563    1.39  0.16590
## Year2011           0.183176   0.047755    3.84  0.00013 ***
## Year2012           0.097099   0.050459    1.92  0.05438 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.441
## Multiple R-squared:  0.0209, Adjusted R-squared:  0.0166
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
##  353 weights are ~= 1. The remaining 3791 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.163   0.856   0.948   0.892   0.985   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          2.41e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.034  1           1.017
## Year              1.034 16           1.001
```




```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.2320 -0.2976  0.0271  0.3030  1.5810
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)          1.00898    0.04286   23.54  < 2e-16 ***
## FirstAuthorFemale1   0.08034    0.01522    5.28  1.4e-07 ***
## Year1997            -0.00053    0.06158   -0.01  0.99313
## Year1998             0.07876    0.05831    1.35  0.17685
## Year1999             0.08651    0.05799    1.49  0.13582
## Year2000             0.19897    0.05552    3.58  0.00034 ***
## Year2001             0.14268    0.05654    2.52  0.01166 *
## Year2002             0.11896    0.05155    2.31  0.02107 *
## Year2003             0.05432    0.05208    1.04  0.29694
## Year2004             0.12868    0.05062    2.54  0.01106 *
## Year2005             0.07995    0.05104    1.57  0.11734
## Year2006             0.11160    0.05160    2.16  0.03061 *
```

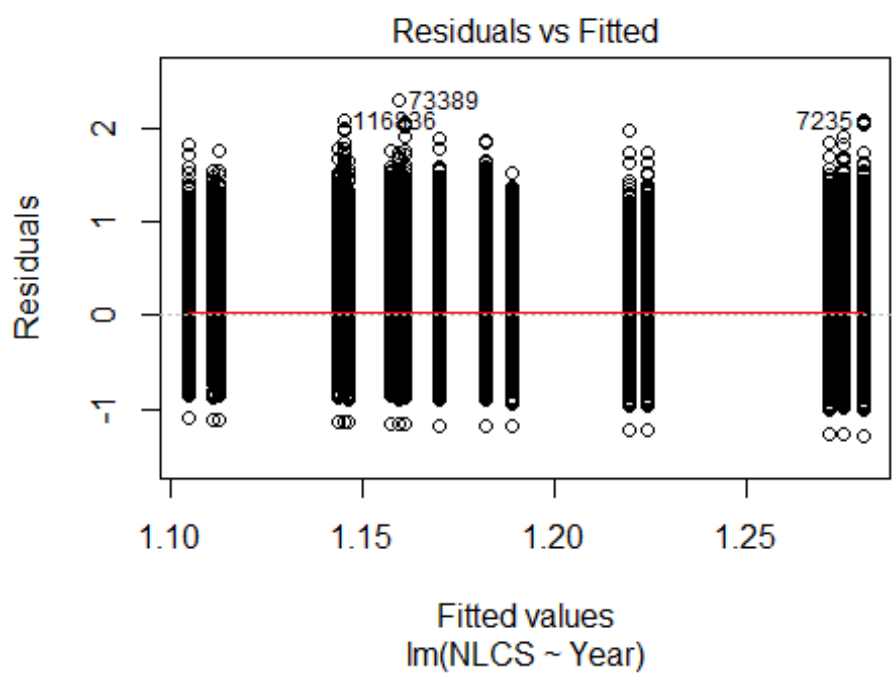
```
## Year2007            0.10135    0.05047    2.01  0.04468 *
## Year2008            0.11167    0.05118    2.18  0.02916 *
## Year2009            0.11834    0.05080    2.33  0.01989 *
## Year2010            0.07429    0.05072    1.46  0.14311
## Year2011            0.18642    0.04791    3.89  0.00010 ***
## Year2012            0.09914    0.05066    1.96  0.05042 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.441
## Multiple R-squared:  0.0196, Adjusted R-squared:  0.0155
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
##  346 weights are ~= 1. The remaining 3798 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.172   0.856   0.948   0.892   0.986   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          2.41e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##                   GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.035  1           1.017
## Year             1.035 16           1.001
```



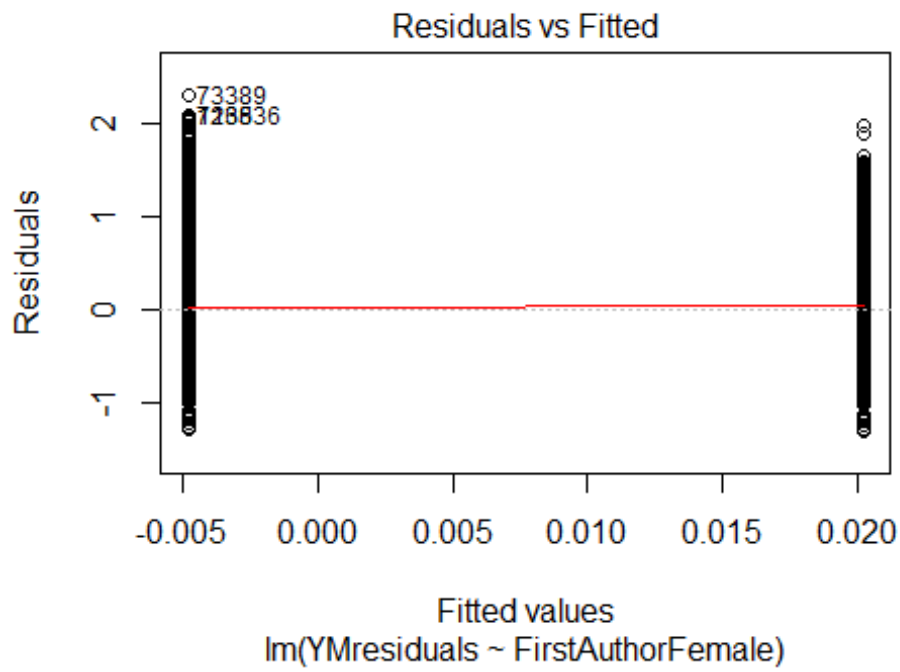

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.2719 -0.3030  0.0292  0.3022  1.6418
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.02515    0.04294   23.87  < 2e-16 ***
## LastAuthorFemale1  0.04925    0.01621    3.04  0.00240 **
## Year1997           0.00114    0.06241    0.02  0.98549
## Year1998           0.07734    0.05854    1.32  0.18651
## Year1999           0.09109    0.05788    1.57  0.11559
## Year2000           0.20360    0.05603    3.63  0.00028 ***
## Year2001           0.14401    0.05660    2.54  0.01099 *
## Year2002           0.12495    0.05179    2.41  0.01588 *
## Year2003           0.05604    0.05235    1.07  0.28447
## Year2004           0.13316    0.05099    2.61  0.00904 **
## Year2005           0.07917    0.05142    1.54  0.12370
## Year2006           0.11503    0.05185    2.22  0.02657 *
```

```
## Year2007           0.10638    0.05069    2.10  0.03592 *
## Year2008           0.11503    0.05131    2.24  0.02502 *
## Year2009           0.12436    0.05107    2.44  0.01493 *
## Year2010           0.07676    0.05081    1.51  0.13094
## Year2011           0.19747    0.04794    4.12  3.9e-05 ***
## Year2012           0.10859    0.05084    2.14  0.03274 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.442
## Multiple R-squared:  0.0149, Adjusted R-squared:  0.0108
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
##  352 weights are ~= 1. The remaining 3792 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.137   0.858   0.949   0.891   0.985   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          2.41e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Sample size for the above analysis:  4144"
## [1] ""
## [1] ""
## [1] "####################################"
## [1] "Analysis of AJSC 2746"
## [1] "####################################"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 5515 5428 5447 5528 5693 6192 5606 4481 4653 5195 6539 6987 7835 7987 8111
## 2011 2012
## 8333 8027
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 3214 3325 3157 3478 3617 3660 4410 3500 3677 4127 5017 5626 6290 6435 6566
## 2011 2012
```

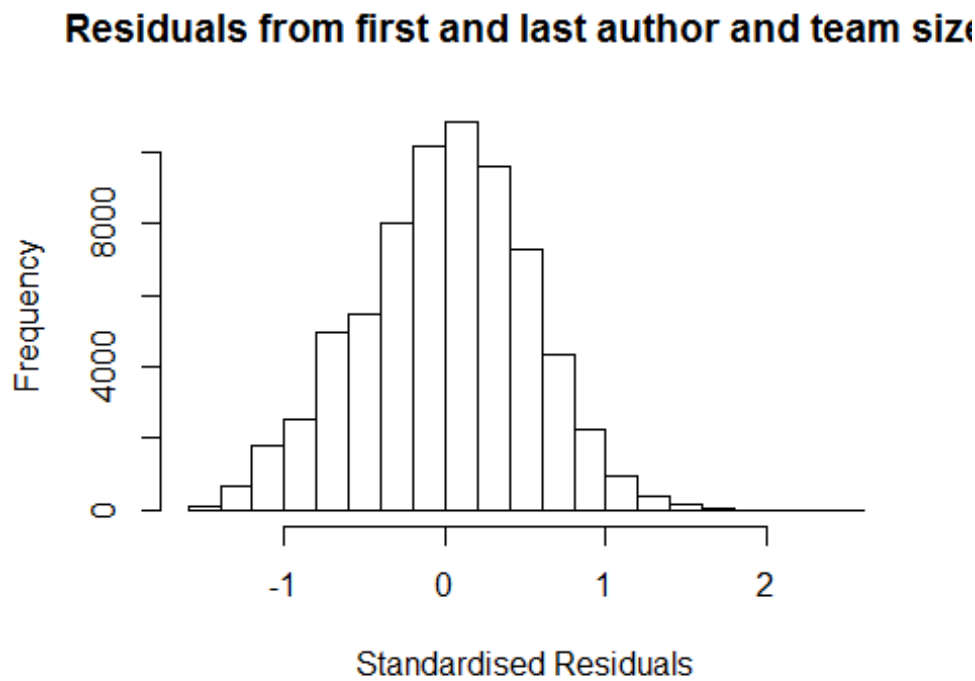
```
## 6727 6526
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 2829 2933 2789 3030 3174 3219 3865 3073 3202 3585 4405 4917 5515 5678 5765
## 2011 2012
## 5877 5699
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
##  Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 200, df = 16, p-value <2e-16
```




```
##
##  Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 7.9, df = 1, p-value = 0.005
```

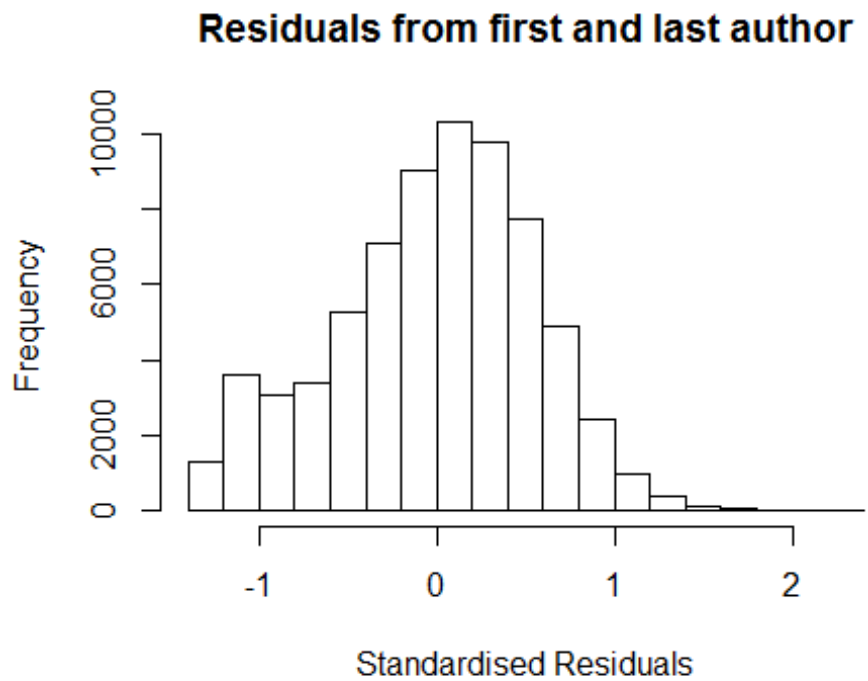



```
## [1] "Female first author team size 2018 geometric mean: 4.99125256211924"
## [1] "Male first author team size 2018 geometric mean: 4.84924878066027"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 5500000, p-value = 0.7
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 4.97330291867359"
## [1] "Male last author team size 2018 geometric mean: 4.87258859497463"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 4100000, p-value = 0.2
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##                     GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.031  1           1.016
## LastAuthorFemale  1.024  1           1.012
## UniqueAuthors     1.023  4           1.003
## Year              1.033 16           1.001
```



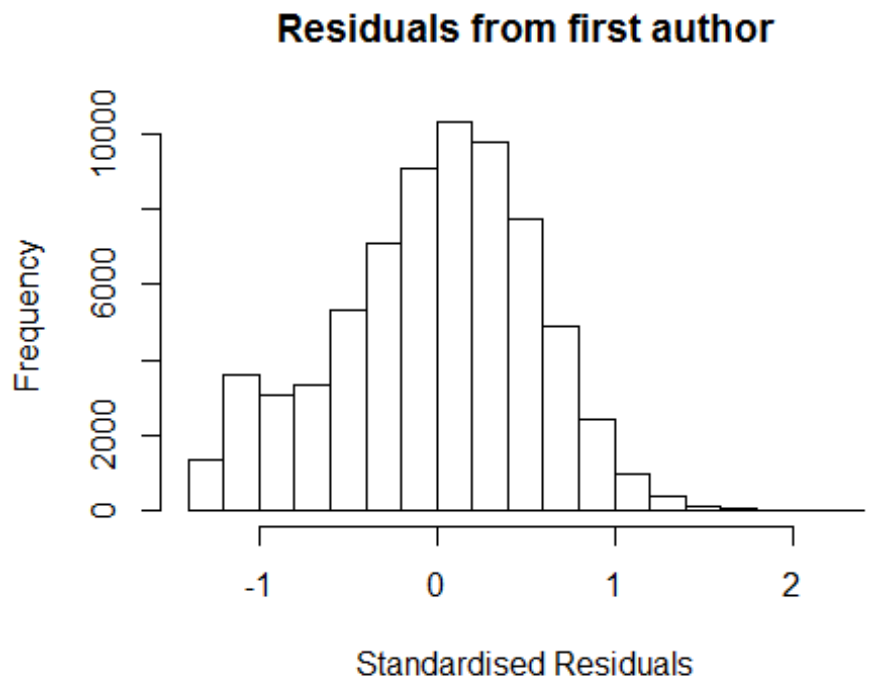

```
## [1] "List of  1 outliers with residuals above  2.5"
##           ScopusId  NLCS Year OneField Fields residuals
## 116836 84855893653 3.226 2012     2746      1     2.585
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
UniqueAuthors +
##     Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##     k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.5555 -0.3541  0.0193  0.3529  2.5853
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         0.85514    0.01367   62.57  < 2e-16 ***
## FirstAuthorFemale1  0.00394    0.00520    0.76  0.44881
## LastAuthorFemale1   0.01136    0.00617    1.84  0.06541 .
## UniqueAuthors2      0.27012    0.00990   27.28  < 2e-16 ***
## UniqueAuthors3      0.39154    0.00944   41.49  < 2e-16 ***
## UniqueAuthors4      0.49076    0.00935   52.52  < 2e-16 ***
## UniqueAuthors5      0.68151    0.00839   81.25  < 2e-16 ***
## Year1997            0.00351    0.01583    0.22  0.82471
## Year1998           -0.01195    0.01546   -0.77  0.43959
## Year1999           -0.05585    0.01514   -3.69  0.00022 ***
```

```
## Year2000           -0.07925    0.01471   -5.39  7.2e-08 ***
## Year2001           -0.09692    0.01468   -6.60  4.0e-11 ***
## Year2002           -0.15350    0.01426  -10.76  < 2e-16 ***
## Year2003           -0.19300    0.01469  -13.14  < 2e-16 ***
## Year2004           -0.19672    0.01464  -13.44  < 2e-16 ***
## Year2005           -0.20380    0.01427  -14.28  < 2e-16 ***
## Year2006           -0.18260    0.01388  -13.15  < 2e-16 ***
## Year2007           -0.17558    0.01359  -12.92  < 2e-16 ***
## Year2008           -0.16903    0.01344  -12.58  < 2e-16 ***
## Year2009           -0.16213    0.01345  -12.05  < 2e-16 ***
## Year2010           -0.18735    0.01349  -13.89  < 2e-16 ***
## Year2011           -0.19814    0.01347  -14.71  < 2e-16 ***
## Year2012           -0.21441    0.01375  -15.60  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.517
## Multiple R-squared:  0.157,  Adjusted R-squared:  0.157
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
##  2 observations c(3554,68879) are outliers with |weight| = 0 ( < 1.4e-06);
##  5874 weights are ~= 1. The remaining 63679 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0318  0.8630  0.9490  0.9050  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.44e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.014  1           1.007
## LastAuthorFemale  1.009  1           1.004
## Year              1.012 16           1.000
```



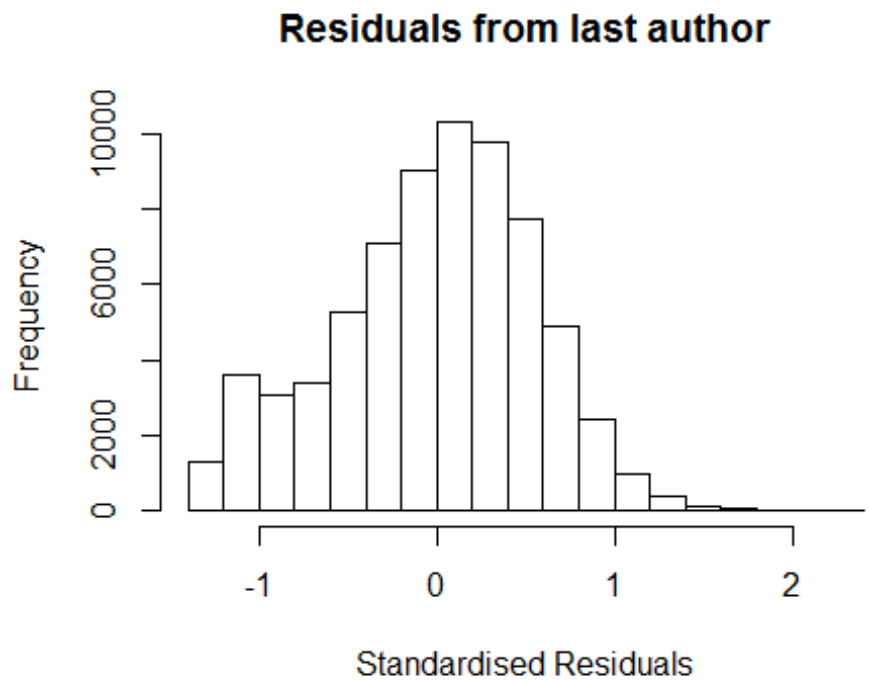

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.3225 -0.3775  0.0367  0.3825  2.2840
##
## Coefficients:
##                      Estimate Std. Error t value Pr(>|t|)
## (Intercept)          1.292996   0.012002  107.73  < 2e-16 ***
## FirstAuthorFemale1   0.029023   0.005595    5.19  2.1e-07 ***
## LastAuthorFemale1   -0.007350   0.006675   -1.10  0.27085
## Year1997             0.000523   0.016892    0.03  0.97528
## Year1998            -0.009611   0.016203   -0.59  0.55310
## Year1999            -0.054679   0.016100   -3.40  0.00068 ***
## Year2000            -0.060434   0.015713   -3.85  0.00012 ***
## Year2001            -0.090511   0.015687   -5.77  8.0e-09 ***
## Year2002            -0.138731   0.015259   -9.09  < 2e-16 ***
## Year2003            -0.181432   0.015692  -11.56  < 2e-16 ***
## Year2004            -0.183328   0.015619  -11.74  < 2e-16 ***
## Year2005            -0.177439   0.015160  -11.70  < 2e-16 ***
```

```
## Year2006           -0.144047   0.014816   -9.72  < 2e-16 ***
## Year2007           -0.128040   0.014465   -8.85  < 2e-16 ***
## Year2008           -0.118378   0.014296   -8.28  < 2e-16 ***
## Year2009           -0.105607   0.014290   -7.39  1.5e-13 ***
## Year2010           -0.124044   0.014344   -8.65  < 2e-16 ***
## Year2011           -0.129734   0.014315   -9.06  < 2e-16 ***
## Year2012           -0.140677   0.014587   -9.64  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.551
## Multiple R-squared:  0.00829,    Adjusted R-squared:  0.00804
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  5676 weights are ~= 1. The remaining 63879 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0468  0.8640  0.9480  0.9020  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.44e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.009  1           1.005
## Year              1.009 16           1.000
```




```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.3208 -0.3774  0.0365  0.3826  2.2849
##
## Coefficients:
##                      Estimate Std. Error t value Pr(>|t|)
## (Intercept)          1.292304   0.011989  107.79  < 2e-16 ***
## FirstAuthorFemale1   0.027999   0.005607    4.99  6.0e-07 ***
## Year1997             0.000513   0.016891    0.03  0.97578
## Year1998            -0.009656   0.016203   -0.60  0.55123
## Year1999            -0.054780   0.016100   -3.40  0.00067 ***
## Year2000            -0.060369   0.015713   -3.84  0.00012 ***
## Year2001            -0.090553   0.015687   -5.77  7.8e-09 ***
## Year2002            -0.138828   0.015258   -9.10  < 2e-16 ***
## Year2003            -0.181481   0.015692  -11.57  < 2e-16 ***
## Year2004            -0.183348   0.015619  -11.74  < 2e-16 ***
## Year2005            -0.177531   0.015159  -11.71  < 2e-16 ***
## Year2006            -0.144162   0.014815   -9.73  < 2e-16 ***
```

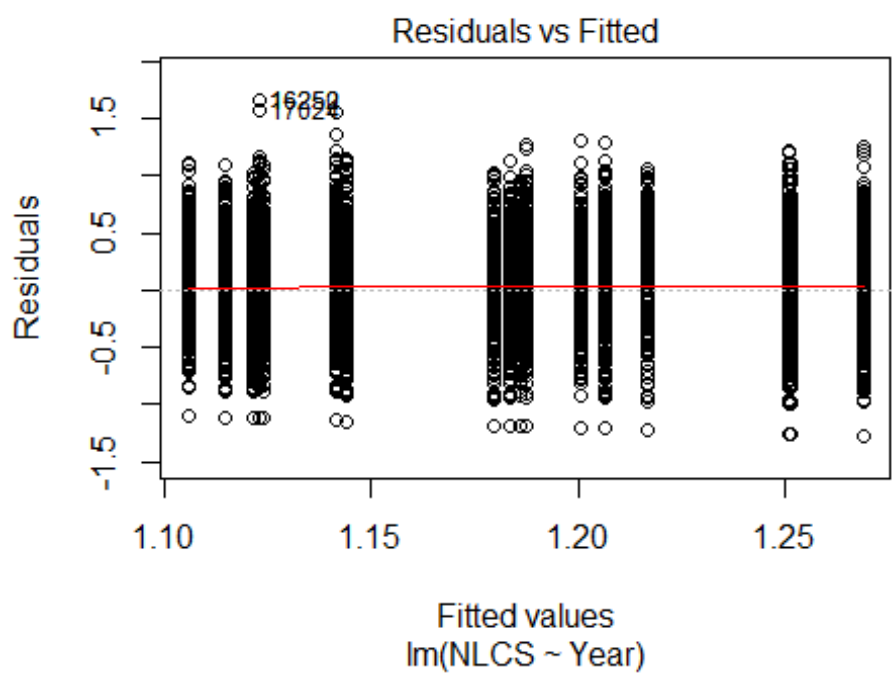
```
## Year2007           -0.128157   0.014464   -8.86  < 2e-16 ***
## Year2008           -0.118522   0.014296   -8.29  < 2e-16 ***
## Year2009           -0.105794   0.014288   -7.40  1.3e-13 ***
## Year2010           -0.124183   0.014343   -8.66  < 2e-16 ***
## Year2011           -0.129941   0.014313   -9.08  < 2e-16 ***
## Year2012           -0.140924   0.014583   -9.66  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.551
## Multiple R-squared:  0.00828,    Adjusted R-squared:  0.00803
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  5688 weights are ~= 1. The remaining 63867 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0466  0.8640  0.9480  0.9020  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.44e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##                   GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.003  1           1.002
## Year             1.003 16           1.000
```



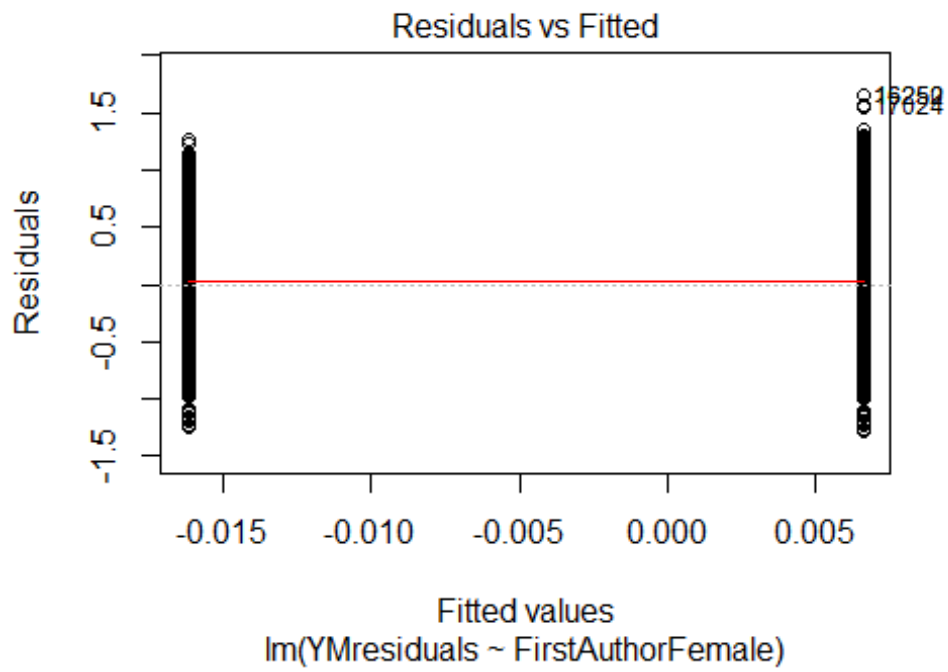

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.2963 -0.3771  0.0381  0.3824  2.2791
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.296078   0.011982  108.17  < 2e-16 ***
## LastAuthorFemale1 -0.001990   0.006659   -0.30  0.76511
## Year1997            0.000231   0.016895    0.01  0.98910
## Year1998           -0.009240   0.016202   -0.57  0.56847
## Year1999           -0.054123   0.016100   -3.36  0.00078 ***
## Year2000           -0.059452   0.015716   -3.78  0.00016 ***
## Year2001           -0.089601   0.015687   -5.71  1.1e-08 ***
## Year2002           -0.137729   0.015262   -9.02  < 2e-16 ***
## Year2003           -0.180187   0.015692  -11.48  < 2e-16 ***
## Year2004           -0.181579   0.015615  -11.63  < 2e-16 ***
## Year2005           -0.176146   0.015156  -11.62  < 2e-16 ***
## Year2006           -0.142334   0.014815   -9.61  < 2e-16 ***
```

```
## Year2007          -0.126149   0.014460   -8.72  < 2e-16 ***
## Year2008          -0.115808   0.014287   -8.11  5.3e-16 ***
## Year2009          -0.103069   0.014281   -7.22  5.4e-13 ***
## Year2010          -0.121112   0.014334   -8.45  < 2e-16 ***
## Year2011          -0.126734   0.014302   -8.86  < 2e-16 ***
## Year2012          -0.137163   0.014569   -9.41  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.551
## Multiple R-squared:  0.00791,    Adjusted R-squared:  0.00767
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  5549 weights are ~= 1. The remaining 64006 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0485  0.8640  0.9490  0.9020  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.44e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Sample size for the above analysis:  69555"
## [1] ""
## [1] ""
## [1] "####################################"
## [1] "Analysis of AJSC 2747"
## [1] "####################################"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  821  769  700  718  815  995  909  790  797  871 1004  959 1051 1038 1039
## 2011 2012
## 1116 1006
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  328  336  320  390  456  504  583  529  553  610  548  572  630  623  605
## 2011 2012
```

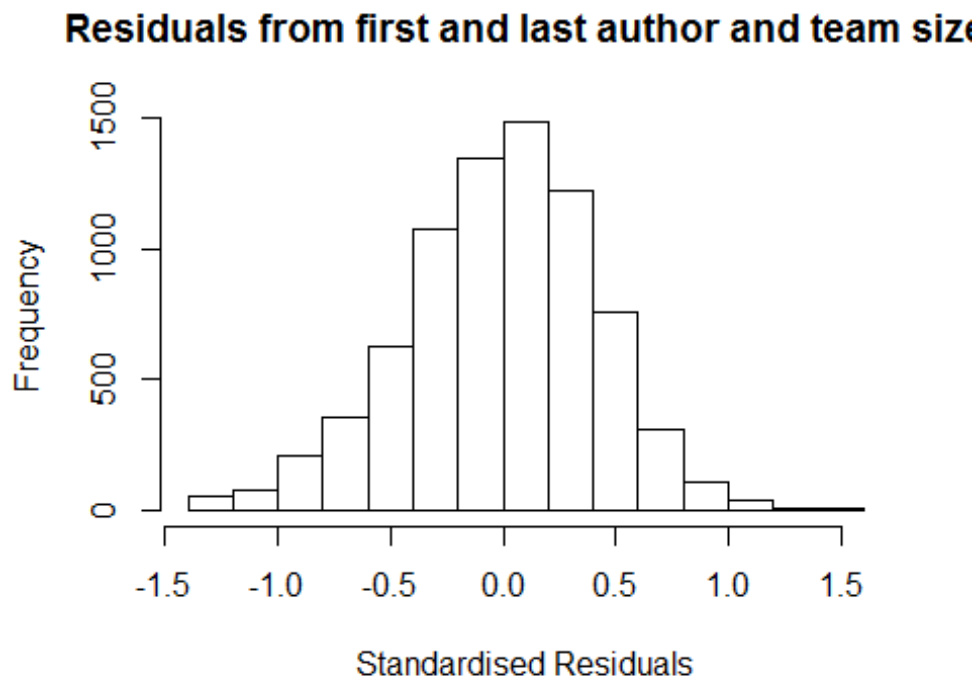
```
##  713  612
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  284  290  277  345  397  444  513  441  456  512  485  500  548  537  509
## 2011 2012
##  606  527
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
##  Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 87, df = 16, p-value = 7e-12
```




```
##
##  Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 4.1, df = 1, p-value = 0.04
```

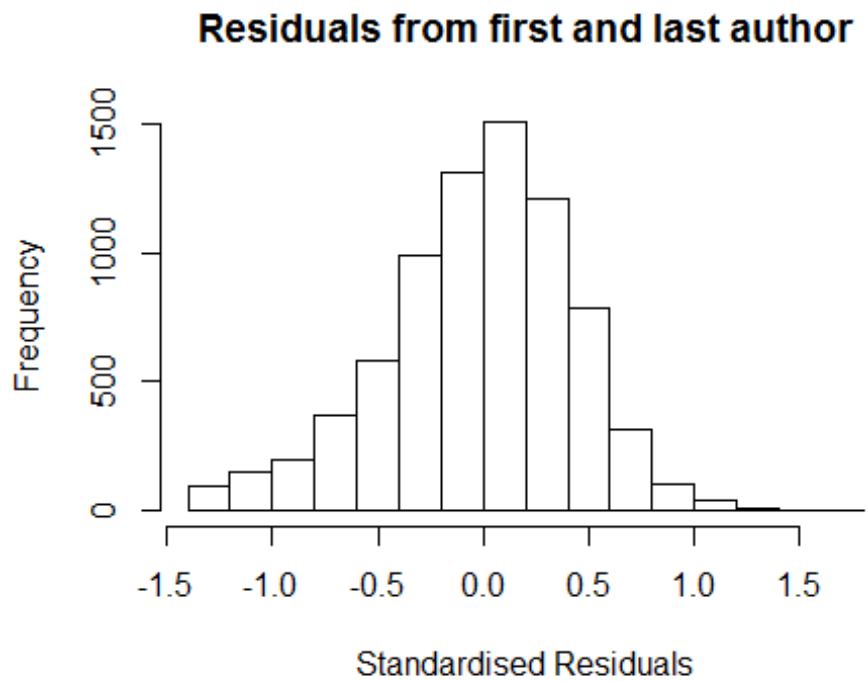



```
## [1] "Female first author team size 2018 geometric mean: 6.79905302846094"
## [1] "Male first author team size 2018 geometric mean: 6.65254430388983"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 47000, p-value = 0.6
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 6.58120416490644"
## [1] "Male last author team size 2018 geometric mean: 6.76639570082063"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 37000, p-value = 0.7
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.030  1           1.015
## LastAuthorFemale  1.044  1           1.022
## UniqueAuthors     1.059  4           1.007
## Year              1.087 16           1.003
```



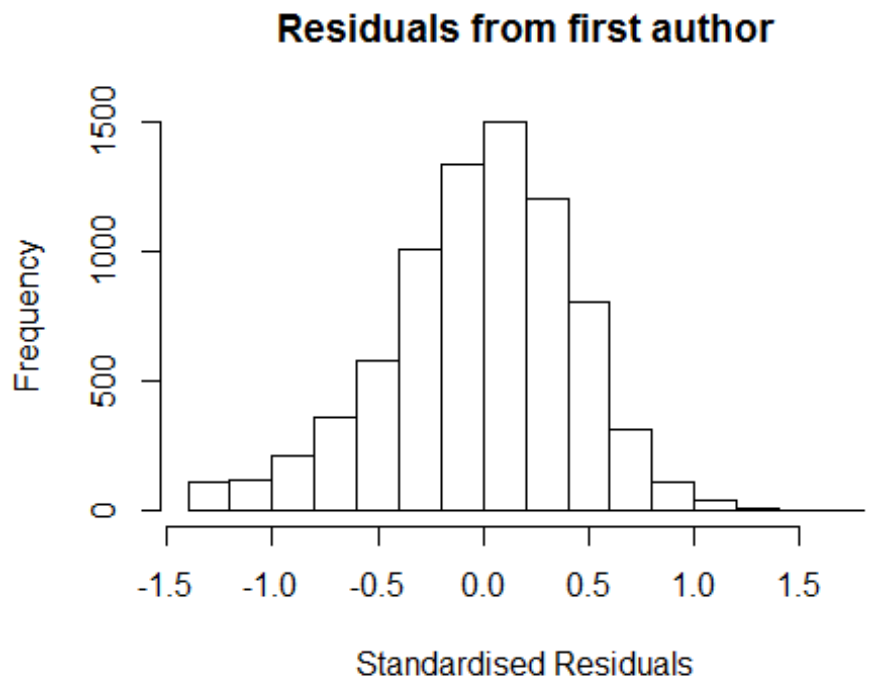

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
UniqueAuthors +
##     Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##     k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.3779 -0.2822  0.0127  0.2804  1.5809
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)         0.942789   0.041219   22.87  < 2e-16 ***
## FirstAuthorFemale1 -0.015830   0.010753   -1.47  0.14101
## LastAuthorFemale1  -0.040201   0.012882   -3.12  0.00181 **
## UniqueAuthors2      0.117023   0.035803    3.27  0.00109 **
## UniqueAuthors3      0.182405   0.033265    5.48  4.3e-08 ***
## UniqueAuthors4      0.301798   0.031182    9.68  < 2e-16 ***
## UniqueAuthors5      0.412508   0.028188   14.63  < 2e-16 ***
## Year1997           -0.066873   0.039337   -1.70  0.08917 .
## Year1998           -0.062128   0.040541   -1.53  0.12545
## Year1999           -0.075901   0.035725   -2.12  0.03365 *
```

```
## Year2000            0.013120   0.035246    0.37  0.70973
## Year2001           -0.038976   0.035471   -1.10  0.27189
## Year2002           -0.070529   0.035246   -2.00  0.04542 *
## Year2003           -0.120062   0.034555   -3.47  0.00051 ***
## Year2004           -0.150780   0.034230   -4.40  1.1e-05 ***
## Year2005           -0.141231   0.033895   -4.17  3.1e-05 ***
## Year2006           -0.133754   0.034834   -3.84  0.00012 ***
## Year2007           -0.046113   0.034512   -1.34  0.18154
## Year2008            0.000134   0.034239    0.00  0.99688
## Year2009            0.022555   0.034995    0.64  0.51926
## Year2010           -0.066439   0.034943   -1.90  0.05729 .
## Year2011           -0.109940   0.034620   -3.18  0.00150 **
## Year2012           -0.156150   0.034684   -4.50  6.8e-06 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.411
## Multiple R-squared:  0.104,  Adjusted R-squared:  0.102
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
##  682 weights are ~= 1. The remaining 6989 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.105   0.868   0.949   0.901   0.985   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.30e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.024  1           1.012
## LastAuthorFemale  1.032  1           1.016
## Year              1.035 16           1.001
```



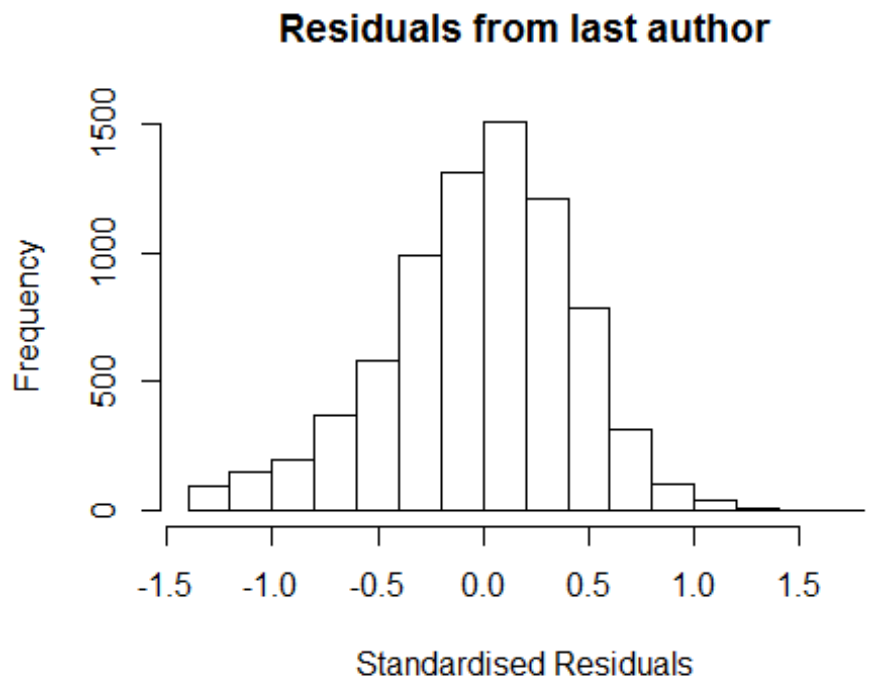

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.3153 -0.2870  0.0204  0.2866  1.6358
##
## Coefficients:
##                      Estimate Std. Error t value Pr(>|t|)
## (Intercept)          1.29e+00   3.03e-02   42.49  < 2e-16 ***
## FirstAuthorFemale1 -1.79e-02   1.11e-02   -1.60  0.10870
## LastAuthorFemale1  -6.48e-02   1.35e-02   -4.81  1.5e-06 ***
## Year1997           -7.06e-02   4.04e-02   -1.75  0.08062 .
## Year1998           -5.61e-02   4.16e-02   -1.35  0.17726
## Year1999           -7.60e-02   3.70e-02   -2.05  0.04007 *
## Year2000           -6.98e-05   3.67e-02    0.00  0.99848
## Year2001           -5.78e-02   3.69e-02   -1.57  0.11725
## Year2002           -9.34e-02   3.66e-02   -2.55  0.01082 *
## Year2003           -1.23e-01   3.57e-02   -3.45  0.00057 ***
## Year2004           -1.39e-01   3.54e-02   -3.92  9.0e-05 ***
## Year2005           -1.30e-01   3.53e-02   -3.69  0.00023 ***
```

```
## Year2006           -1.52e-01   3.63e-02   -4.19  2.8e-05 ***
## Year2007           -4.98e-02   3.60e-02   -1.39  0.16572
## Year2008           -9.16e-04   3.55e-02   -0.03  0.97942
## Year2009            2.62e-02   3.62e-02    0.73  0.46847
## Year2010           -5.15e-02   3.68e-02   -1.40  0.16184
## Year2011           -1.12e-01   3.64e-02   -3.09  0.00204 **
## Year2012           -1.45e-01   3.60e-02   -4.02  5.8e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.416
## Multiple R-squared:  0.0205, Adjusted R-squared:  0.0182
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
##  632 weights are ~= 1. The remaining 7039 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0879  0.8650  0.9490  0.8960  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.30e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.016  1           1.008
## Year              1.016 16           1.000
```




```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.3045 -0.2844  0.0188  0.2860  1.6486
##
## Coefficients:
##                      Estimate Std. Error t value Pr(>|t|)
## (Intercept)          1.282033   0.030180   42.48  < 2e-16 ***
## FirstAuthorFemale1 -0.025280   0.011127   -2.27  0.02312 *
## Year1997           -0.069266   0.040330   -1.72  0.08593 .
## Year1998           -0.056518   0.041528   -1.36  0.17357
## Year1999           -0.076006   0.036957   -2.06  0.03975 *
## Year2000            0.000459   0.036609    0.01  0.99000
## Year2001           -0.057960   0.036777   -1.58  0.11507
## Year2002           -0.095795   0.036614   -2.62  0.00890 **
## Year2003           -0.126542   0.035656   -3.55  0.00039 ***
## Year2004           -0.139066   0.035349   -3.93  8.4e-05 ***
## Year2005           -0.132544   0.035221   -3.76  0.00017 ***
## Year2006           -0.152163   0.036218   -4.20  2.7e-05 ***
```

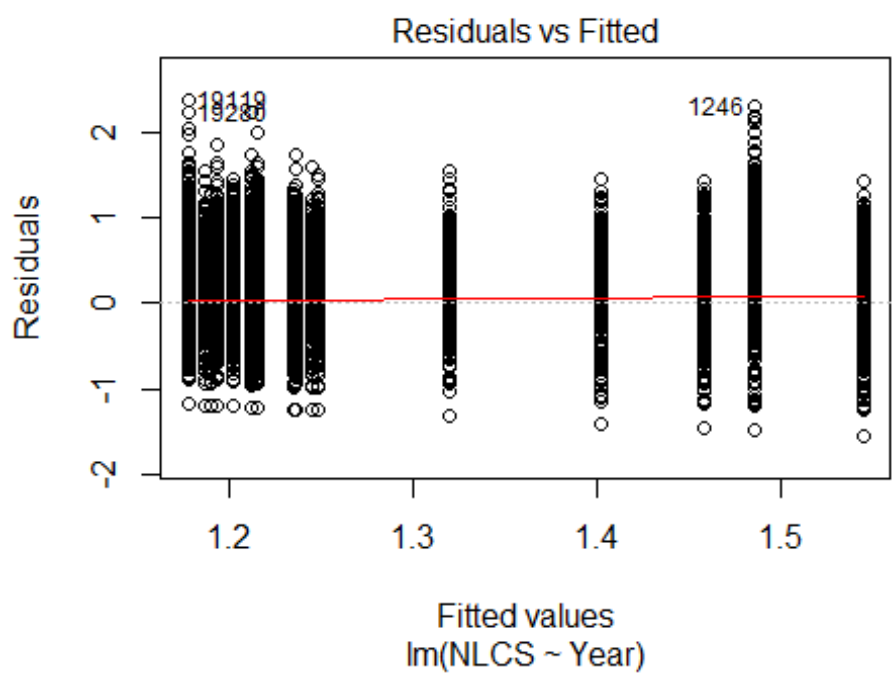
```
## Year2007           -0.054376   0.035823   -1.52  0.12908
## Year2008           -0.005880   0.035441   -0.17  0.86823
## Year2009            0.022425   0.036002    0.62  0.53339
## Year2010           -0.055515   0.036740   -1.51  0.13082
## Year2011           -0.115673   0.036305   -3.19  0.00145 **
## Year2012           -0.150670   0.035905   -4.20  2.7e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.416
## Multiple R-squared:  0.0173, Adjusted R-squared:  0.0151
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
##  629 weights are ~= 1. The remaining 7042 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0817  0.8640  0.9490  0.8960  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.30e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##                   GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.023  1           1.011
## Year             1.023 16           1.001
```



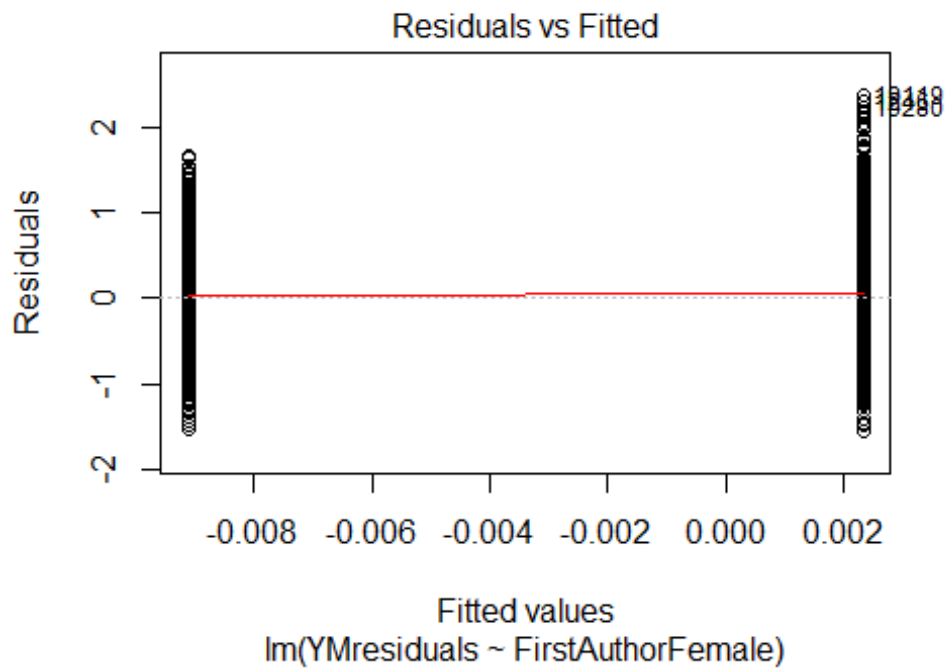

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.3096 -0.2859  0.0195  0.2834  1.6420
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.285676   0.030311   42.42  < 2e-16 ***
## LastAuthorFemale1 -0.067682   0.013414   -5.05  4.6e-07 ***
## Year1997          -0.071055   0.040409   -1.76  0.07872 .
## Year1998          -0.056711   0.041589   -1.36  0.17273
## Year1999          -0.076924   0.037000   -2.08  0.03765 *
## Year2000          -0.000628   0.036733   -0.02  0.98637
## Year2001          -0.058759   0.036933   -1.59  0.11166
## Year2002          -0.093914   0.036654   -2.56  0.01042 *
## Year2003          -0.123825   0.035740   -3.46  0.00053 ***
## Year2004          -0.140089   0.035415   -3.96  7.7e-05 ***
## Year2005          -0.131160   0.035316   -3.71  0.00021 ***
## Year2006          -0.153647   0.036315   -4.23  2.4e-05 ***
```

```
## Year2007          -0.051585   0.035934   -1.44  0.15117
## Year2008          -0.002848   0.035477   -0.08  0.93602
## Year2009           0.023949   0.036166    0.66  0.50787
## Year2010          -0.053376   0.036829   -1.45  0.14730
## Year2011          -0.114370   0.036366   -3.14  0.00167 **
## Year2012          -0.147661   0.035956   -4.11  4.1e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.416
## Multiple R-squared:  0.0202, Adjusted R-squared:  0.018
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
##  616 weights are ~= 1. The remaining 7055 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0848  0.8650  0.9490  0.8960  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.30e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Sample size for the above analysis:  7671"
## [1] ""
## [1] ""
## [1] "#####################################"
## [1] "Analysis of AJSC 2748"
## [1] "#####################################"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 1001 1372 1227 1165 1305 1406 1233 1019 1019 1113 1433 1741 1843 1762 1651
## 2011 2012
## 1630 1656
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  556  811  689  634  738  722  871  707  710  834 1090 1371 1428 1389 1310
## 2011 2012
```

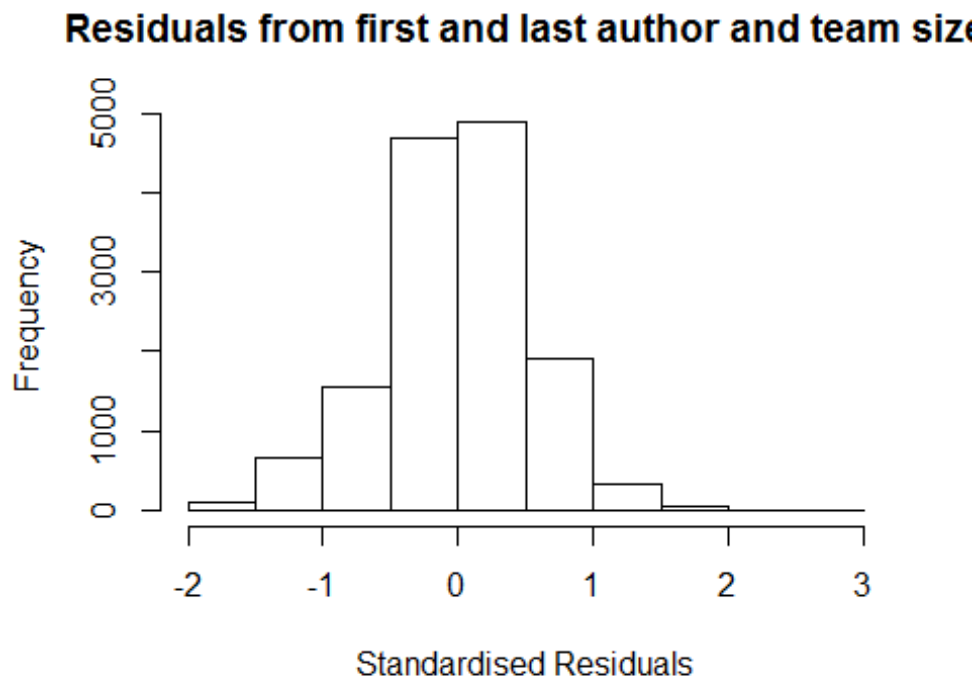
```
## 1297 1343
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  498  718  605  537  633  625  752  617  608  710  925 1147 1212 1198 1115
## 2011 2012
## 1112 1164
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
##  Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 610, df = 16, p-value <2e-16
```




```
##
##  Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 96, df = 1, p-value <2e-16
```

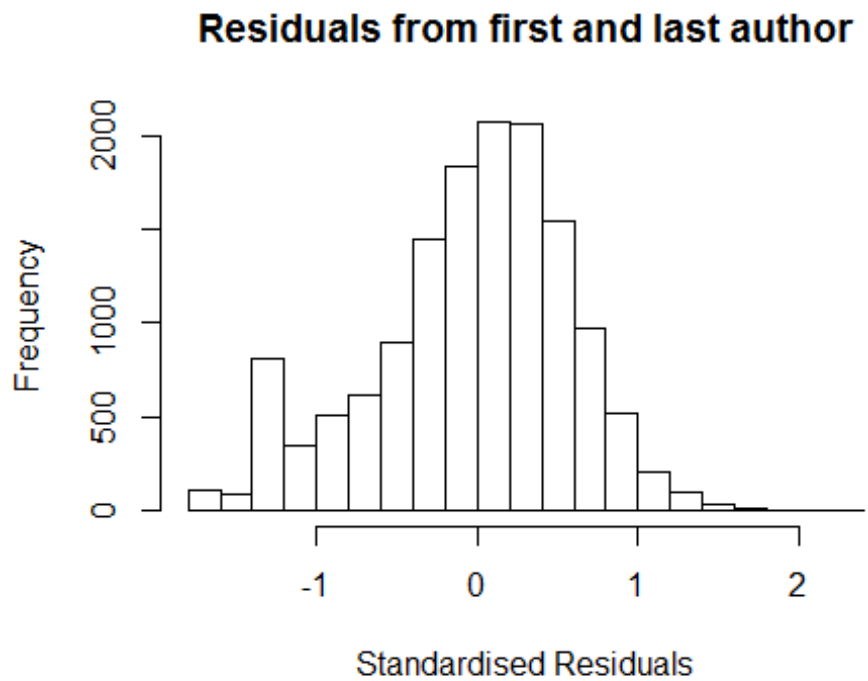



```
## [1] "Female first author team size 2018 geometric mean: 5.20270787039278"
## [1] "Male first author team size 2018 geometric mean: 5.56109178782433"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 180000, p-value = 0.001
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 5.30718345079352"
## [1] "Male last author team size 2018 geometric mean: 5.46713691903977"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 120000, p-value = 0.1
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.062  1           1.031
## LastAuthorFemale  1.048  1           1.024
## UniqueAuthors     1.097  4           1.012
## Year              1.123 16           1.004
```



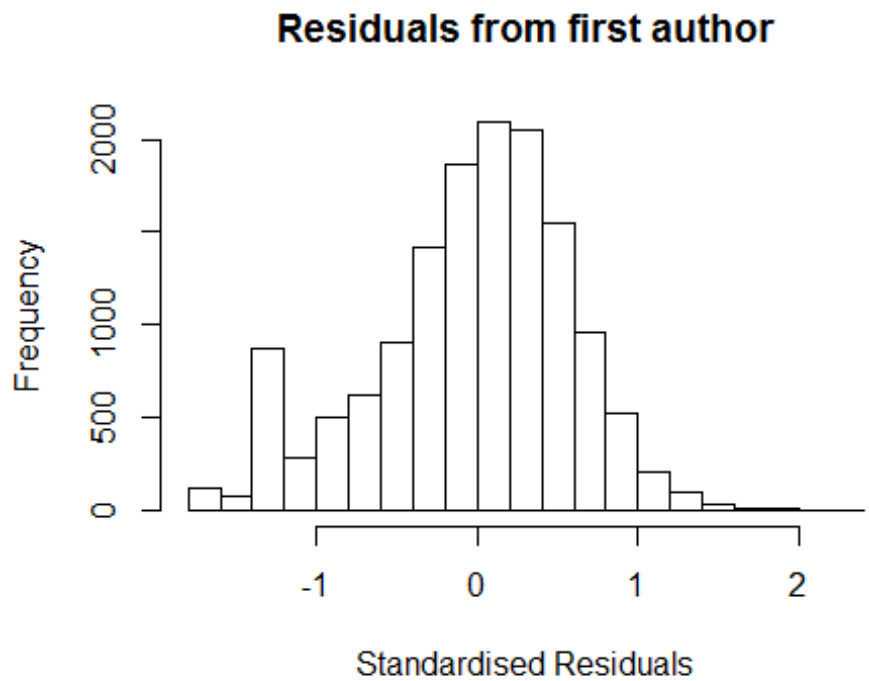

```
## [1] "List of  1 outliers with residuals above  2.5"
##        ScopusId  NLCS Year OneField Fields residuals
## 2362 0031012642 3.478 1997     2748      1      2.57
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
UniqueAuthors +
##     Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##     k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.8912 -0.3686  0.0126  0.3518  2.5705
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)           0.8473     0.0351   24.17  < 2e-16 ***
## FirstAuthorFemale1   -0.0447     0.0108   -4.13  3.6e-05 ***
## LastAuthorFemale1    -0.0194     0.0136   -1.43    0.153
## UniqueAuthors2        0.5921     0.0263   22.51  < 2e-16 ***
## UniqueAuthors3        0.7104     0.0240   29.63  < 2e-16 ***
## UniqueAuthors4        0.8157     0.0235   34.66  < 2e-16 ***
## UniqueAuthors5        0.9836     0.0212   46.30  < 2e-16 ***
## Year1997              0.0602     0.0502    1.20    0.231
## Year1998             -0.0662     0.0379   -1.75    0.081 .
## Year1999             -0.1612     0.0383   -4.21  2.6e-05 ***
```

```
## Year2000            -0.2326     0.0355   -6.55  5.8e-11 ***
## Year2001            -0.3327     0.0358   -9.29  < 2e-16 ***
## Year2002            -0.3396     0.0338  -10.05  < 2e-16 ***
## Year2003            -0.3881     0.0347  -11.19  < 2e-16 ***
## Year2004            -0.4191     0.0340  -12.34  < 2e-16 ***
## Year2005            -0.3799     0.0345  -11.00  < 2e-16 ***
## Year2006            -0.3772     0.0328  -11.49  < 2e-16 ***
## Year2007            -0.3801     0.0321  -11.86  < 2e-16 ***
## Year2008            -0.4044     0.0320  -12.62  < 2e-16 ***
## Year2009            -0.4147     0.0332  -12.50  < 2e-16 ***
## Year2010            -0.4156     0.0329  -12.63  < 2e-16 ***
## Year2011            -0.4362     0.0324  -13.47  < 2e-16 ***
## Year2012            -0.4598     0.0325  -14.14  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.494
## Multiple R-squared:  0.261,  Adjusted R-squared:  0.259
## Convergence in 15 IRWLS iterations
##
## Robustness weights:
##  2 observations c(1074,9914) are outliers with |weight| = 0 ( < 7.1e-06);
##  1100 weights are ~= 1. The remaining 13074 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0046  0.8600  0.9440  0.8880  0.9840  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          7.05e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.038  1           1.019
## LastAuthorFemale  1.024  1           1.012
## Year              1.046 16           1.001
```



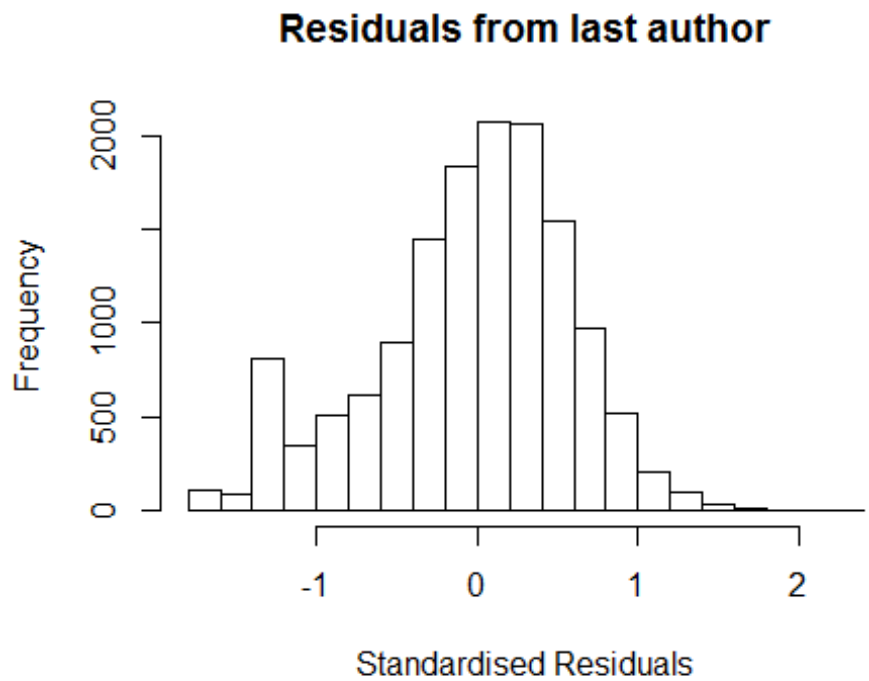

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##    Min     1Q Median     3Q    Max
## -1.617 -0.379  0.039  0.383  2.369
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)          1.58147    0.02721   58.12  < 2e-16 ***
## FirstAuthorFemale1 -0.00579    0.01177   -0.49   0.6225
## LastAuthorFemale1  -0.04185    0.01534   -2.73   0.0064 **
## Year1997            0.03590    0.05189    0.69   0.4890
## Year1998           -0.08500    0.03660   -2.32   0.0202 *
## Year1999           -0.12063    0.03845   -3.14   0.0017 **
## Year2000           -0.21426    0.03481   -6.15  7.7e-10 ***
## Year2001           -0.32028    0.03648   -8.78  < 2e-16 ***
## Year2002           -0.30521    0.03344   -9.13  < 2e-16 ***
## Year2003           -0.35099    0.03645   -9.63  < 2e-16 ***
## Year2004           -0.36448    0.03374  -10.80  < 2e-16 ***
## Year2005           -0.31332    0.03451   -9.08  < 2e-16 ***
```

```
## Year2006           -0.34123    0.03315  -10.29  < 2e-16 ***
## Year2007           -0.32397    0.03226  -10.04  < 2e-16 ***
## Year2008           -0.35392    0.03348  -10.57  < 2e-16 ***
## Year2009           -0.39909    0.03546  -11.26  < 2e-16 ***
## Year2010           -0.34284    0.03369  -10.18  < 2e-16 ***
## Year2011           -0.36066    0.03206  -11.25  < 2e-16 ***
## Year2012           -0.34717    0.03206  -10.83  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.552
## Multiple R-squared:  0.0386, Adjusted R-squared:  0.0374
## Convergence in 17 IRWLS iterations
##
## Robustness weights:
##  1195 weights are ~= 1. The remaining 12981 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0259  0.8510  0.9480  0.8880  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          7.05e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.032  1           1.016
## Year              1.032 16           1.001
```




```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.6136 -0.3758  0.0401  0.3820  2.3722
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)           1.5790     0.0273   57.93  < 2e-16 ***
## FirstAuthorFemale1   -0.0119     0.0118   -1.01   0.3123
## Year1997              0.0346     0.0519    0.67   0.5045
## Year1998             -0.0849     0.0367   -2.32   0.0206 *
## Year1999             -0.1199     0.0385   -3.12   0.0018 **
## Year2000             -0.2146     0.0349   -6.15  7.7e-10 ***
## Year2001             -0.3211     0.0365   -8.79  < 2e-16 ***
## Year2002             -0.3051     0.0335   -9.11  < 2e-16 ***
## Year2003             -0.3515     0.0365   -9.62  < 2e-16 ***
## Year2004             -0.3661     0.0338  -10.84  < 2e-16 ***
## Year2005             -0.3157     0.0345   -9.15  < 2e-16 ***
## Year2006             -0.3428     0.0332  -10.34  < 2e-16 ***
```

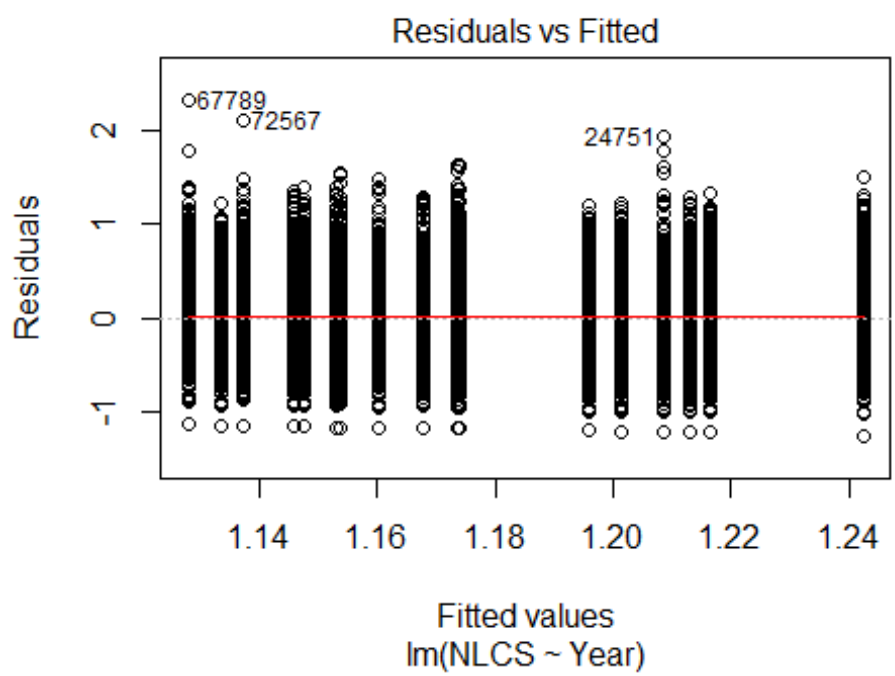
```
## Year2007            -0.3255     0.0323  -10.08  < 2e-16 ***
## Year2008            -0.3563     0.0335  -10.64  < 2e-16 ***
## Year2009            -0.4001     0.0355  -11.27  < 2e-16 ***
## Year2010            -0.3447     0.0337  -10.22  < 2e-16 ***
## Year2011            -0.3630     0.0321  -11.31  < 2e-16 ***
## Year2012            -0.3502     0.0320  -10.93  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.552
## Multiple R-squared:  0.0382, Adjusted R-squared:  0.037
## Convergence in 17 IRWLS iterations
##
## Robustness weights:
##  1164 weights are ~= 1. The remaining 13012 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0251  0.8510  0.9480  0.8890  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          7.05e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##                   GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.019  1           1.009
## Year             1.019 16           1.001
```



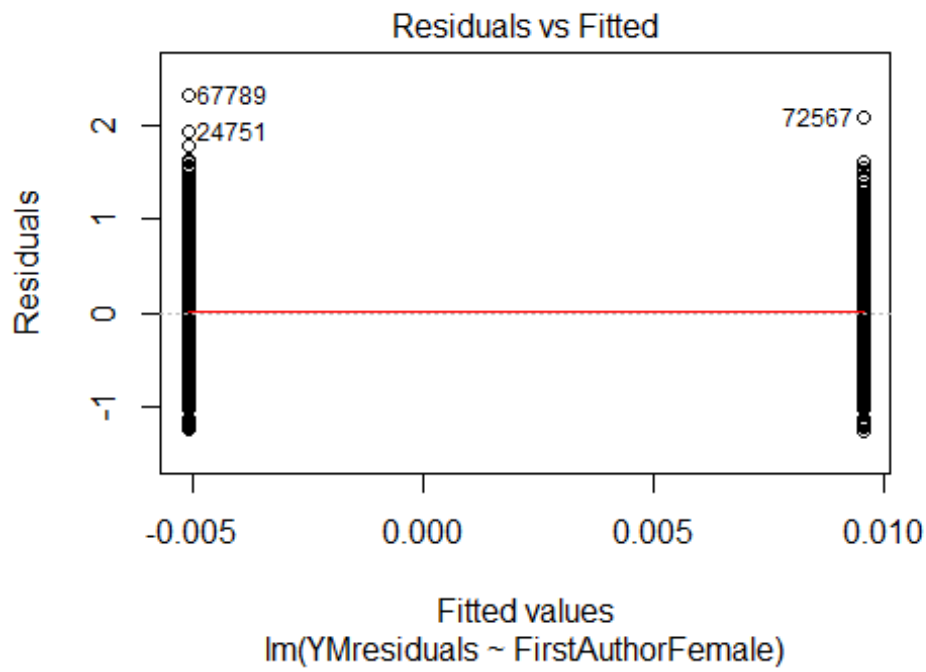

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.6157 -0.3781  0.0393  0.3833  2.3699
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)          1.5809     0.0272   58.15  < 2e-16 ***
## LastAuthorFemale1   -0.0432     0.0154   -2.81   0.0050 **
## Year1997             0.0348     0.0518    0.67   0.5011
## Year1998            -0.0851     0.0366   -2.33   0.0200 *
## Year1999            -0.1212     0.0384   -3.15   0.0016 **
## Year2000            -0.2146     0.0348   -6.17  7.1e-10 ***
## Year2001            -0.3206     0.0365   -8.79  < 2e-16 ***
## Year2002            -0.3057     0.0334   -9.15  < 2e-16 ***
## Year2003            -0.3517     0.0364   -9.65  < 2e-16 ***
## Year2004            -0.3651     0.0337  -10.83  < 2e-16 ***
## Year2005            -0.3140     0.0345   -9.11  < 2e-16 ***
## Year2006            -0.3420     0.0331  -10.34  < 2e-16 ***
```

```
## Year2007           -0.3248     0.0322  -10.10  < 2e-16 ***
## Year2008           -0.3548     0.0334  -10.63  < 2e-16 ***
## Year2009           -0.3998     0.0354  -11.29  < 2e-16 ***
## Year2010           -0.3437     0.0336  -10.23  < 2e-16 ***
## Year2011           -0.3616     0.0320  -11.31  < 2e-16 ***
## Year2012           -0.3481     0.0320  -10.89  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.553
## Multiple R-squared:  0.0385, Adjusted R-squared:  0.0374
## Convergence in 16 IRWLS iterations
##
## Robustness weights:
##  1168 weights are ~= 1. The remaining 13008 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0264  0.8520  0.9480  0.8890  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          7.05e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Sample size for the above analysis:  14176"
## [1] ""
## [1] ""
## [1] "####################################"
## [1] "Analysis of AJSC 2800"
## [1] "####################################"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 3653 3746 3883 3737 3744 3698 3694 3692 3990 4052 4548 4514 4076 3489 3810
## 2011 2012
## 3918 3822
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 2268 2336 2481 2301 2233 2028 2476 2448 2699 2722 2983 3010 2700 2271 2506
## 2011 2012
```

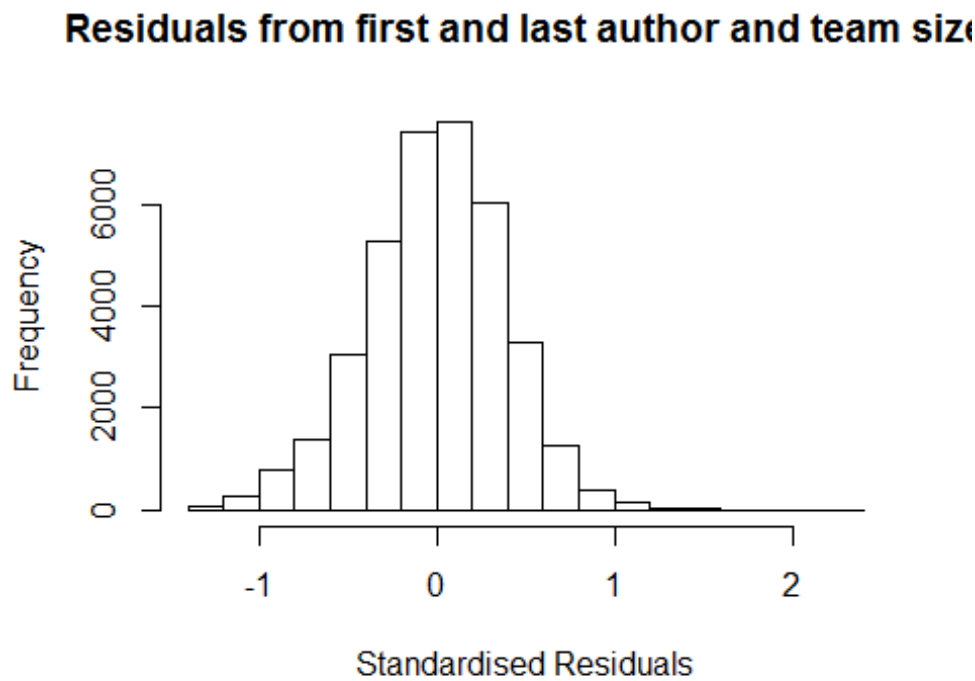
```
## 2615 2556
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 2013 2073 2201 2018 1937 1753 2146 2148 2351 2346 2571 2617 2325 1966 2175
## 2011 2012
## 2244 2196
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
##  Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 190, df = 16, p-value <2e-16
```




```
##
##  Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 95, df = 1, p-value <2e-16
```

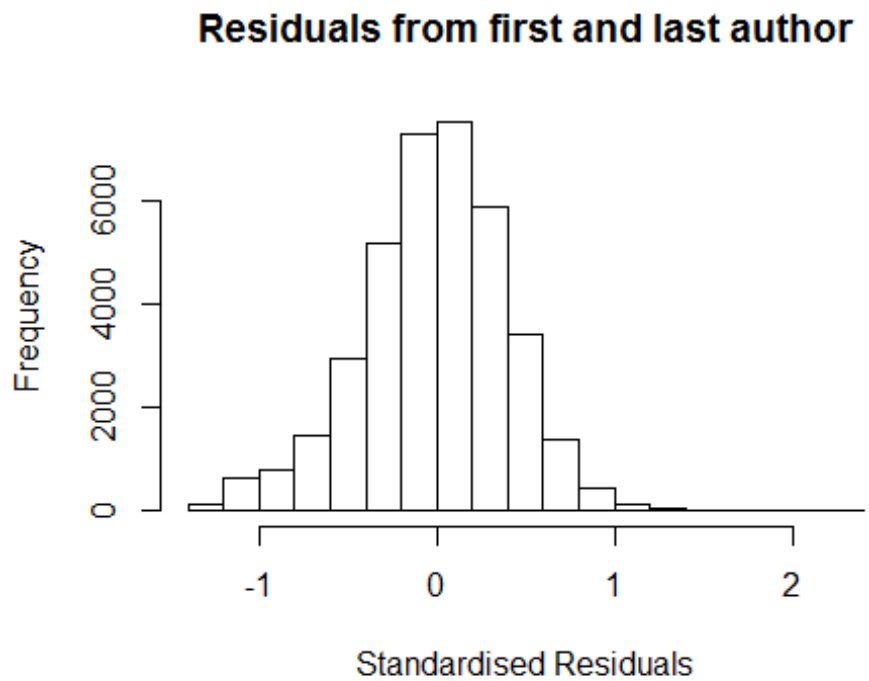



```
## [1] "Female first author team size 2018 geometric mean: 4.62059654418437"
## [1] "Male first author team size 2018 geometric mean: 4.18312912893973"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 870000, p-value = 7e-04
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 4.40329308600654"
## [1] "Male last author team size 2018 geometric mean: 4.34979473947918"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 680000, p-value = 0.8
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.019  1           1.009
## LastAuthorFemale  1.013  1           1.006
## UniqueAuthors     1.052  4           1.006
## Year              1.057 16           1.002
```



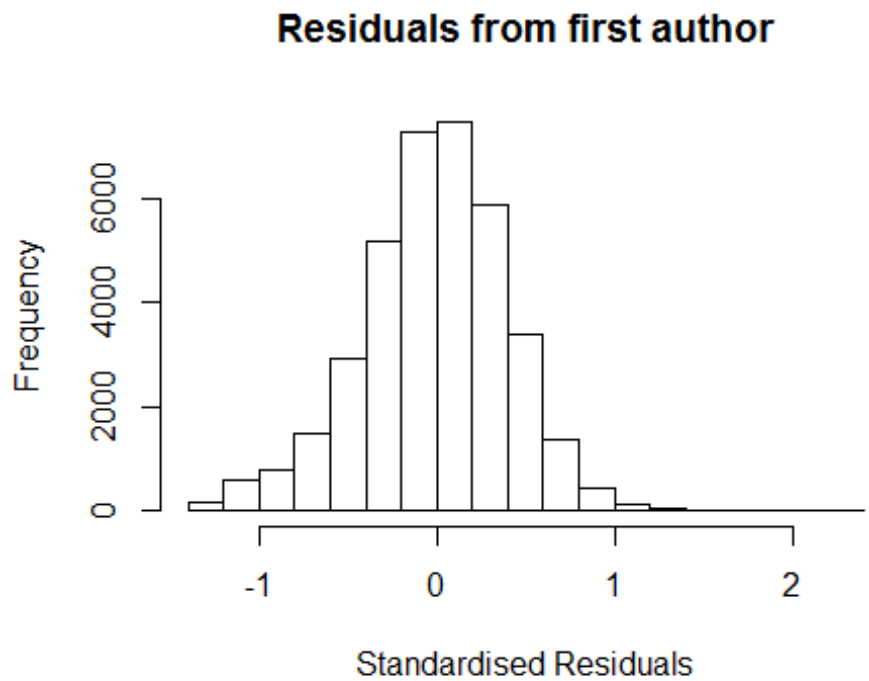

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
UniqueAuthors +
##     Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##     k.max = 1000))
##  \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.38748 -0.25190  0.00652  0.25511  2.27262
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         0.91768    0.01416   64.83  < 2e-16 ***
## FirstAuthorFemale1  0.00200    0.00422    0.48   0.6347
## LastAuthorFemale1  -0.02493    0.00488   -5.11  3.3e-07 ***
## UniqueAuthors2      0.28922    0.01169   24.74  < 2e-16 ***
## UniqueAuthors3      0.32664    0.01172   27.87  < 2e-16 ***
## UniqueAuthors4      0.36295    0.01194   30.41  < 2e-16 ***
## UniqueAuthors5      0.46779    0.01160   40.33  < 2e-16 ***
## Year1997            0.01809    0.01313    1.38   0.1682
## Year1998           -0.00809    0.01293   -0.63   0.5319
## Year1999           -0.02627    0.01277   -2.06   0.0396 *
```

```
## Year2000           -0.03668    0.01271   -2.89   0.0039 **
## Year2001           -0.03930    0.01324   -2.97   0.0030 **
## Year2002           -0.07766    0.01274   -6.10  1.1e-09 ***
## Year2003           -0.10685    0.01269   -8.42  < 2e-16 ***
## Year2004           -0.07420    0.01245   -5.96  2.5e-09 ***
## Year2005           -0.10595    0.01232   -8.60  < 2e-16 ***
## Year2006           -0.10683    0.01205   -8.87  < 2e-16 ***
## Year2007           -0.10909    0.01191   -9.16  < 2e-16 ***
## Year2008           -0.09908    0.01257   -7.88  3.4e-15 ***
## Year2009           -0.10320    0.01289   -8.00  1.2e-15 ***
## Year2010           -0.11143    0.01326   -8.41  < 2e-16 ***
## Year2011           -0.13131    0.01287  -10.21  < 2e-16 ***
## Year2012           -0.11913    0.01312   -9.08  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.375
## Multiple R-squared:  0.097,  Adjusted R-squared:  0.0965
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
##  4 observations c(11930,11934,33716,36045)
##   are outliers with |weight| = 0 ( < 2.7e-06);
##  3027 weights are ~= 1. The remaining 34049 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0026  0.8690  0.9510  0.9010  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          2.70e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.015  1           1.007
## LastAuthorFemale  1.010  1           1.005
## Year              1.009 16           1.000
```



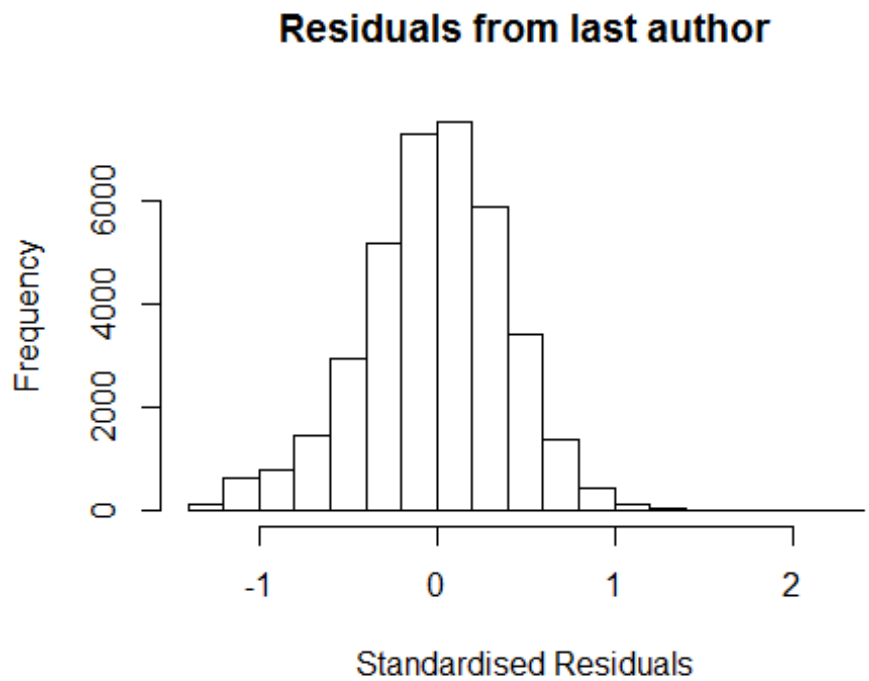

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.26554 -0.25767  0.00661  0.25868  2.30395
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)          1.22934    0.00959  128.21  < 2e-16 ***
## FirstAuthorFemale1   0.01749    0.00434    4.03  5.5e-05 ***
## LastAuthorFemale1   -0.03215    0.00502   -6.41  1.5e-10 ***
## Year1997             0.01871    0.01326    1.41   0.1582
## Year1998            -0.00470    0.01290   -0.36   0.7154
## Year1999            -0.01537    0.01271   -1.21   0.2267
## Year2000            -0.02153    0.01256   -1.71   0.0865 .
## Year2001            -0.00834    0.01325   -0.63   0.5291
## Year2002            -0.05336    0.01296   -4.12  3.8e-05 ***
## Year2003            -0.07532    0.01296   -5.81  6.2e-09 ***
## Year2004            -0.03864    0.01262   -3.06   0.0022 **
## Year2005            -0.07282    0.01253   -5.81  6.2e-09 ***
```

```
## Year2006           -0.07308    0.01208   -6.05  1.5e-09 ***
## Year2007           -0.06767    0.01196   -5.66  1.6e-08 ***
## Year2008           -0.05090    0.01259   -4.04  5.3e-05 ***
## Year2009           -0.05570    0.01307   -4.26  2.0e-05 ***
## Year2010           -0.06143    0.01340   -4.58  4.6e-06 ***
## Year2011           -0.08629    0.01311   -6.58  4.6e-11 ***
## Year2012           -0.06962    0.01336   -5.21  1.9e-07 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.384
## Multiple R-squared:  0.00733,    Adjusted R-squared:  0.00685
## Convergence in 10 IRWLS iterations
##
## Robustness weights:
##  3 observations c(11930,33716,36045)
##   are outliers with |weight| = 0 ( < 2.7e-06);
##  3086 weights are ~= 1. The remaining 33991 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0008  0.8680  0.9510  0.8970  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          2.70e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.007  1           1.003
## Year              1.007 16           1.000
```




```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.25616 -0.25825  0.00506  0.25941  2.31026
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)          1.22462    0.00955  128.20  < 2e-16 ***
## FirstAuthorFemale1   0.01363    0.00433    3.15   0.0016 **
## Year1997             0.01791    0.01326    1.35   0.1768
## Year1998            -0.00494    0.01290   -0.38   0.7017
## Year1999            -0.01625    0.01271   -1.28   0.2011
## Year2000            -0.02268    0.01256   -1.81   0.0709 .
## Year2001            -0.01035    0.01325   -0.78   0.4349
## Year2002            -0.05456    0.01296   -4.21  2.6e-05 ***
## Year2003            -0.07642    0.01296   -5.90  3.7e-09 ***
## Year2004            -0.03968    0.01262   -3.15   0.0017 **
## Year2005            -0.07409    0.01253   -5.91  3.4e-09 ***
## Year2006            -0.07429    0.01207   -6.15  7.7e-10 ***
```

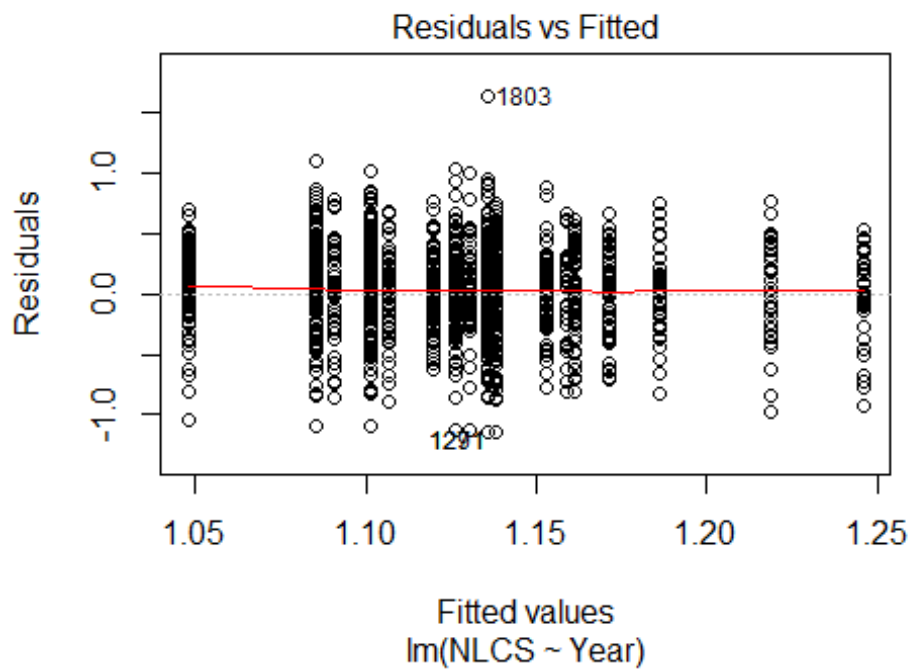
```
## Year2007           -0.06852    0.01196   -5.73  1.0e-08 ***
## Year2008           -0.05208    0.01260   -4.13  3.6e-05 ***
## Year2009           -0.05653    0.01307   -4.33  1.5e-05 ***
## Year2010           -0.06267    0.01341   -4.67  3.0e-06 ***
## Year2011           -0.08787    0.01310   -6.71  2.0e-11 ***
## Year2012           -0.07120    0.01336   -5.33  9.8e-08 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.384
## Multiple R-squared:  0.00623,    Adjusted R-squared:  0.00578
## Convergence in 10 IRWLS iterations
##
## Robustness weights:
##  3 observations c(11930,33716,36045)
##   are outliers with |weight| = 0 ( < 2.7e-06);
##  3091 weights are ~= 1. The remaining 33986 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0005  0.8690  0.9510  0.8970  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          2.70e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##                   GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.002  1           1.001
## Year             1.002 16           1.000
```



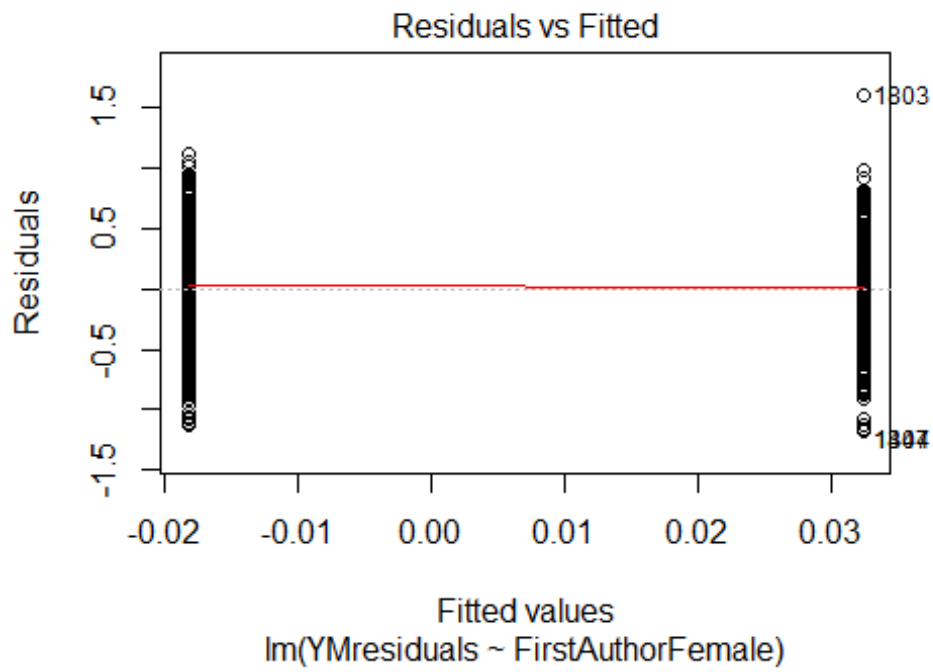

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.25332 -0.25914  0.00586  0.25886  2.29766
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.23476    0.00949  130.07  < 2e-16 ***
## LastAuthorFemale1 -0.02936    0.00500   -5.87  4.4e-09 ***
## Year1997            0.01857    0.01326    1.40   0.1615
## Year1998           -0.00538    0.01290   -0.42   0.6767
## Year1999           -0.01568    0.01272   -1.23   0.2175
## Year2000           -0.02180    0.01257   -1.73   0.0828 .
## Year2001           -0.00864    0.01325   -0.65   0.5144
## Year2002           -0.05327    0.01297   -4.11  4.0e-05 ***
## Year2003           -0.07508    0.01296   -5.79  7.0e-09 ***
## Year2004           -0.03846    0.01262   -3.05   0.0023 **
## Year2005           -0.07276    0.01253   -5.80  6.5e-09 ***
## Year2006           -0.07279    0.01208   -6.02  1.7e-09 ***
```

```
## Year2007          -0.06736    0.01198   -5.63  1.9e-08 ***
## Year2008          -0.05062    0.01259   -4.02  5.8e-05 ***
## Year2009          -0.05473    0.01307   -4.19  2.8e-05 ***
## Year2010          -0.06040    0.01340   -4.51  6.6e-06 ***
## Year2011          -0.08542    0.01310   -6.52  7.2e-11 ***
## Year2012          -0.06862    0.01335   -5.14  2.8e-07 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.384
## Multiple R-squared:  0.00691,    Adjusted R-squared:  0.00646
## Convergence in 10 IRWLS iterations
##
## Robustness weights:
##  3 observations c(11930,33716,36045)
##   are outliers with |weight| = 0 ( < 2.7e-06);
##  3079 weights are ~= 1. The remaining 33998 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0012  0.8690  0.9510  0.8970  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          2.70e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                  psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Sample size for the above analysis:  37080"
## [1] ""
## [1] ""
## [1] "##################################"
## [1] "Analysis of AJSC 2801"
## [1] "##################################"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##   43   64   75   51   69   72   61   57   73   71   87  107  128  199  229
## 2011 2012
##  285  337
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
```

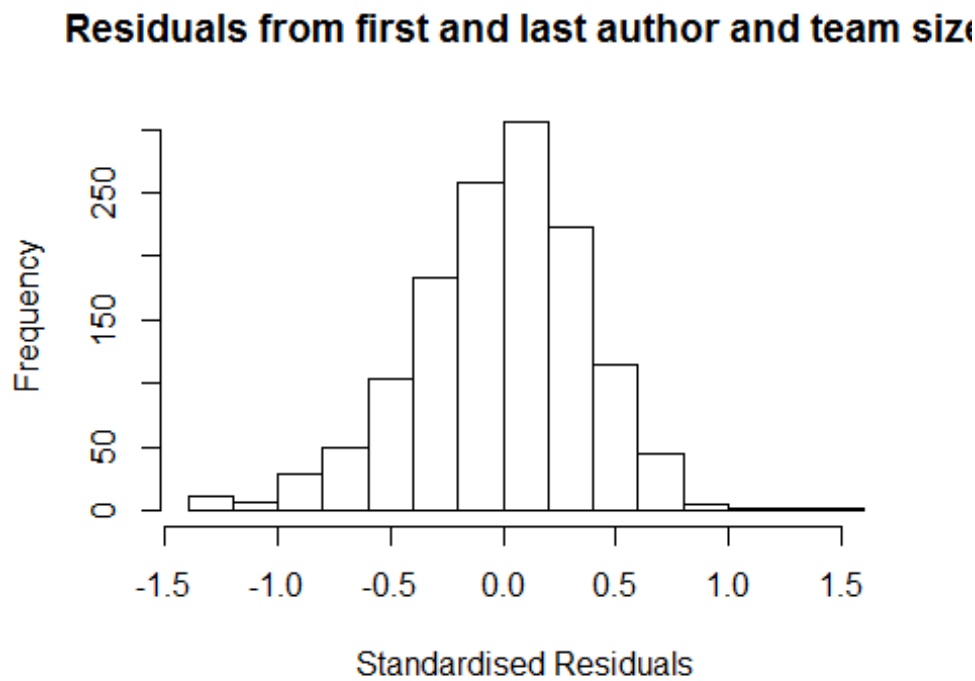
```
##   30   39   54   41   53   46   51   45   57   61   76   81  101  149  177
## 2011 2012
##  219  264
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##   25   36   41   37   45   42   43   33   50   44   62   67   89  134  158
## 2011 2012
##  195  237
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
##  Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 24, df = 16, p-value = 0.09
```




```
##
##  Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 3.1, df = 1, p-value = 0.08
```

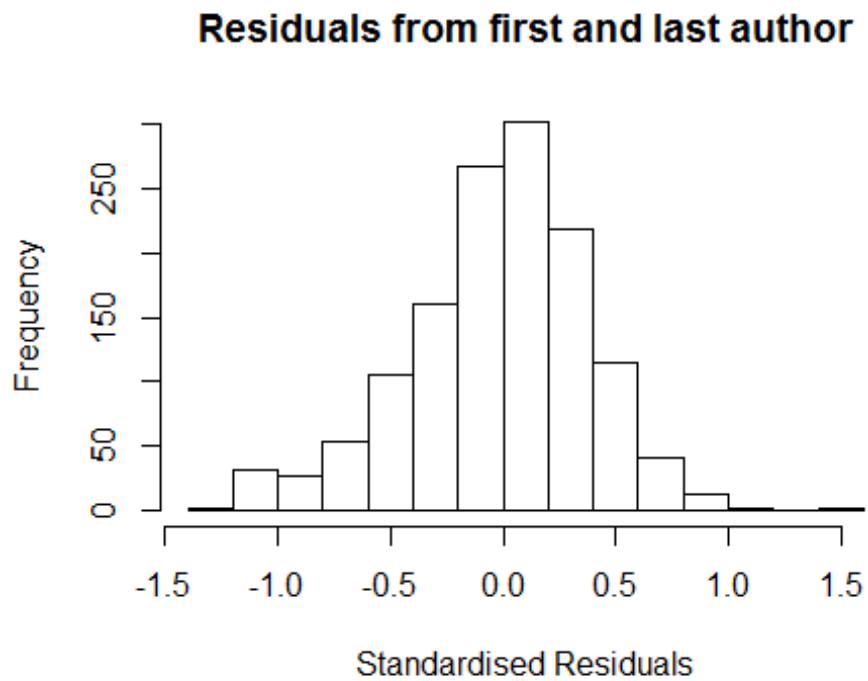



```
## [1] "Female first author team size 2018 geometric mean: 5.24289353852042"
## [1] "Male first author team size 2018 geometric mean: 4.51844310022277"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 9300, p-value = 0.1
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 4.66476542310077"
## [1] "Male last author team size 2018 geometric mean: 4.95550158020123"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 7000, p-value = 0.2
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.050  1           1.025
## LastAuthorFemale  1.100  1           1.049
## UniqueAuthors     1.369  4           1.040
## Year              1.469 16           1.012
```



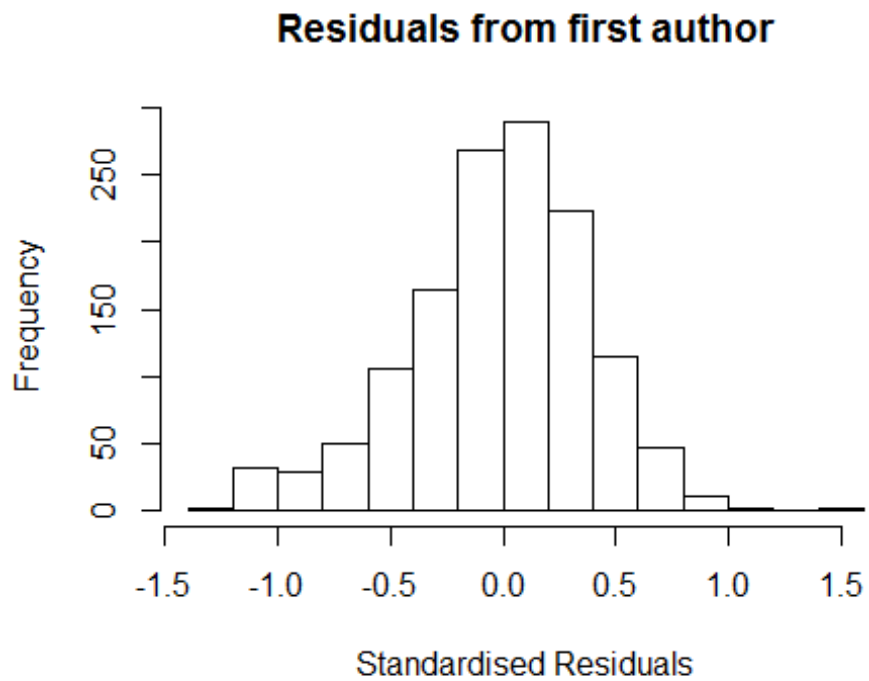

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
UniqueAuthors +
##     Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##     k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.3250 -0.2527  0.0136  0.2391  1.4525
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)          0.9769     0.1116    8.75  < 2e-16 ***
## FirstAuthorFemale1   0.0258     0.0214    1.21   0.2277
## LastAuthorFemale1    0.0347     0.0237    1.47   0.1424
## UniqueAuthors2       0.1803     0.0688    2.62   0.0089 **
## UniqueAuthors3       0.2092     0.0686    3.05   0.0023 **
## UniqueAuthors4       0.2331     0.0693    3.36   0.0008 ***
## UniqueAuthors5       0.4082     0.0651    6.27  4.8e-10 ***
## Year1997            -0.0640     0.1134   -0.56   0.5729
## Year1998            -0.0468     0.1135   -0.41   0.6798
## Year1999             0.0191     0.1067    0.18   0.8582
```

```
## Year2000            -0.0786     0.1047   -0.75   0.4527
## Year2001            -0.0843     0.1057   -0.80   0.4253
## Year2002            -0.0913     0.1072   -0.85   0.3943
## Year2003            -0.1138     0.1178   -0.97   0.3343
## Year2004            -0.1811     0.1065   -1.70   0.0894 .
## Year2005            -0.0668     0.1090   -0.61   0.5403
## Year2006            -0.1861     0.1021   -1.82   0.0685 .
## Year2007            -0.1608     0.1055   -1.52   0.1275
## Year2008            -0.1760     0.1028   -1.71   0.0873 .
## Year2009            -0.1722     0.1030   -1.67   0.0948 .
## Year2010            -0.1207     0.1003   -1.20   0.2293
## Year2011            -0.1333     0.0993   -1.34   0.1800
## Year2012            -0.1599     0.0987   -1.62   0.1056
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.361
## Multiple R-squared:  0.115,  Adjusted R-squared:   0.1
## Convergence in 16 IRWLS iterations
##
## Robustness weights:
##  109 weights are ~= 1. The remaining 1229 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0687  0.8670  0.9530  0.8980  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          7.47e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.059  1           1.029
## LastAuthorFemale  1.067  1           1.033
## Year              1.128 16           1.004
```



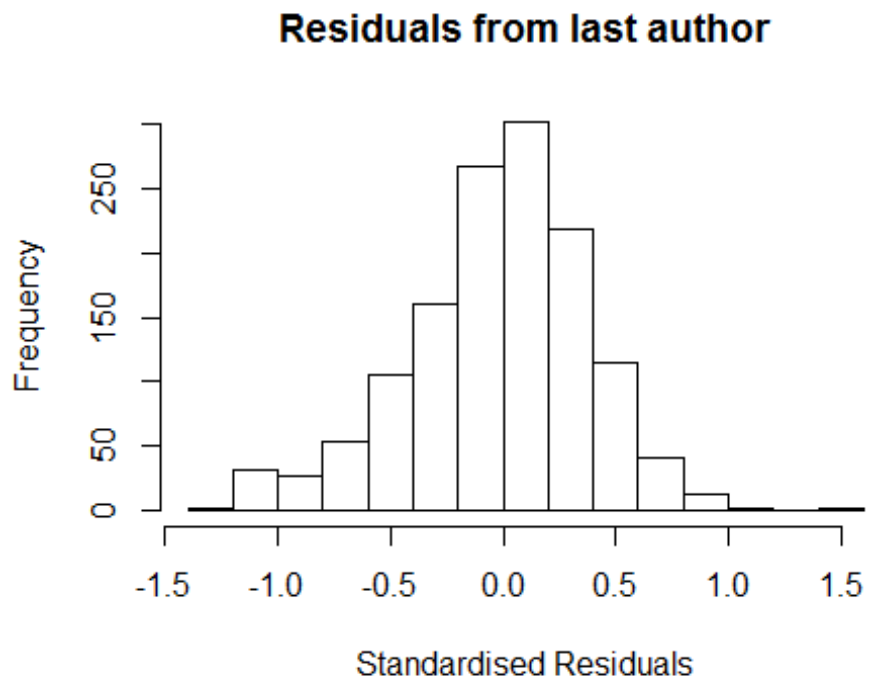

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.2272 -0.2511  0.0133  0.2490  1.5553
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)           1.2270     0.0926   13.25   <2e-16 ***
## FirstAuthorFemale1    0.0498     0.0223    2.23    0.026 *
## LastAuthorFemale1     0.0259     0.0243    1.07    0.287
## Year1997             -0.0592     0.1136   -0.52    0.603
## Year1998             -0.0912     0.1122   -0.81    0.417
## Year1999              0.0489     0.1071    0.46    0.648
## Year2000             -0.0725     0.1061   -0.68    0.494
## Year2001             -0.0363     0.1078   -0.34    0.736
## Year2002             -0.0881     0.1095   -0.80    0.422
## Year2003             -0.1219     0.1208   -1.01    0.313
## Year2004             -0.1383     0.1069   -1.29    0.196
## Year2005             -0.0331     0.1054   -0.31    0.753
```

```
## Year2006            -0.1190     0.0998   -1.19    0.233
## Year2007            -0.1422     0.1027   -1.38    0.166
## Year2008            -0.1163     0.1016   -1.15    0.252
## Year2009            -0.1268     0.1012   -1.25    0.211
## Year2010            -0.0756     0.0978   -0.77    0.440
## Year2011            -0.0930     0.0968   -0.96    0.337
## Year2012            -0.1356     0.0967   -1.40    0.161
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.368
## Multiple R-squared:  0.0161, Adjusted R-squared:  0.00263
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
##  123 weights are ~= 1. The remaining 1215 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0346  0.8610  0.9500  0.8910  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          7.47e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.056  1           1.028
## Year              1.056 16           1.002
```




```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.2092 -0.2489  0.0135  0.2456  1.5715
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)           1.2338     0.0913   13.51   <2e-16 ***
## FirstAuthorFemale1    0.0528     0.0224    2.36    0.018 *
## Year1997             -0.0622     0.1131   -0.55    0.582
## Year1998             -0.0889     0.1112   -0.80    0.424
## Year1999              0.0472     0.1062    0.44    0.657
## Year2000             -0.0746     0.1051   -0.71    0.478
## Year2001             -0.0405     0.1067   -0.38    0.704
## Year2002             -0.0886     0.1087   -0.82    0.415
## Year2003             -0.1256     0.1201   -1.05    0.296
## Year2004             -0.1358     0.1060   -1.28    0.201
## Year2005             -0.0335     0.1044   -0.32    0.748
## Year2006             -0.1207     0.0990   -1.22    0.223
```

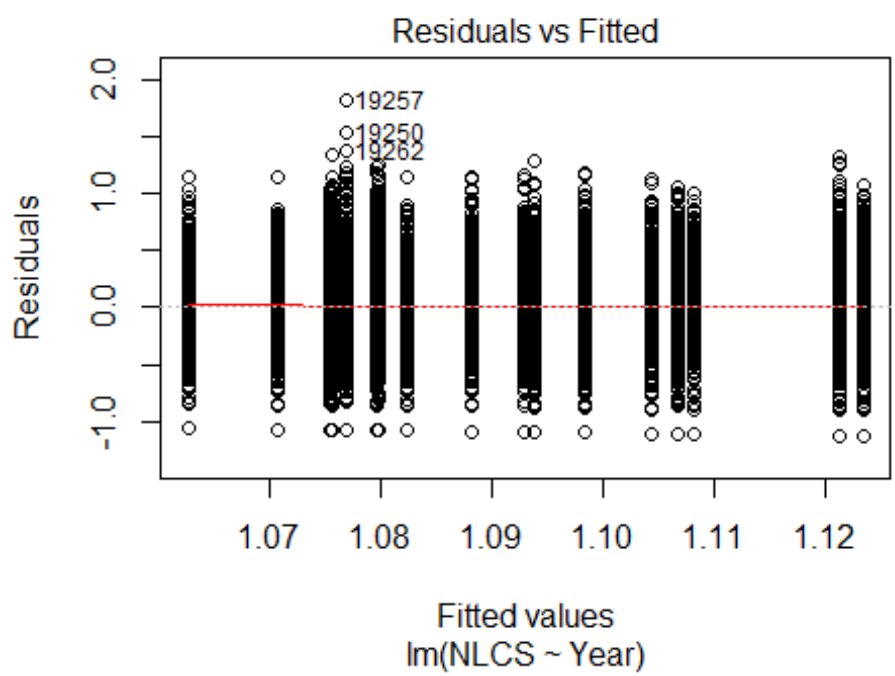
```
## Year2007            -0.1417     0.1022   -1.39    0.166
## Year2008            -0.1182     0.1008   -1.17    0.241
## Year2009            -0.1295     0.1003   -1.29    0.197
## Year2010            -0.0774     0.0969   -0.80    0.425
## Year2011            -0.0931     0.0960   -0.97    0.332
## Year2012            -0.1354     0.0959   -1.41    0.158
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.368
## Multiple R-squared:  0.0152, Adjusted R-squared:  0.00247
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
##  121 weights are ~= 1. The remaining 1217 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0282  0.8630  0.9490  0.8910  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          7.47e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##                   GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.066  1           1.033
## Year             1.066 16           1.002
```



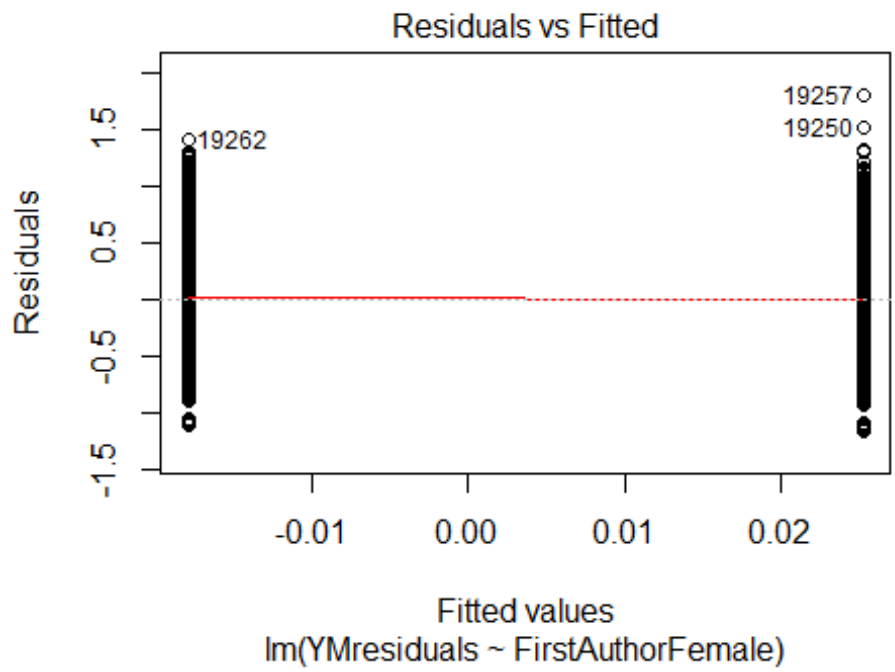

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.2041 -0.2553  0.0156  0.2425  1.5806
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)          1.2335     0.0955   12.91   <2e-16 ***
## LastAuthorFemale1    0.0327     0.0244    1.34     0.18
## Year1997            -0.0486     0.1159   -0.42     0.67
## Year1998            -0.0885     0.1141   -0.78     0.44
## Year1999             0.0502     0.1097    0.46     0.65
## Year2000            -0.0617     0.1084   -0.57     0.57
## Year2001            -0.0243     0.1102   -0.22     0.83
## Year2002            -0.0774     0.1121   -0.69     0.49
## Year2003            -0.1083     0.1221   -0.89     0.38
## Year2004            -0.1264     0.1083   -1.17     0.24
## Year2005            -0.0211     0.1079   -0.20     0.85
## Year2006            -0.1087     0.1019   -1.07     0.29
```

```
## Year2007           -0.1419     0.1052   -1.35     0.18
## Year2008           -0.1023     0.1037   -0.99     0.32
## Year2009           -0.1154     0.1032   -1.12     0.26
## Year2010           -0.0621     0.1000   -0.62     0.53
## Year2011           -0.0818     0.0990   -0.83     0.41
## Year2012           -0.1251     0.0991   -1.26     0.21
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.368
## Multiple R-squared:  0.0123, Adjusted R-squared:  -0.000429
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
##  123 weights are ~= 1. The remaining 1215 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0254  0.8600  0.9490  0.8910  0.9860  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          7.47e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Sample size for the above analysis:  1338"
## [1] ""
## [1] ""
## [1] "####################################"
## [1] "Analysis of AJSC 2802"
## [1] "####################################"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 1080  946  909  818  888  981 1027  913 1015 1055 1232 1281 1300 1345 1298
## 2011 2012
## 1535 1474
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  796  732  657  611  633  674  832  722  793  868  997 1049 1048 1084 1046
## 2011 2012
```

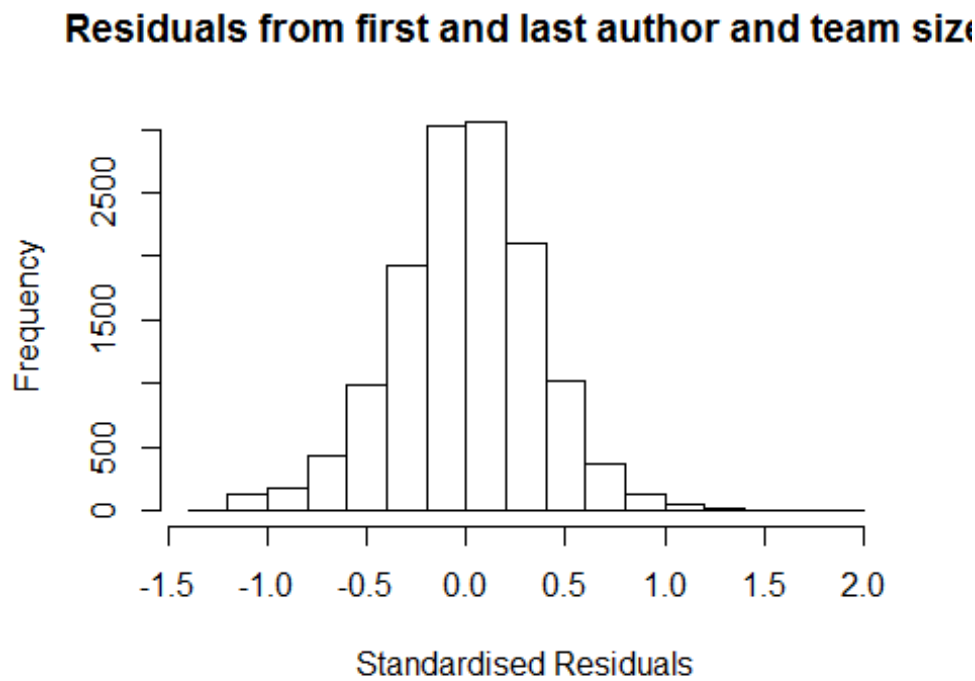
```
## 1246 1155
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  723  655  577  552  567  600  762  659  698  791  882  940  955  973  939
## 2011 2012
## 1120 1041
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
##  Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 100, df = 16, p-value = 7e-15
```




```
##
##  Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 34, df = 1, p-value = 4e-09
```

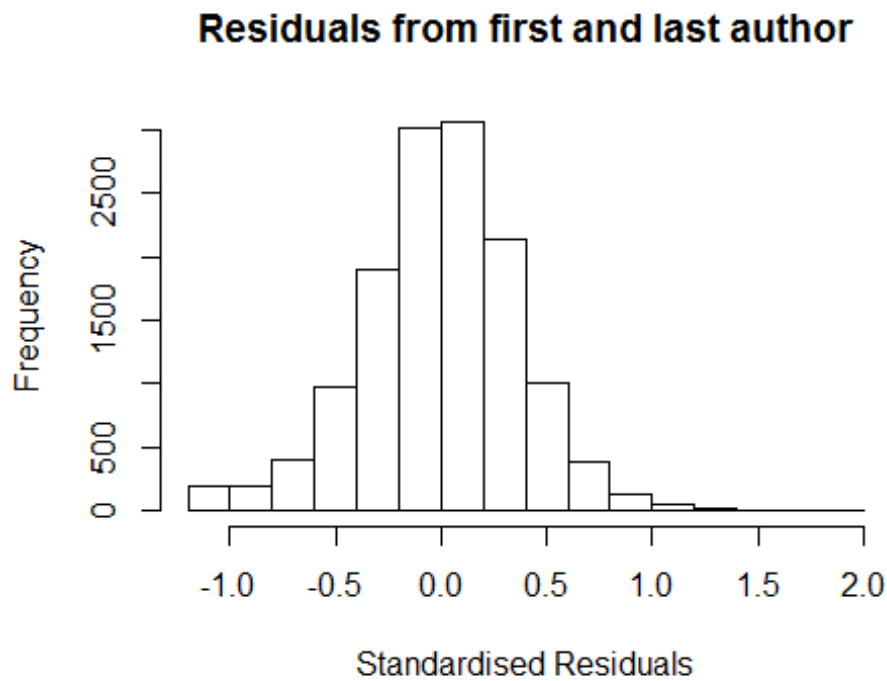



```
## [1] "Female first author team size 2018 geometric mean: 4.00488686024521"
## [1] "Male first author team size 2018 geometric mean: 3.78272296131065"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 150000, p-value = 0.1
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 3.96749270116032"
## [1] "Male last author team size 2018 geometric mean: 3.85353227460788"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 150000, p-value = 0.4
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.031  1           1.015
## LastAuthorFemale  1.033  1           1.016
## UniqueAuthors     1.087  4           1.011
## Year              1.101 16           1.003
```



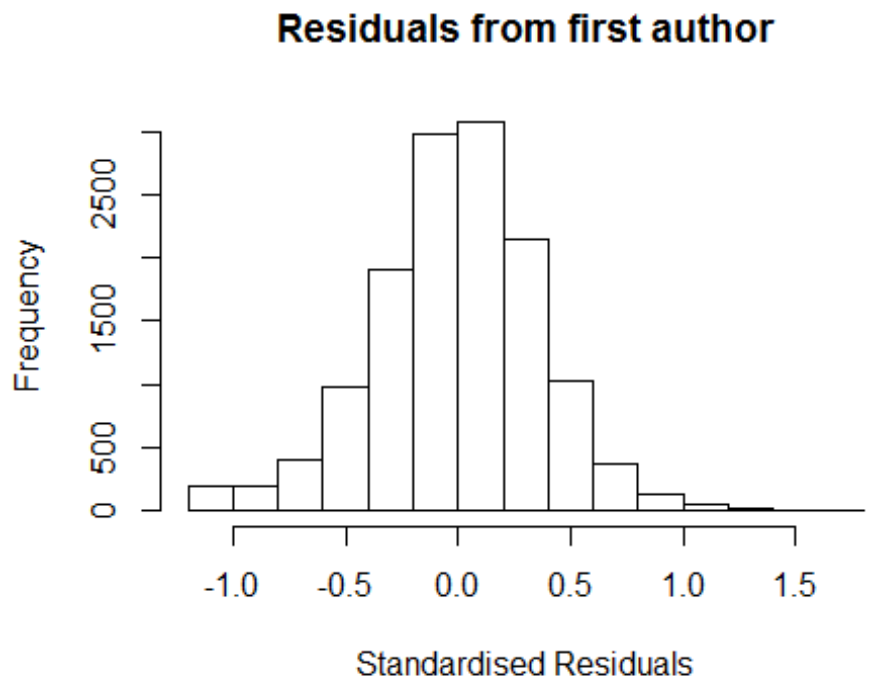

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
UniqueAuthors +
##     Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##     k.max = 1000))
##  \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.22453 -0.22560  0.00126  0.22365  1.84253
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.02455    0.01916   53.48  < 2e-16 ***
## FirstAuthorFemale1  0.02545    0.00624    4.08  4.6e-05 ***
## LastAuthorFemale1   0.02248    0.00661    3.40  0.00067 ***
## UniqueAuthors2      0.08089    0.01441    5.61  2.0e-08 ***
## UniqueAuthors3      0.08633    0.01464    5.90  3.8e-09 ***
## UniqueAuthors4      0.11824    0.01510    7.83  5.3e-15 ***
## UniqueAuthors5      0.16966    0.01466   11.57  < 2e-16 ***
## Year1997            0.00559    0.02114    0.26  0.79134
## Year1998           -0.03394    0.02130   -1.59  0.11106
## Year1999           -0.02322    0.02075   -1.12  0.26313
```

```
## Year2000           -0.02047    0.02058   -0.99  0.31990
## Year2001           -0.01761    0.02067   -0.85  0.39422
## Year2002           -0.04131    0.01975   -2.09  0.03649 *
## Year2003           -0.05257    0.02019   -2.60  0.00924 **
## Year2004           -0.03817    0.01902   -2.01  0.04474 *
## Year2005           -0.07279    0.01859   -3.91  9.1e-05 ***
## Year2006           -0.04375    0.01824   -2.40  0.01648 *
## Year2007           -0.05214    0.01790   -2.91  0.00359 **
## Year2008           -0.06420    0.01811   -3.54  0.00039 ***
## Year2009           -0.06949    0.01816   -3.83  0.00013 ***
## Year2010           -0.04534    0.01808   -2.51  0.01218 *
## Year2011           -0.07642    0.01870   -4.09  4.4e-05 ***
## Year2012           -0.07064    0.01915   -3.69  0.00023 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.336
## Multiple R-squared:  0.0241, Adjusted R-squared:  0.0225
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
##  observation 12140 is an outlier with |weight| = 0 ( < 7.4e-06);
##  1134 weights are ~= 1. The remaining 12299 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0018  0.8640  0.9510  0.8950  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          7.44e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.029  1           1.014
## LastAuthorFemale  1.030  1           1.015
## Year              1.020 16           1.001
```



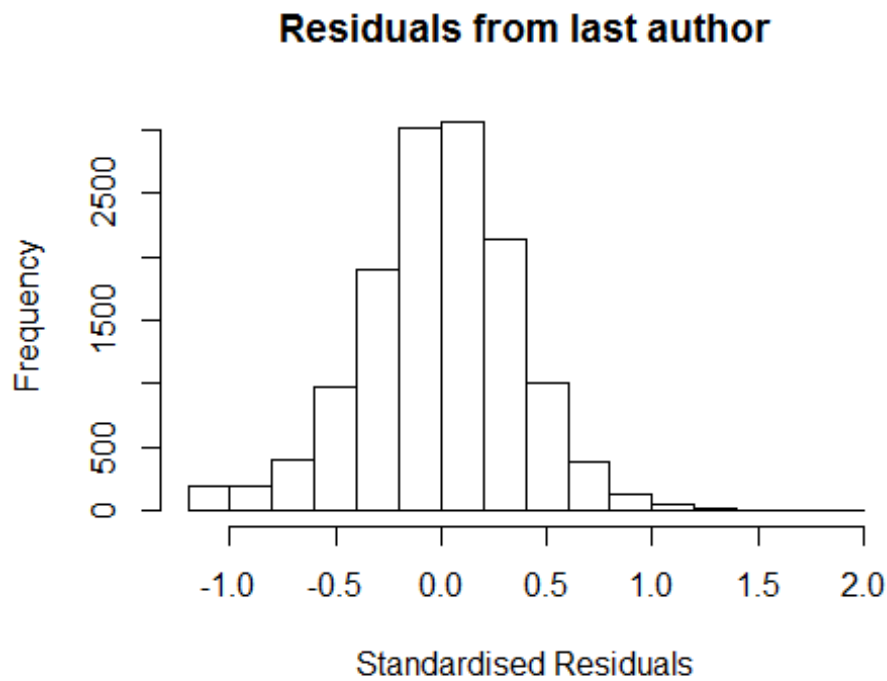

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.16147 -0.22659  0.00342  0.22587  1.80546
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)          1.11128    0.01440   77.18  < 2e-16 ***
## FirstAuthorFemale1   0.03035    0.00628    4.83  1.4e-06 ***
## LastAuthorFemale1    0.01983    0.00663    2.99   0.0028 **
## Year1997             0.00584    0.02096    0.28   0.7804
## Year1998            -0.03559    0.02116   -1.68   0.0926 .
## Year1999            -0.02242    0.02057   -1.09   0.2758
## Year2000            -0.01676    0.02050   -0.82   0.4136
## Year2001            -0.00663    0.02048   -0.32   0.7462
## Year2002            -0.03279    0.01965   -1.67   0.0952 .
## Year2003            -0.04271    0.02005   -2.13   0.0332 *
## Year2004            -0.02608    0.01904   -1.37   0.1708
## Year2005            -0.05797    0.01851   -3.13   0.0017 **
```

```
## Year2006           -0.03121    0.01807   -1.73   0.0841 .
## Year2007           -0.03741    0.01788   -2.09   0.0364 *
## Year2008           -0.04669    0.01795   -2.60   0.0093 **
## Year2009           -0.04950    0.01799   -2.75   0.0059 **
## Year2010           -0.02208    0.01796   -1.23   0.2188
## Year2011           -0.05010    0.01840   -2.72   0.0065 **
## Year2012           -0.04328    0.01893   -2.29   0.0223 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.338
## Multiple R-squared:  0.00507,    Adjusted R-squared:  0.00373
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  observation 12140 is an outlier with |weight| = 0 ( < 7.4e-06);
##  1123 weights are ~= 1. The remaining 12310 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0092  0.8640  0.9500  0.8940  0.9860  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          7.44e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##                   GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.01  1           1.005
## Year              1.01 16           1.000
```




```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.14895 -0.22664  0.00422  0.22544  1.79658
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)          1.11541    0.01432   77.90  < 2e-16 ***
## FirstAuthorFemale1   0.03354    0.00623    5.38  7.5e-08 ***
## Year1997             0.00585    0.02099    0.28   0.7806
## Year1998            -0.03581    0.02116   -1.69   0.0906 .
## Year1999            -0.02163    0.02055   -1.05   0.2927
## Year2000            -0.01691    0.02050   -0.83   0.4093
## Year2001            -0.00574    0.02047   -0.28   0.7793
## Year2002            -0.03246    0.01967   -1.65   0.0989 .
## Year2003            -0.04191    0.02005   -2.09   0.0366 *
## Year2004            -0.02518    0.01905   -1.32   0.1861
## Year2005            -0.05773    0.01849   -3.12   0.0018 **
## Year2006            -0.02987    0.01806   -1.65   0.0982 .
```

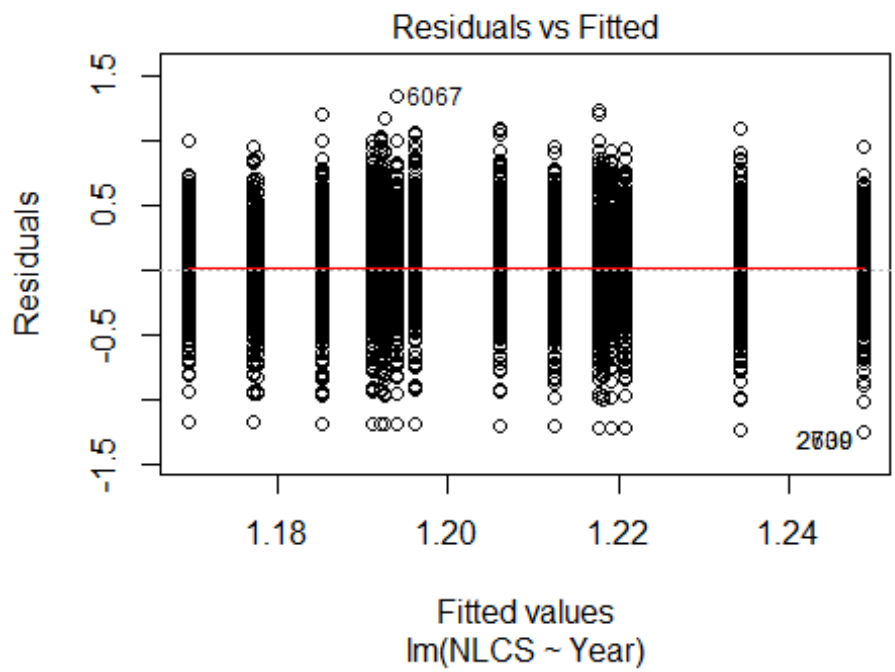
```
## Year2007           -0.03592    0.01786   -2.01   0.0444 *
## Year2008           -0.04565    0.01794   -2.54   0.0109 *
## Year2009           -0.04905    0.01799   -2.73   0.0064 **
## Year2010           -0.02046    0.01795   -1.14   0.2542
## Year2011           -0.04853    0.01837   -2.64   0.0083 **
## Year2012           -0.04192    0.01893   -2.21   0.0268 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.338
## Multiple R-squared:  0.00438,    Adjusted R-squared:  0.00312
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  observation 12140 is an outlier with |weight| = 0 ( < 7.4e-06);
##  1123 weights are ~= 1. The remaining 12310 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0071  0.8650  0.9510  0.8940  0.9860  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          7.44e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##                   GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.011  1           1.005
## Year             1.011 16           1.000
```



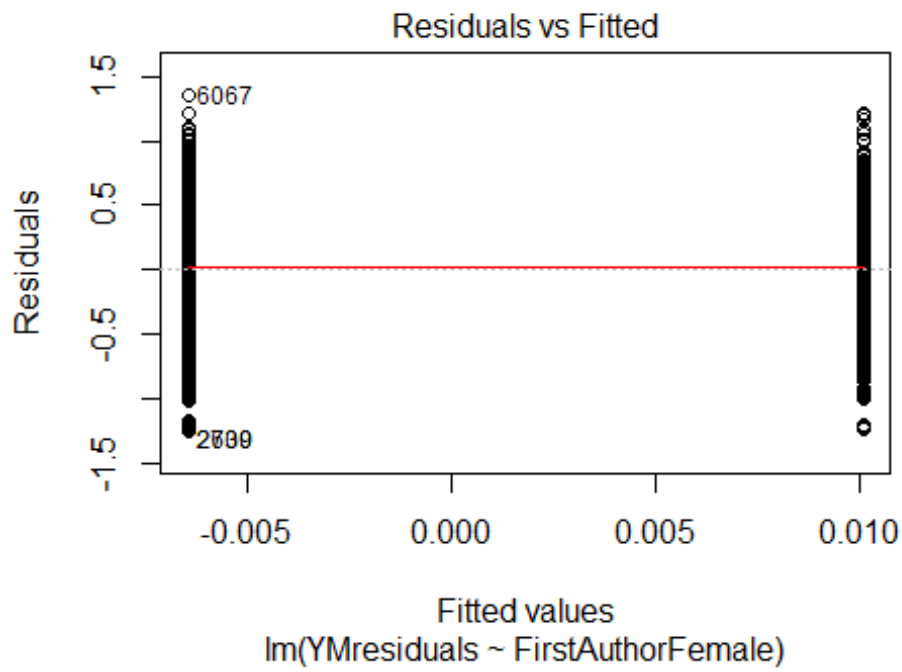

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.15241 -0.22821  0.00508  0.22559  1.82319
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.12010    0.01427   78.51  < 2e-16 ***
## LastAuthorFemale1   0.02530    0.00658    3.84  0.00012 ***
## Year1997            0.00701    0.02100    0.33  0.73874
## Year1998           -0.03502    0.02116   -1.65  0.09800 .
## Year1999           -0.02072    0.02055   -1.01  0.31340
## Year2000           -0.01565    0.02050   -0.76  0.44529
## Year2001           -0.00529    0.02049   -0.26  0.79634
## Year2002           -0.03103    0.01962   -1.58  0.11374
## Year2003           -0.04115    0.02006   -2.05  0.04025 *
## Year2004           -0.02377    0.01906   -1.25  0.21243
## Year2005           -0.05598    0.01856   -3.02  0.00256 **
## Year2006           -0.02864    0.01807   -1.59  0.11298
```

```
## Year2007          -0.03487    0.01790   -1.95  0.05145 .
## Year2008          -0.04389    0.01796   -2.44  0.01454 *
## Year2009          -0.04587    0.01796   -2.55  0.01067 *
## Year2010          -0.01869    0.01794   -1.04  0.29751
## Year2011          -0.04629    0.01837   -2.52  0.01174 *
## Year2012          -0.03834    0.01890   -2.03  0.04254 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.338
## Multiple R-squared:  0.00328,    Adjusted R-squared:  0.00201
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  observation 12140 is an outlier with |weight| = 0 ( < 7.4e-06);
##  1151 weights are ~= 1. The remaining 12282 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0067  0.8640  0.9510  0.8940  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          7.44e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Sample size for the above analysis:  13434"
## [1] ""
## [1] ""
## [1] "####################################"
## [1] "Analysis of AJSC 2803"
## [1] "####################################"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  538  494  467  503  473  472  506  449  497  539  613  684  736  689  752
## 2011 2012
##  763  798
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  432  372  355  394  360  376  418  356  400  442  494  561  604  552  615
```

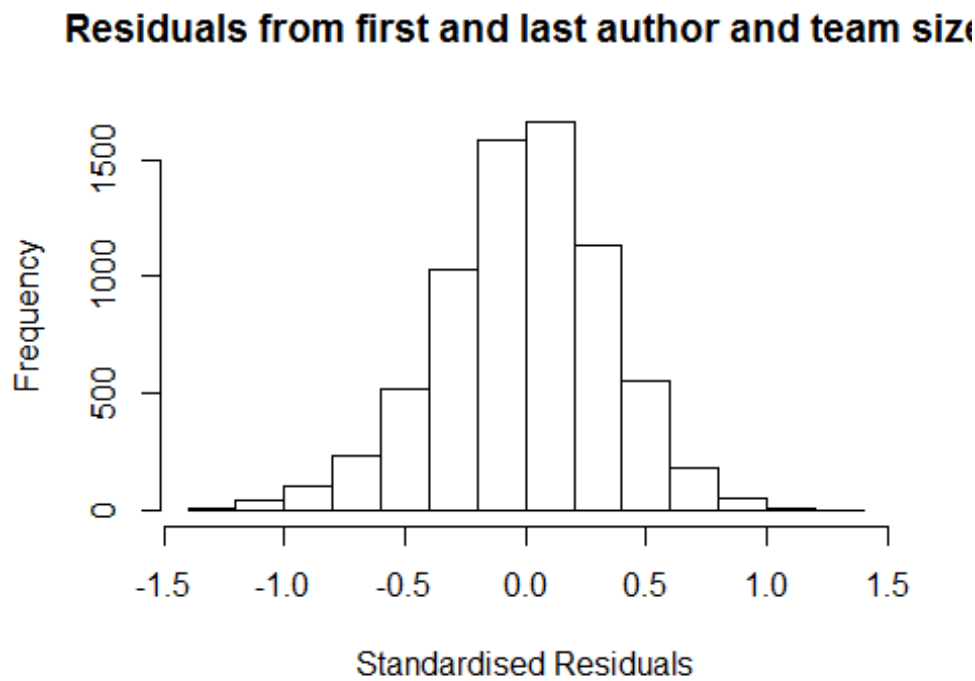
```
## 2011 2012
##  587  635
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  382  329  300  365  323  328  376  315  358  396  438  500  544  498  553
## 2011 2012
##  529  576
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
##  Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 48, df = 16, p-value = 5e-05
```




```
##
##  Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 17, df = 1, p-value = 5e-05
```

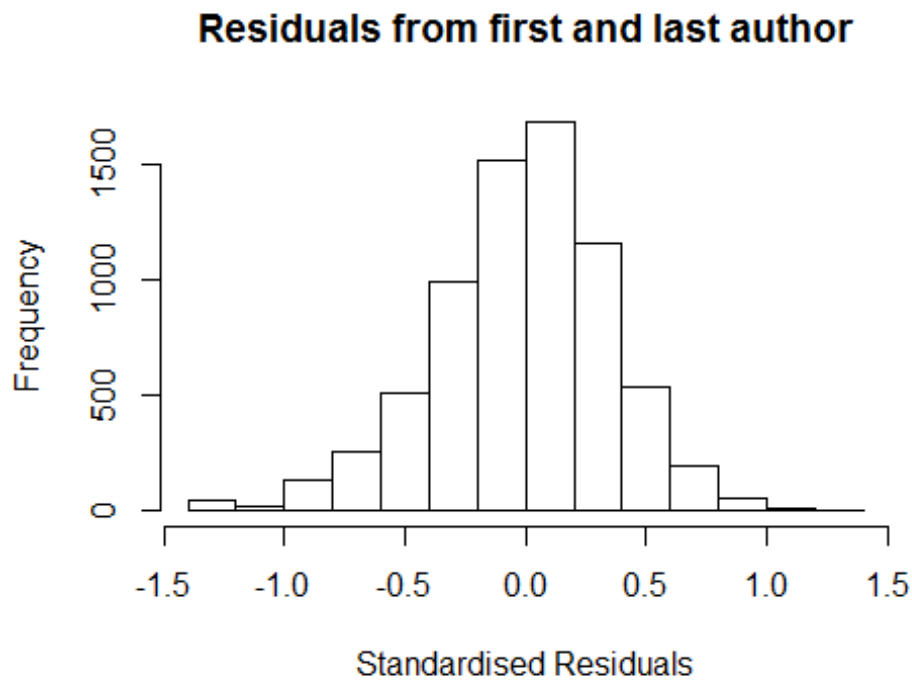



```
## [1] "Female first author team size 2018 geometric mean: 5.07839147143003"
## [1] "Male first author team size 2018 geometric mean: 4.88346081567113"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 51000, p-value = 0.6
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 4.85337633854117"
## [1] "Male last author team size 2018 geometric mean: 5.06975962959433"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 44000, p-value = 0.1
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.030  1           1.015
## LastAuthorFemale  1.021  1           1.010
## UniqueAuthors     1.082  4           1.010
## Year              1.099 16           1.003
```



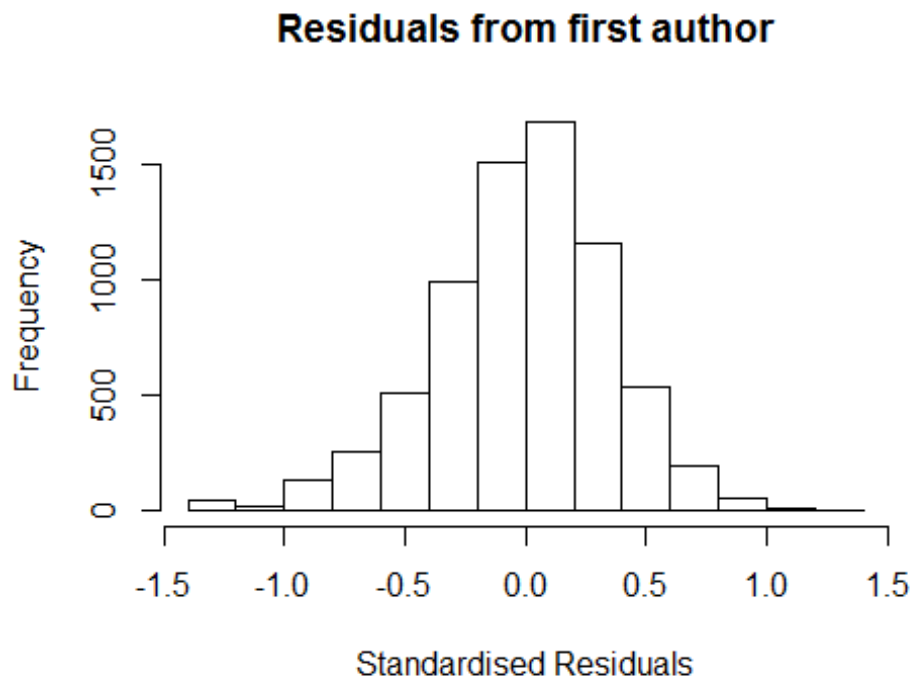

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
UniqueAuthors +
##     Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##     k.max = 1000))
##  \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.31695 -0.22714  0.00328  0.22353  1.27441
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.07182    0.03040   35.25   <2e-16 ***
## FirstAuthorFemale1  0.01867    0.00848    2.20    0.028 *
## LastAuthorFemale1   0.00844    0.00913    0.92    0.355
## UniqueAuthors2      0.06120    0.02728    2.24    0.025 *
## UniqueAuthors3      0.10670    0.02628    4.06    5e-05 ***
## UniqueAuthors4      0.14254    0.02653    5.37    8e-08 ***
## UniqueAuthors5      0.22282    0.02498    8.92   <2e-16 ***
## Year1997            0.00645    0.02959    0.22    0.828
## Year1998            0.01970    0.02863    0.69    0.491
## Year1999           -0.01315    0.02718   -0.48    0.629
```

```
## Year2000            0.02231    0.02726    0.82    0.413
## Year2001            0.01280    0.02742    0.47    0.641
## Year2002           -0.01111    0.02705   -0.41    0.681
## Year2003           -0.05216    0.02647   -1.97    0.049 *
## Year2004           -0.05277    0.02489   -2.12    0.034 *
## Year2005           -0.06606    0.02529   -2.61    0.009 **
## Year2006           -0.03005    0.02508   -1.20    0.231
## Year2007           -0.05959    0.02479   -2.40    0.016 *
## Year2008           -0.04222    0.02426   -1.74    0.082 .
## Year2009           -0.03585    0.02536   -1.41    0.158
## Year2010           -0.02609    0.02476   -1.05    0.292
## Year2011           -0.02699    0.02529   -1.07    0.286
## Year2012           -0.03732    0.02557   -1.46    0.144
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.332
## Multiple R-squared:  0.0463, Adjusted R-squared:  0.0434
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
##  653 weights are ~= 1. The remaining 6457 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0805  0.8660  0.9490  0.8970  0.9860  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.41e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.026  1           1.013
## LastAuthorFemale  1.010  1           1.005
## Year              1.025 16           1.001
```



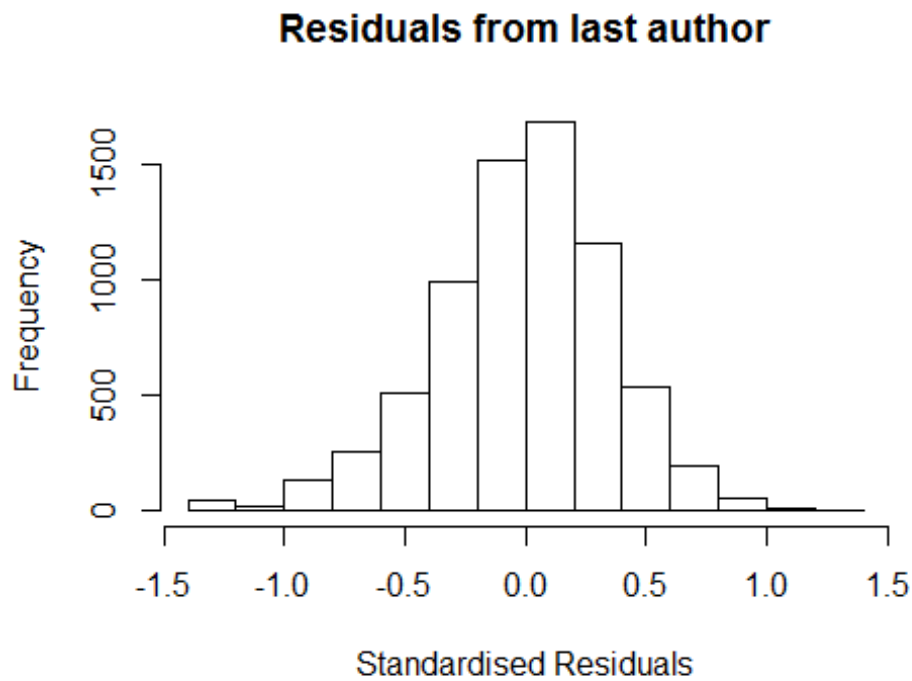

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.2506 -0.2285  0.0087  0.2261  1.3341
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.22e+00   1.97e-02   61.84   <2e-16 ***
## FirstAuthorFemale1  1.73e-02   8.63e-03    2.01    0.045 *
## LastAuthorFemale1  -6.01e-03   9.18e-03   -0.65    0.513
## Year1997            6.79e-03   2.92e-02    0.23    0.816
## Year1998            2.56e-02   2.81e-02    0.91    0.362
## Year1999           -2.00e-03   2.70e-02   -0.07    0.941
## Year2000            3.31e-02   2.73e-02    1.21    0.225
## Year2001            2.13e-02   2.75e-02    0.78    0.438
## Year2002            6.49e-05   2.71e-02    0.00    0.998
## Year2003           -3.78e-02   2.70e-02   -1.40    0.161
## Year2004           -3.08e-02   2.52e-02   -1.22    0.223
## Year2005           -4.15e-02   2.56e-02   -1.62    0.105
```

```
## Year2006           -1.26e-02   2.52e-02   -0.50    0.618
## Year2007           -2.85e-02   2.51e-02   -1.14    0.255
## Year2008           -1.72e-02   2.44e-02   -0.70    0.482
## Year2009           -8.43e-03   2.55e-02   -0.33    0.741
## Year2010           -7.16e-03   2.50e-02   -0.29    0.775
## Year2011           -1.23e-03   2.51e-02   -0.05    0.961
## Year2012           -1.55e-02   2.55e-02   -0.61    0.542
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.337
## Multiple R-squared:  0.00349,    Adjusted R-squared:  0.000957
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  603 weights are ~= 1. The remaining 6507 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0819  0.8660  0.9500  0.8960  0.9860  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.41e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.021  1           1.010
## Year              1.021 16           1.001
```




```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.24926 -0.22813  0.00881  0.22565  1.33566
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.216418   0.019584   62.11   <2e-16 ***
## FirstAuthorFemale1  0.016773   0.008609    1.95    0.051 .
## Year1997            0.006377   0.029179    0.22    0.827
## Year1998            0.025497   0.028098    0.91    0.364
## Year1999           -0.002355   0.026995   -0.09    0.930
## Year2000            0.032840   0.027250    1.21    0.228
## Year2001            0.021041   0.027455    0.77    0.443
## Year2002           -0.000408   0.027076   -0.02    0.988
## Year2003           -0.038183   0.026956   -1.42    0.157
## Year2004           -0.031272   0.025214   -1.24    0.215
## Year2005           -0.041999   0.025571   -1.64    0.101
## Year2006           -0.013074   0.025235   -0.52    0.604
```

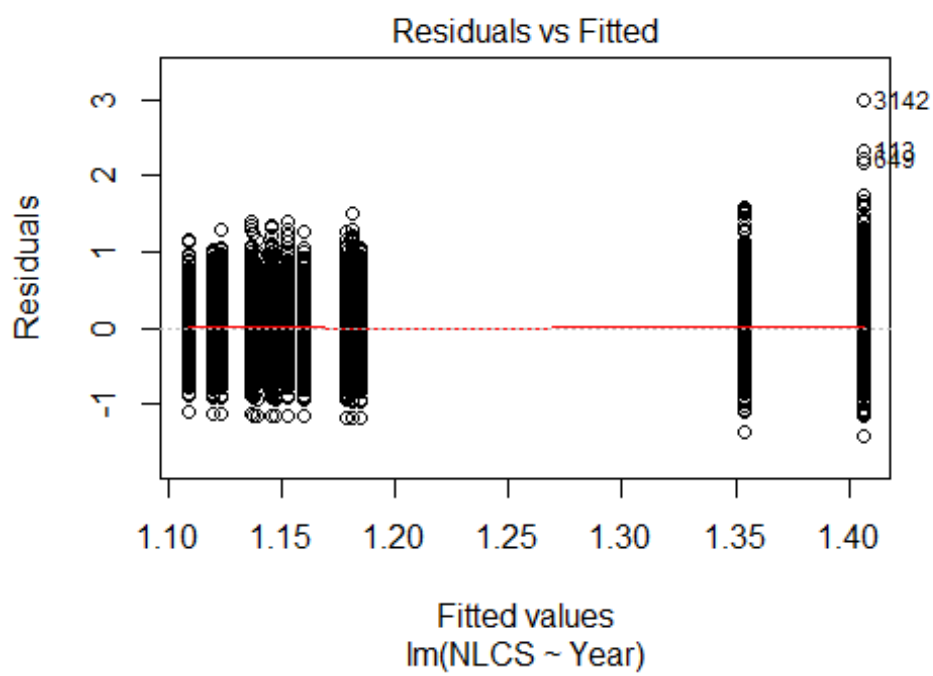
```
## Year2007           -0.028956   0.025035   -1.16    0.247
## Year2008           -0.017515   0.024411   -0.72    0.473
## Year2009           -0.008921   0.025487   -0.35    0.726
## Year2010           -0.007782   0.024999   -0.31    0.756
## Year2011           -0.001794   0.025112   -0.07    0.943
## Year2012           -0.016061   0.025494   -0.63    0.529
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.337
## Multiple R-squared:  0.00343,    Adjusted R-squared:  0.00104
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  599 weights are ~= 1. The remaining 6511 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0809  0.8660  0.9510  0.8970  0.9860  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.41e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##                   GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.005  1           1.002
## Year             1.005 16           1.000
```



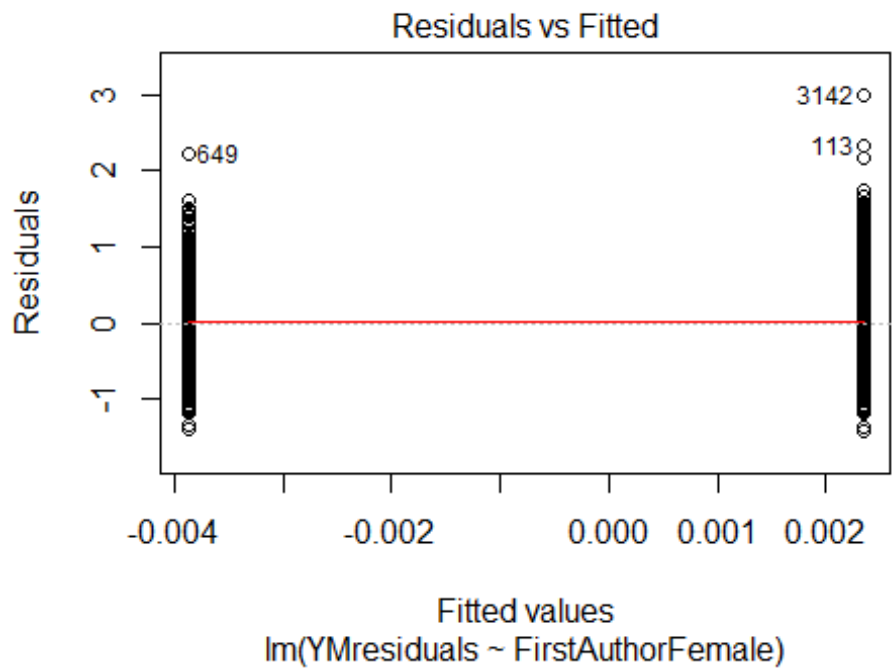

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.25627 -0.22718  0.00928  0.22484  1.32653
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.22224    0.01956   62.49   <2e-16 ***
## LastAuthorFemale1 -0.00422    0.00916   -0.46     0.64
## Year1997            0.00712    0.02925    0.24     0.81
## Year1998            0.02536    0.02810    0.90     0.37
## Year1999           -0.00151    0.02703   -0.06     0.96
## Year2000            0.03403    0.02726    1.25     0.21
## Year2001            0.02211    0.02747    0.80     0.42
## Year2002            0.00172    0.02709    0.06     0.95
## Year2003           -0.03609    0.02695   -1.34     0.18
## Year2004           -0.02861    0.02526   -1.13     0.26
## Year2005           -0.04002    0.02562   -1.56     0.12
## Year2006           -0.00977    0.02522   -0.39     0.70
```

```
## Year2007          -0.02702    0.02508   -1.08     0.28
## Year2008          -0.01500    0.02441   -0.61     0.54
## Year2009          -0.00606    0.02549   -0.24     0.81
## Year2010          -0.00453    0.02499   -0.18     0.86
## Year2011           0.00134    0.02509    0.05     0.96
## Year2012          -0.01233    0.02544   -0.48     0.63
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.337
## Multiple R-squared:  0.00291,    Adjusted R-squared:  0.000515
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  588 weights are ~= 1. The remaining 6522 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0865  0.8670  0.9500  0.8970  0.9860  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.41e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Sample size for the above analysis:  7110"
## [1] ""
## [1] ""
## [1] "####################################"
## [1] "Analysis of AJSC 2804"
## [1] "####################################"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 3055 2416 1542 1478 1590 1538 1532 1338 1387 1555 1596 1697 1798 2050 2067
## 2011 2012
## 2322 2227
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 1113 1026  937  938  895  777  959  862  855 1014 1049 1126 1164 1320 1344
## 2011 2012
```

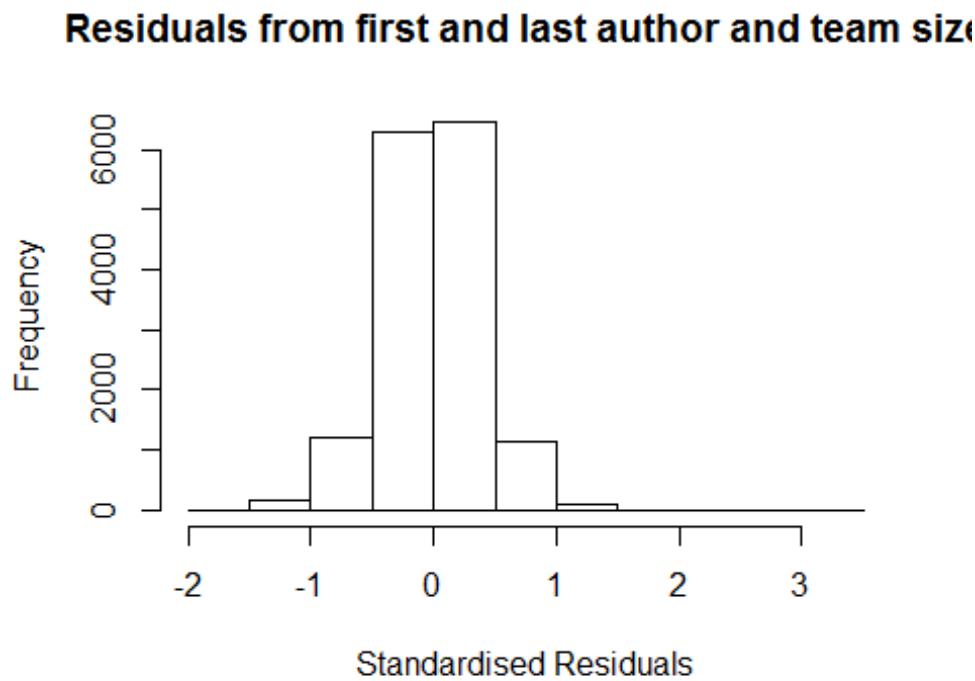
```
## 1513 1512
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  923  848  796  782  751  649  804  691  702  851  879  930  964 1100 1127
## 2011 2012
## 1288 1283
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
##  Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 1100, df = 16, p-value <2e-16
```




```
##
##  Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 130, df = 1, p-value <2e-16
```

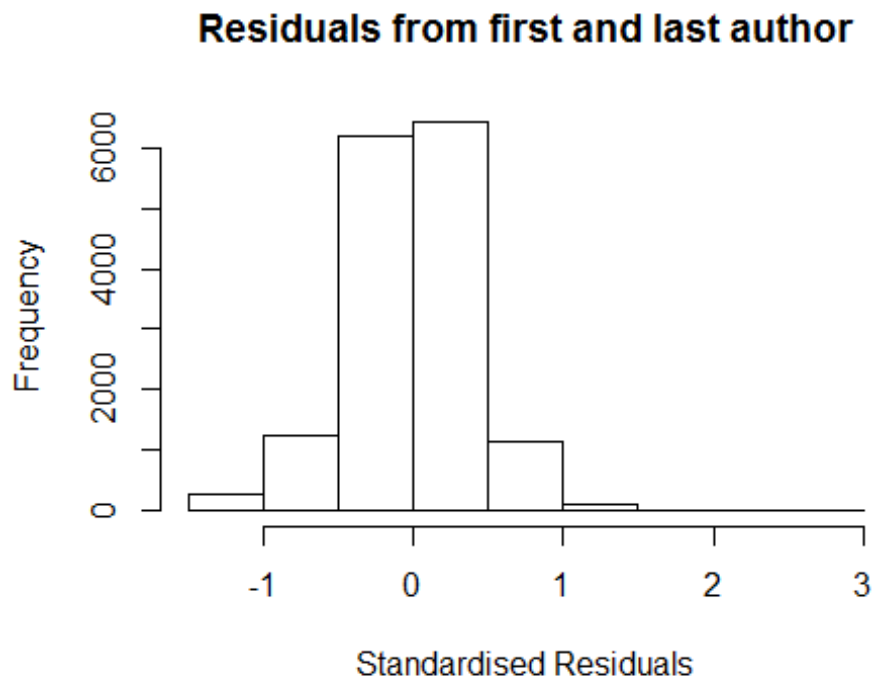



```
## [1] "Female first author team size 2018 geometric mean: 5.60830190167534"
## [1] "Male first author team size 2018 geometric mean: 4.95949821344186"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 170000, p-value = 0.001
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 5.293484486858"
## [1] "Male last author team size 2018 geometric mean: 5.21393304538431"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 130000, p-value = 0.9
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.019  1           1.009
## LastAuthorFemale  1.014  1           1.007
## UniqueAuthors     1.048  4           1.006
## Year              1.062 16           1.002
```



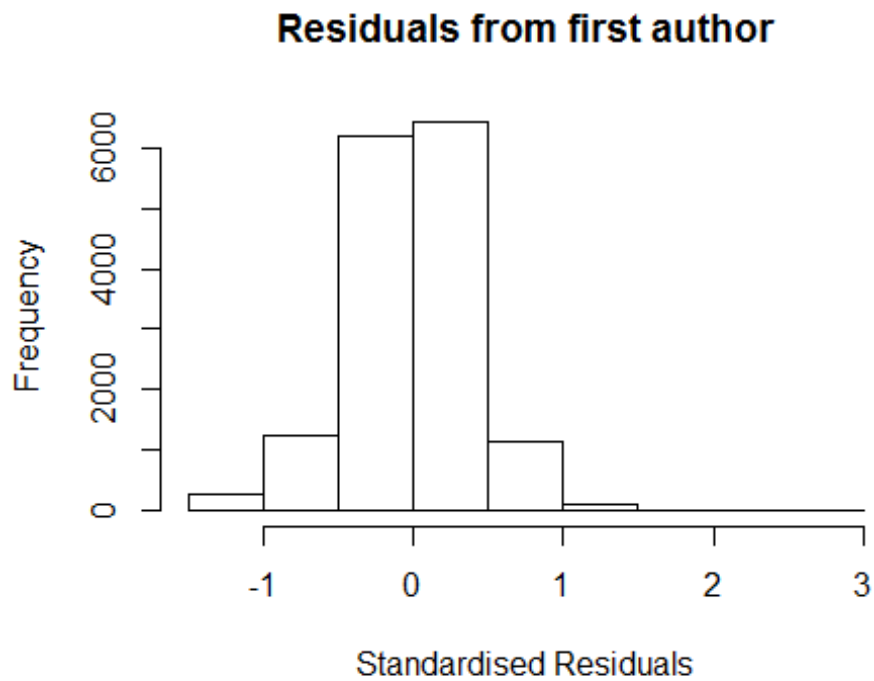

```
## [1] "List of  1 outliers with residuals above  2.5"
##        ScopusId  NLCS Year OneField Fields residuals
## 3142 0029980777 4.398 1996     2731      3     3.167
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
UniqueAuthors +
##     Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##     k.max = 1000))
##  \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.51765 -0.23228  0.00222  0.23310  3.16713
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.23087    0.03116   39.50  < 2e-16 ***
## FirstAuthorFemale1 -0.01403    0.00603   -2.33    0.020 *
## LastAuthorFemale1  -0.00452    0.00682   -0.66    0.508
## UniqueAuthors2      0.12431    0.02505    4.96  7.1e-07 ***
## UniqueAuthors3      0.17584    0.02487    7.07  1.6e-12 ***
## UniqueAuthors4      0.20991    0.02493    8.42  < 2e-16 ***
## UniqueAuthors5      0.28678    0.02445   11.73  < 2e-16 ***
## Year1997           -0.05786    0.02691   -2.15    0.032 *
## Year1998           -0.24193    0.02420  -10.00  < 2e-16 ***
## Year1999           -0.31465    0.02379  -13.23  < 2e-16 ***
```

```
## Year2000           -0.25112    0.02426  -10.35  < 2e-16 ***
## Year2001           -0.30960    0.02405  -12.87  < 2e-16 ***
## Year2002           -0.28380    0.02376  -11.94  < 2e-16 ***
## Year2003           -0.30508    0.02364  -12.91  < 2e-16 ***
## Year2004           -0.31826    0.02412  -13.19  < 2e-16 ***
## Year2005           -0.32592    0.02349  -13.87  < 2e-16 ***
## Year2006           -0.31597    0.02361  -13.38  < 2e-16 ***
## Year2007           -0.30233    0.02316  -13.06  < 2e-16 ***
## Year2008           -0.27945    0.02343  -11.93  < 2e-16 ***
## Year2009           -0.28717    0.02318  -12.39  < 2e-16 ***
## Year2010           -0.25981    0.02305  -11.27  < 2e-16 ***
## Year2011           -0.28795    0.02304  -12.50  < 2e-16 ***
## Year2012           -0.30363    0.02326  -13.05  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.349
## Multiple R-squared:  0.0847, Adjusted R-squared:  0.0834
## Convergence in 15 IRWLS iterations
##
## Robustness weights:
##  5 observations c(46,254,798,817,1459)
##   are outliers with |weight| = 0 ( < 6.5e-06);
##  1305 weights are ~= 1. The remaining 14058 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0032  0.8650  0.9510  0.8920  0.9860  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          6.51e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.018  1           1.009
## LastAuthorFemale  1.012  1           1.006
## Year              1.018 16           1.001
```



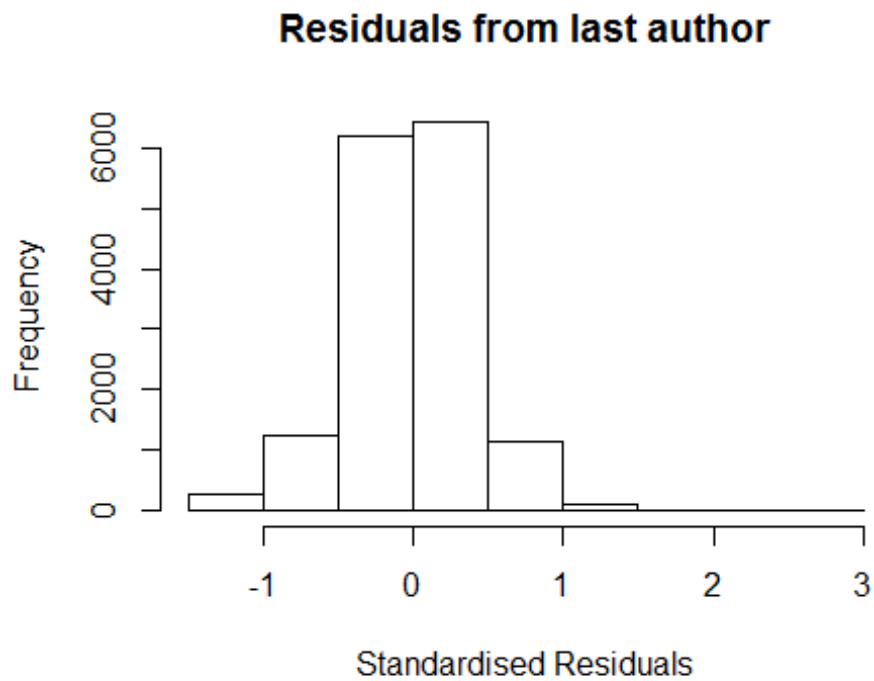

```
## [1] "List of  1 outliers with residuals above  2.5"
##        ScopusId  NLCS Year OneField Fields residuals
## 3142 0029980777 4.398 1996     2731      3     2.975
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.42310 -0.23887  0.00094  0.23542  2.97490
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.42310    0.02089   68.11   <2e-16 ***
## FirstAuthorFemale1 -0.00838    0.00614   -1.37    0.172
## LastAuthorFemale1  -0.01035    0.00689   -1.50    0.133
## Year1997           -0.05261    0.02712   -1.94    0.052 .
## Year1998           -0.23683    0.02443   -9.69   <2e-16 ***
## Year1999           -0.30312    0.02406  -12.60   <2e-16 ***
## Year2000           -0.24305    0.02451   -9.92   <2e-16 ***
## Year2001           -0.29558    0.02423  -12.20   <2e-16 ***
## Year2002           -0.26753    0.02411  -11.10   <2e-16 ***
## Year2003           -0.28743    0.02394  -12.00   <2e-16 ***
## Year2004           -0.28799    0.02431  -11.85   <2e-16 ***
## Year2005           -0.30004    0.02380  -12.60   <2e-16 ***
```

```
## Year2006           -0.28840    0.02390  -12.07   <2e-16 ***
## Year2007           -0.27865    0.02331  -11.96   <2e-16 ***
## Year2008           -0.24587    0.02368  -10.38   <2e-16 ***
## Year2009           -0.25920    0.02342  -11.07   <2e-16 ***
## Year2010           -0.23185    0.02322   -9.98   <2e-16 ***
## Year2011           -0.25810    0.02320  -11.13   <2e-16 ***
## Year2012           -0.27532    0.02348  -11.73   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.355
## Multiple R-squared:  0.0438, Adjusted R-squared:  0.0427
## Convergence in 15 IRWLS iterations
##
## Robustness weights:
##  4 observations c(46,254,798,817)
##   are outliers with |weight| = 0 ( < 6.5e-06);
##  1316 weights are ~= 1. The remaining 14048 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0004  0.8650  0.9510  0.8910  0.9860  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          6.51e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.013  1           1.006
## Year              1.013 16           1.000
```




```
## [1] "List of  1 outliers with residuals above  2.5"
##        ScopusId  NLCS Year OneField Fields residuals
## 3142 0029980777 4.398 1996     2731      3     2.975
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##       Min        1Q    Median        3Q       Max
## -1.421189 -0.239353  0.000363  0.235237  2.976811
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.42119    0.02087   68.11   <2e-16 ***
## FirstAuthorFemale1 -0.00933    0.00612   -1.52    0.127
## Year1997           -0.05255    0.02714   -1.94    0.053 .
## Year1998           -0.23686    0.02445   -9.69   <2e-16 ***
## Year1999           -0.30313    0.02407  -12.59   <2e-16 ***
## Year2000           -0.24300    0.02452   -9.91   <2e-16 ***
## Year2001           -0.29582    0.02424  -12.20   <2e-16 ***
## Year2002           -0.26785    0.02413  -11.10   <2e-16 ***
## Year2003           -0.28775    0.02396  -12.01   <2e-16 ***
## Year2004           -0.28843    0.02432  -11.86   <2e-16 ***
## Year2005           -0.30021    0.02382  -12.60   <2e-16 ***
## Year2006           -0.28855    0.02392  -12.07   <2e-16 ***
```

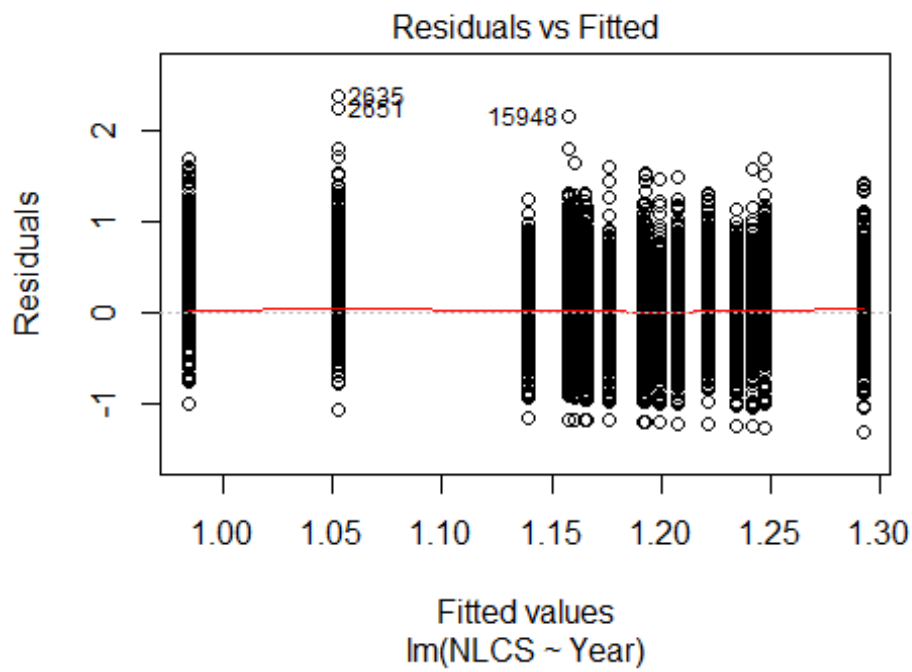
```
## Year2007           -0.27898    0.02332  -11.96   <2e-16 ***
## Year2008           -0.24658    0.02369  -10.41   <2e-16 ***
## Year2009           -0.25967    0.02343  -11.08   <2e-16 ***
## Year2010           -0.23218    0.02324   -9.99   <2e-16 ***
## Year2011           -0.25845    0.02321  -11.13   <2e-16 ***
## Year2012           -0.27571    0.02349  -11.74   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.355
## Multiple R-squared:  0.0436, Adjusted R-squared:  0.0426
## Convergence in 15 IRWLS iterations
##
## Robustness weights:
##  4 observations c(46,254,798,817)
##   are outliers with |weight| = 0 ( < 6.5e-06);
##  1337 weights are ~= 1. The remaining 14027 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0003  0.8650  0.9500  0.8910  0.9860  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          6.51e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##                   GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.007  1           1.003
## Year             1.007 16           1.000
```



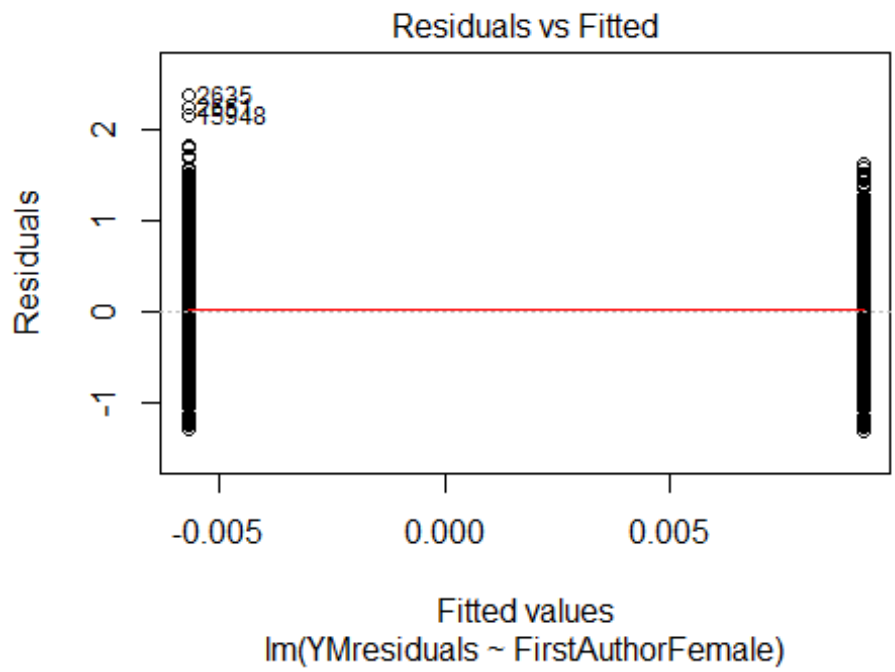

```
## [1] "List of  1 outliers with residuals above  2.5"
##        ScopusId  NLCS Year OneField Fields residuals
## 3142 0029980777 4.398 1996     2731      3     2.975
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##       Min        1Q    Median        3Q       Max
## -1.419938 -0.239496  0.000326  0.234950  2.978062
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.41994    0.02067   68.70   <2e-16 ***
## LastAuthorFemale1 -0.01134    0.00688   -1.65    0.099 .
## Year1997          -0.05266    0.02712   -1.94    0.052 .
## Year1998          -0.23650    0.02441   -9.69   <2e-16 ***
## Year1999          -0.30269    0.02403  -12.59   <2e-16 ***
## Year2000          -0.24274    0.02449   -9.91   <2e-16 ***
## Year2001          -0.29517    0.02421  -12.19   <2e-16 ***
## Year2002          -0.26726    0.02409  -11.09   <2e-16 ***
## Year2003          -0.28721    0.02393  -12.00   <2e-16 ***
## Year2004          -0.28776    0.02429  -11.85   <2e-16 ***
## Year2005          -0.29989    0.02379  -12.61   <2e-16 ***
## Year2006          -0.28826    0.02387  -12.07   <2e-16 ***
```

```
## Year2007          -0.27862    0.02330  -11.96   <2e-16 ***
## Year2008          -0.24603    0.02367  -10.39   <2e-16 ***
## Year2009          -0.25938    0.02341  -11.08   <2e-16 ***
## Year2010          -0.23191    0.02321   -9.99   <2e-16 ***
## Year2011          -0.25854    0.02319  -11.15   <2e-16 ***
## Year2012          -0.27519    0.02346  -11.73   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.355
## Multiple R-squared:  0.0436, Adjusted R-squared:  0.0425
## Convergence in 15 IRWLS iterations
##
## Robustness weights:
##  4 observations c(46,254,798,817)
##   are outliers with |weight| = 0 ( < 6.5e-06);
##  1334 weights are ~= 1. The remaining 14030 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0003  0.8650  0.9500  0.8910  0.9860  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          6.51e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                  psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Sample size for the above analysis:  15368"
## [1] ""
## [1] ""
## [1] "##################################"
## [1] "Analysis of AJSC 2805"
## [1] "##################################"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  472  595  712  624  763  630  601  665  730  835  879  993 1057 1298 1358
## 2011 2012
## 1582 1553
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
```

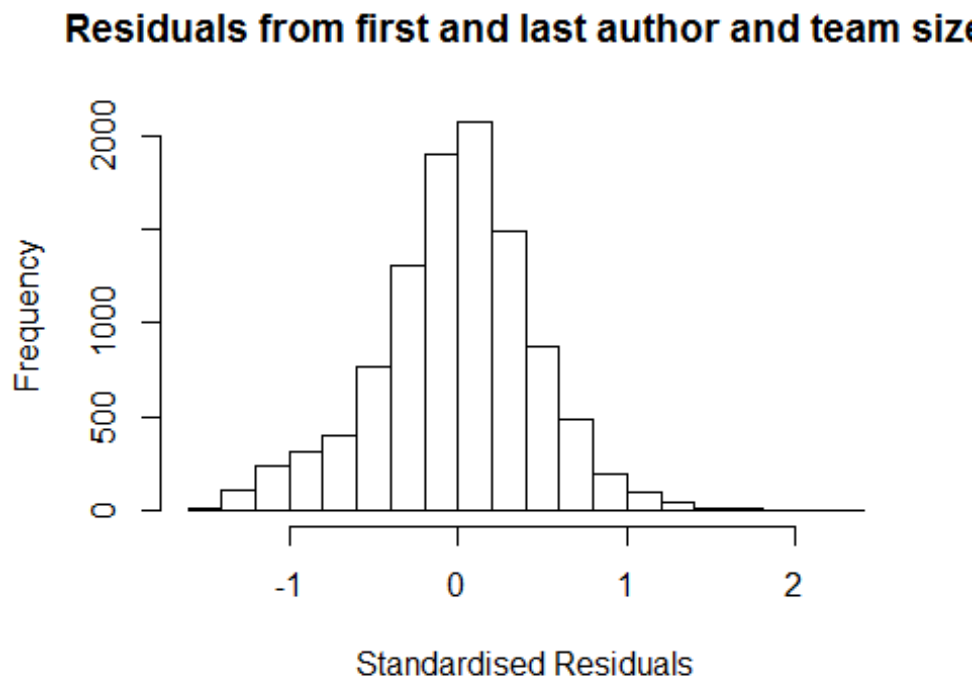
```
##  373  360  364  435  519  430  464  516  585  654  671  791  828 1002 1071
## 2011 2012
## 1237 1215
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  341  340  336  396  464  390  418  464  522  571  591  710  741  893  957
## 2011 2012
## 1090 1089
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
##  Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 820, df = 16, p-value <2e-16
```




```
##
##  Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 23, df = 1, p-value = 2e-06
```

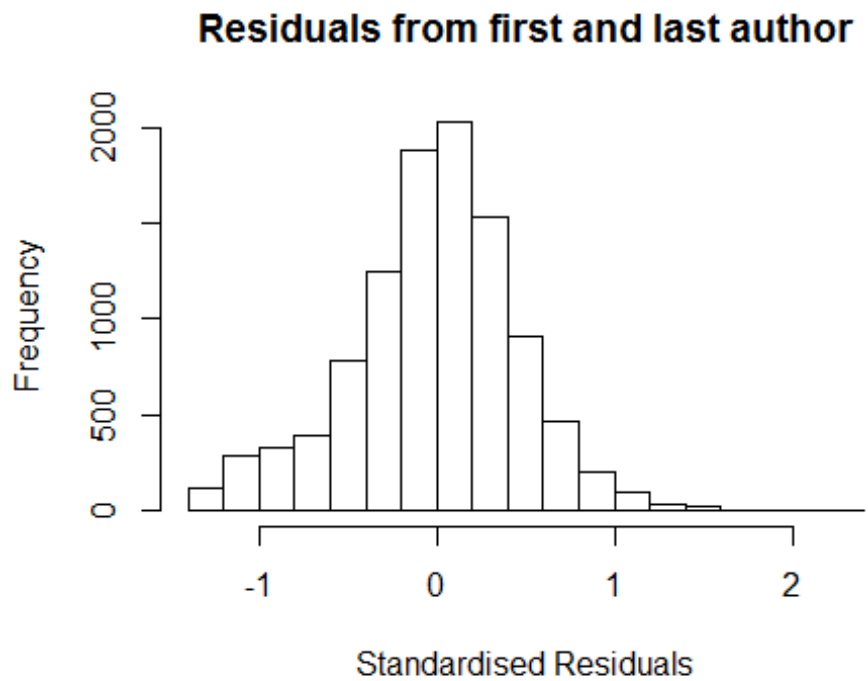



```
## [1] "Female first author team size 2018 geometric mean: 3.87902434238598"
## [1] "Male first author team size 2018 geometric mean: 3.97825210917023"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 160000, p-value = 0.5
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 3.75054375482607"
## [1] "Male last author team size 2018 geometric mean: 4.05359419851693"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 150000, p-value = 0.1
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.045  1           1.022
## LastAuthorFemale  1.055  1           1.027
## UniqueAuthors     1.240  4           1.027
## Year              1.262 16           1.007
```



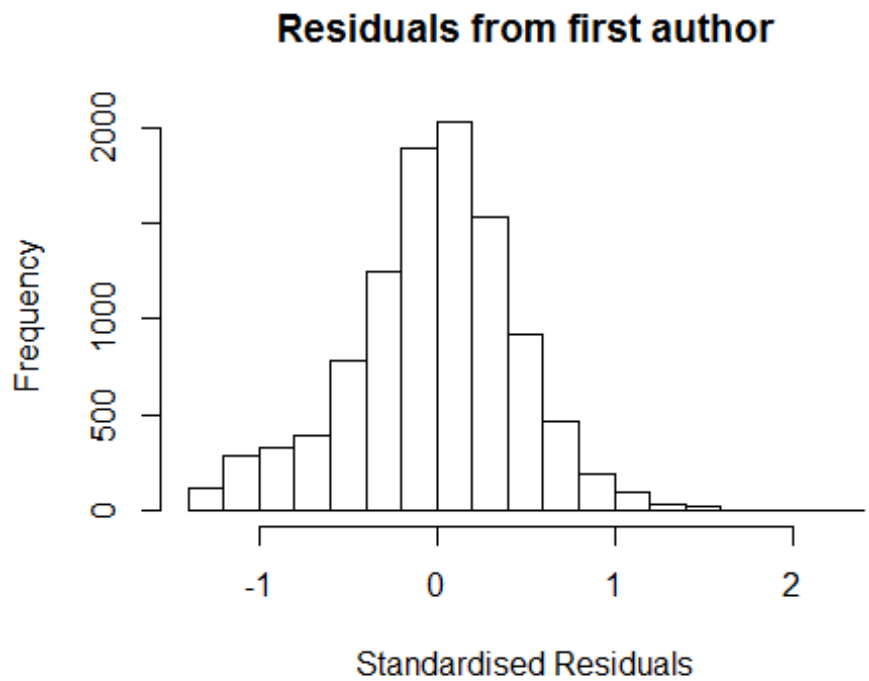

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
UniqueAuthors +
##     Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##     k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.4409 -0.2717  0.0104  0.2738  2.2557
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.06901    0.03363   31.79  < 2e-16 ***
## FirstAuthorFemale1  0.00896    0.00922    0.97  0.33093
## LastAuthorFemale1   0.00713    0.00980    0.73  0.46669
## UniqueAuthors2      0.18325    0.01972    9.29  < 2e-16 ***
## UniqueAuthors3      0.20844    0.02024   10.30  < 2e-16 ***
## UniqueAuthors4      0.22207    0.02121   10.47  < 2e-16 ***
## UniqueAuthors5      0.25844    0.02014   12.83  < 2e-16 ***
## Year1997            0.08948    0.04407    2.03  0.04231 *
## Year1998            0.09739    0.05065    1.92  0.05452 .
## Year1999           -0.10513    0.06354   -1.65  0.09806 .
```

```
## Year2000           -0.26740    0.06552   -4.08  4.5e-05 ***
## Year2001           -0.00261    0.03764   -0.07  0.94471
## Year2002           -0.01696    0.03544   -0.48  0.63224
## Year2003           -0.07849    0.03585   -2.19  0.02860 *
## Year2004           -0.05677    0.03482   -1.63  0.10299
## Year2005           -0.06280    0.03462   -1.81  0.06968 .
## Year2006           -0.08490    0.03417   -2.48  0.01298 *
## Year2007           -0.10705    0.03413   -3.14  0.00171 **
## Year2008           -0.06026    0.03288   -1.83  0.06691 .
## Year2009           -0.06984    0.03295   -2.12  0.03408 *
## Year2010           -0.09693    0.03298   -2.94  0.00330 **
## Year2011           -0.10187    0.03289   -3.10  0.00196 **
## Year2012           -0.11863    0.03316   -3.58  0.00035 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.408
## Multiple R-squared:  0.048,  Adjusted R-squared:  0.0459
## Convergence in 25 IRWLS iterations
##
## Robustness weights:
##  3 observations c(1296,1307,9284)
##   are outliers with |weight| = 0 ( < 9.7e-06);
##  870 weights are ~= 1. The remaining 9440 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0143  0.8510  0.9500  0.8830  0.9860  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          9.70e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.037  1           1.018
## LastAuthorFemale  1.051  1           1.025
## Year              1.030 16           1.001
```



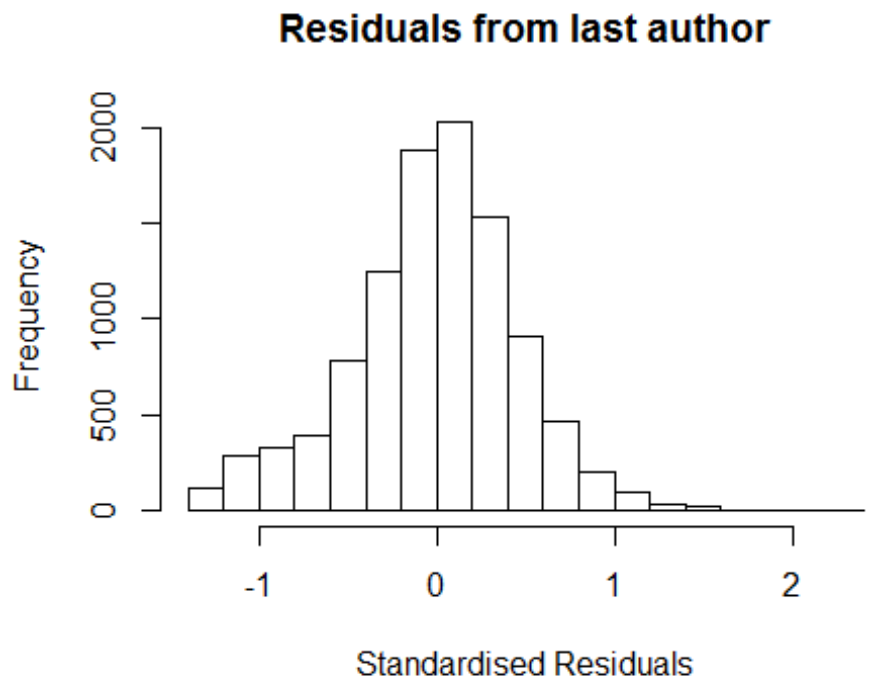

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.3429 -0.2797  0.0129  0.2791  2.3196
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)          1.22417    0.03020   40.54   <2e-16 ***
## FirstAuthorFemale1   0.02021    0.00929    2.18    0.030 *
## LastAuthorFemale1    0.00276    0.00985    0.28    0.780
## Year1997             0.09579    0.04364    2.19    0.028 *
## Year1998             0.09003    0.05139    1.75    0.080 .
## Year1999            -0.11576    0.06081   -1.90    0.057 .
## Year2000            -0.25950    0.06210   -4.18    3e-05 ***
## Year2001             0.01647    0.03759    0.44    0.661
## Year2002             0.01251    0.03508    0.36    0.721
## Year2003            -0.04598    0.03550   -1.30    0.195
## Year2004            -0.01885    0.03474   -0.54    0.587
## Year2005            -0.02098    0.03432   -0.61    0.541
```

```
## Year2006           -0.04915    0.03395   -1.45    0.148
## Year2007           -0.06325    0.03377   -1.87    0.061 .
## Year2008           -0.02480    0.03274   -0.76    0.449
## Year2009           -0.03032    0.03262   -0.93    0.353
## Year2010           -0.05421    0.03272   -1.66    0.098 .
## Year2011           -0.05645    0.03263   -1.73    0.084 .
## Year2012           -0.06853    0.03282   -2.09    0.037 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.413
## Multiple R-squared:  0.0188, Adjusted R-squared:  0.0171
## Convergence in 23 IRWLS iterations
##
## Robustness weights:
##  3 observations c(1296,1307,9284)
##   are outliers with |weight| = 0 ( < 9.7e-06);
##  845 weights are ~= 1. The remaining 9465 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0149  0.8540  0.9490  0.8830  0.9860  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          9.70e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.008  1           1.004
## Year              1.008 16           1.000
```




```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.3415 -0.2803  0.0131  0.2791  2.3189
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)          1.22485    0.03001   40.82  < 2e-16 ***
## FirstAuthorFemale1   0.02083    0.00918    2.27    0.023 *
## Year1997             0.09582    0.04363    2.20    0.028 *
## Year1998             0.09005    0.05138    1.75    0.080 .
## Year1999            -0.11577    0.06082   -1.90    0.057 .
## Year2000            -0.25976    0.06208   -4.18  2.9e-05 ***
## Year2001             0.01655    0.03758    0.44    0.660
## Year2002             0.01248    0.03508    0.36    0.722
## Year2003            -0.04602    0.03550   -1.30    0.195
## Year2004            -0.01898    0.03471   -0.55    0.584
## Year2005            -0.02121    0.03428   -0.62    0.536
## Year2006            -0.04919    0.03394   -1.45    0.147
```

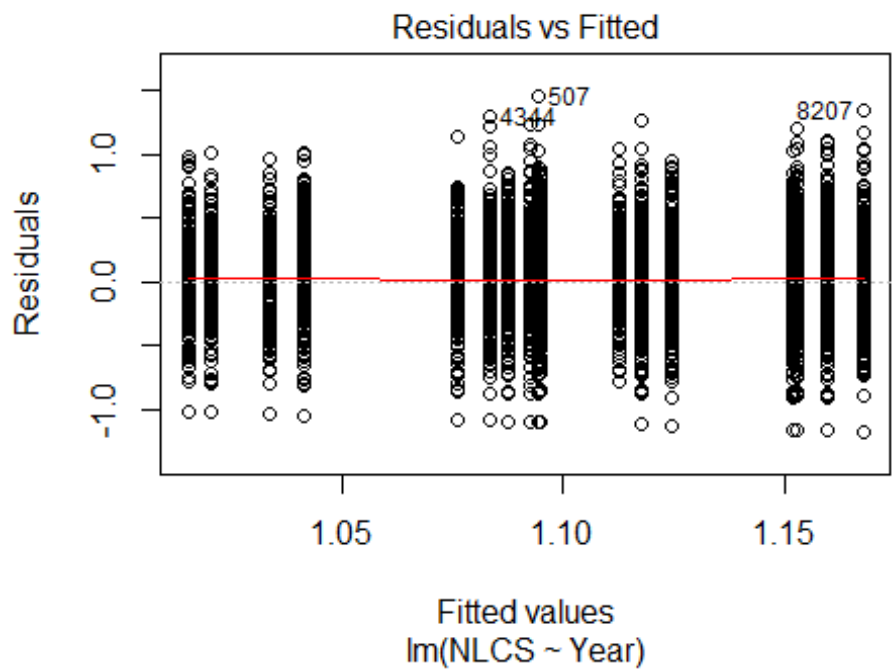
```
## Year2007           -0.06341    0.03374   -1.88    0.060 .
## Year2008           -0.02486    0.03273   -0.76    0.448
## Year2009           -0.03052    0.03258   -0.94    0.349
## Year2010           -0.05433    0.03270   -1.66    0.097 .
## Year2011           -0.05650    0.03262   -1.73    0.083 .
## Year2012           -0.06859    0.03282   -2.09    0.037 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.413
## Multiple R-squared:  0.0188, Adjusted R-squared:  0.0172
## Convergence in 23 IRWLS iterations
##
## Robustness weights:
##  3 observations c(1296,1307,9284)
##   are outliers with |weight| = 0 ( < 9.7e-06);
##  847 weights are ~= 1. The remaining 9463 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.015   0.854   0.949   0.883   0.986   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          9.70e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##                   GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.021  1           1.011
## Year             1.021 16           1.001
```



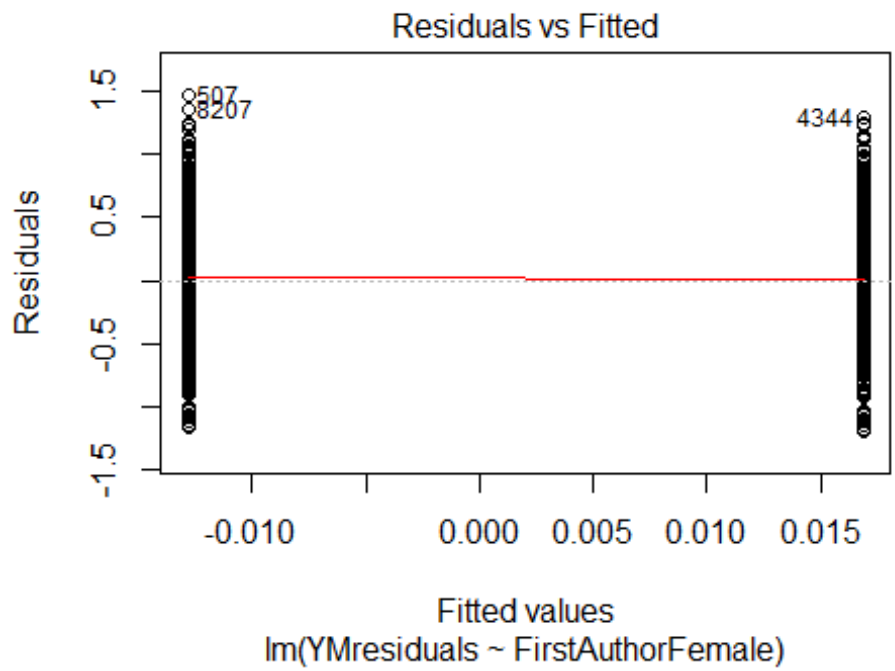

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##    Min     1Q Median     3Q    Max
## -1.333 -0.280  0.013  0.281  2.315
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.22899    0.03002   40.94  < 2e-16 ***
## LastAuthorFemale1   0.00795    0.00975    0.82    0.415
## Year1997            0.09609    0.04360    2.20    0.028 *
## Year1998            0.08963    0.05137    1.74    0.081 .
## Year1999           -0.11559    0.06064   -1.91    0.057 .
## Year2000           -0.25779    0.06200   -4.16  3.2e-05 ***
## Year2001            0.01673    0.03749    0.45    0.656
## Year2002            0.01382    0.03497    0.40    0.693
## Year2003           -0.04495    0.03543   -1.27    0.205
## Year2004           -0.01699    0.03465   -0.49    0.624
## Year2005           -0.01912    0.03423   -0.56    0.576
## Year2006           -0.04770    0.03386   -1.41    0.159
```

```
## Year2007          -0.06177    0.03368   -1.83    0.067 .
## Year2008          -0.02250    0.03263   -0.69    0.490
## Year2009          -0.02854    0.03252   -0.88    0.380
## Year2010          -0.05224    0.03260   -1.60    0.109
## Year2011          -0.05442    0.03253   -1.67    0.094 .
## Year2012          -0.06689    0.03274   -2.04    0.041 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.414
## Multiple R-squared:  0.0183, Adjusted R-squared:  0.0166
## Convergence in 23 IRWLS iterations
##
## Robustness weights:
##  3 observations c(1296,1307,9284)
##   are outliers with |weight| = 0 ( < 9.7e-06);
##  850 weights are ~= 1. The remaining 9460 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0168  0.8550  0.9490  0.8830  0.9860  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          9.70e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Sample size for the above analysis:  10313"
## [1] ""
## [1] ""
## [1] "##################################"
## [1] "Analysis of AJSC 2806"
## [1] "##################################"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  492  504  450  459  423  463  436  391  343  367  393  512  473  448  449
## 2011 2012
##  531  523
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
```

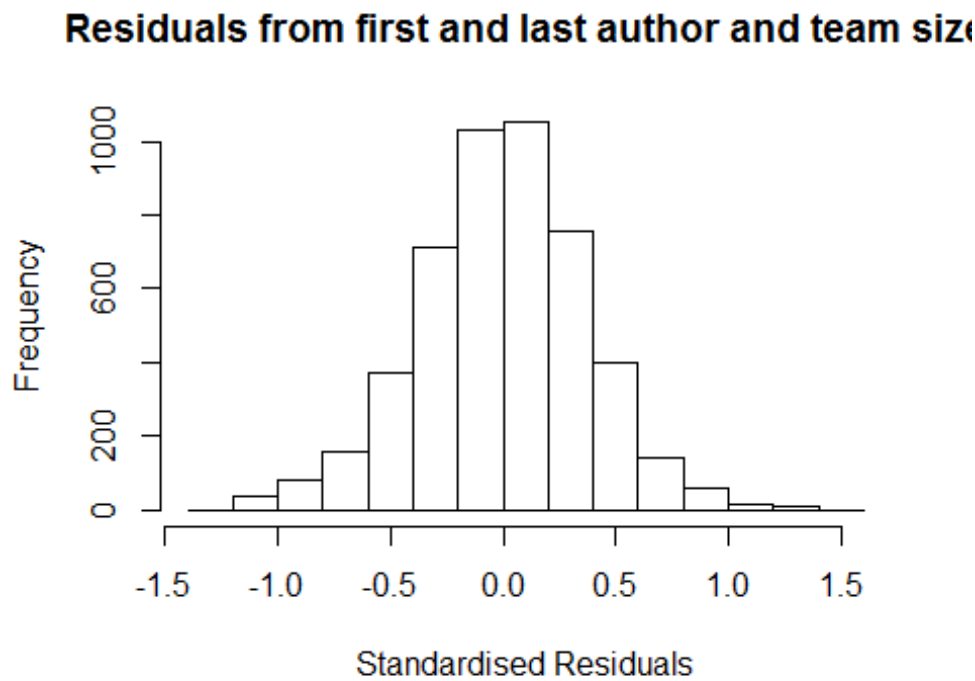
```
##  350  339  325  319  287  267  307  282  242  267  289  373  340  324  346
## 2011 2012
##  410  402
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  313  293  295  283  253  229  273  246  212  234  258  330  303  289  305
## 2011 2012
##  361  364
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
##  Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 49, df = 16, p-value = 4e-05
```




```
##
##  Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 22, df = 1, p-value = 2e-06
```

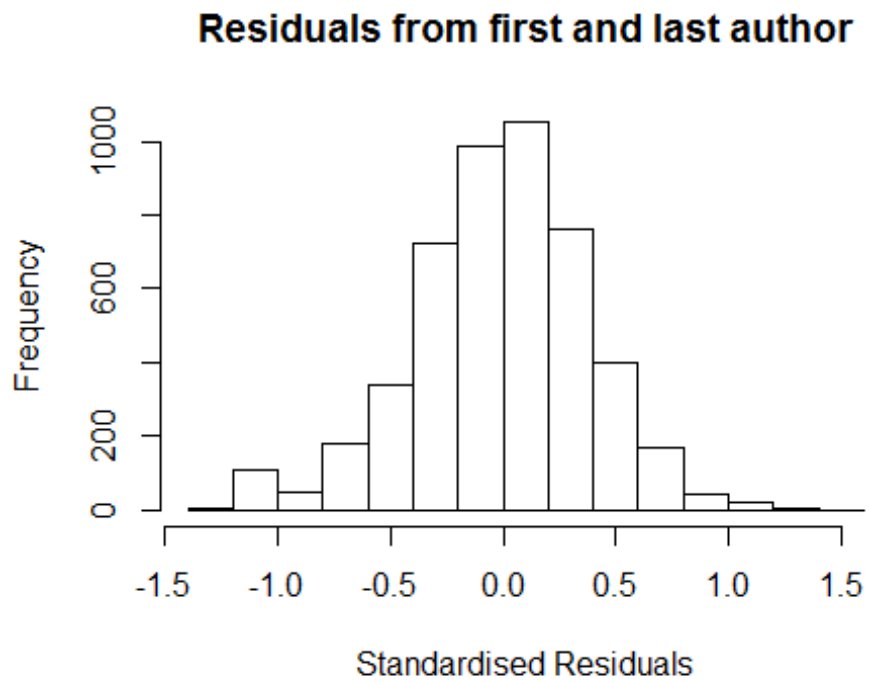



```
## [1] "Female first author team size 2018 geometric mean: 4.65463306792843"
## [1] "Male first author team size 2018 geometric mean: 4.01211733433694"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 15000, p-value = 0.08
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 4.52519016584331"
## [1] "Male last author team size 2018 geometric mean: 4.28033896027641"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 14000, p-value = 0.4
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.050  1           1.024
## LastAuthorFemale  1.026  1           1.013
## UniqueAuthors     1.135  4           1.016
## Year              1.180 16           1.005
```



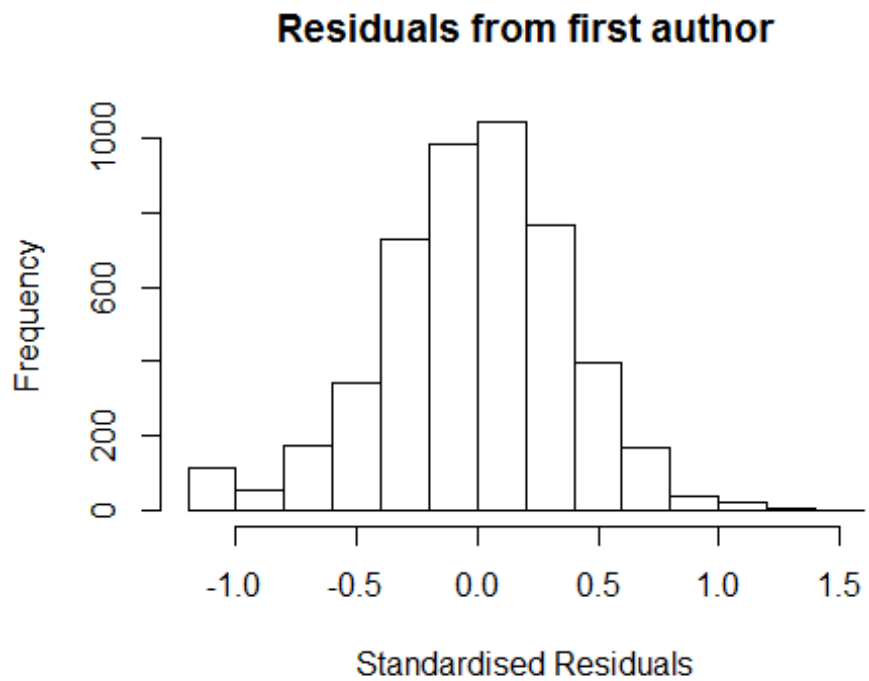

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
UniqueAuthors +
##     Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##     k.max = 1000))
##  \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.27570 -0.23904  0.00296  0.24035  1.44027
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         0.85206    0.03823   22.29  < 2e-16 ***
## FirstAuthorFemale1  0.01072    0.01089    0.98   0.3249
## LastAuthorFemale1   0.02198    0.01161    1.89   0.0583 .
## UniqueAuthors2      0.21290    0.03437    6.20  6.3e-10 ***
## UniqueAuthors3      0.25267    0.03437    7.35  2.3e-13 ***
## UniqueAuthors4      0.28106    0.03458    8.13  5.5e-16 ***
## UniqueAuthors5      0.38464    0.03398   11.32  < 2e-16 ***
## Year1997            0.00232    0.03181    0.07   0.9419
## Year1998            0.00660    0.03084    0.21   0.8306
## Year1999           -0.00800    0.03075   -0.26   0.7947
```

```
## Year2000           -0.01065    0.03024   -0.35   0.7247
## Year2001           -0.08417    0.03151   -2.67   0.0076 **
## Year2002           -0.03597    0.03226   -1.12   0.2649
## Year2003           -0.04901    0.03101   -1.58   0.1140
## Year2004           -0.05974    0.03231   -1.85   0.0645 .
## Year2005           -0.12902    0.03317   -3.89   0.0001 ***
## Year2006           -0.09400    0.03212   -2.93   0.0034 **
## Year2007           -0.08903    0.02951   -3.02   0.0026 **
## Year2008           -0.02188    0.03160   -0.69   0.4887
## Year2009            0.01360    0.03247    0.42   0.6753
## Year2010            0.01080    0.03080    0.35   0.7259
## Year2011            0.02829    0.03109    0.91   0.3629
## Year2012           -0.00594    0.03068   -0.19   0.8465
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.355
## Multiple R-squared:  0.0913, Adjusted R-squared:  0.0871
## Convergence in 18 IRWLS iterations
##
## Robustness weights:
##  415 weights are ~= 1. The remaining 4426 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0618  0.8680  0.9500  0.8980  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          2.07e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.054  1           1.026
## LastAuthorFemale  1.029  1           1.015
## Year              1.059 16           1.002
```



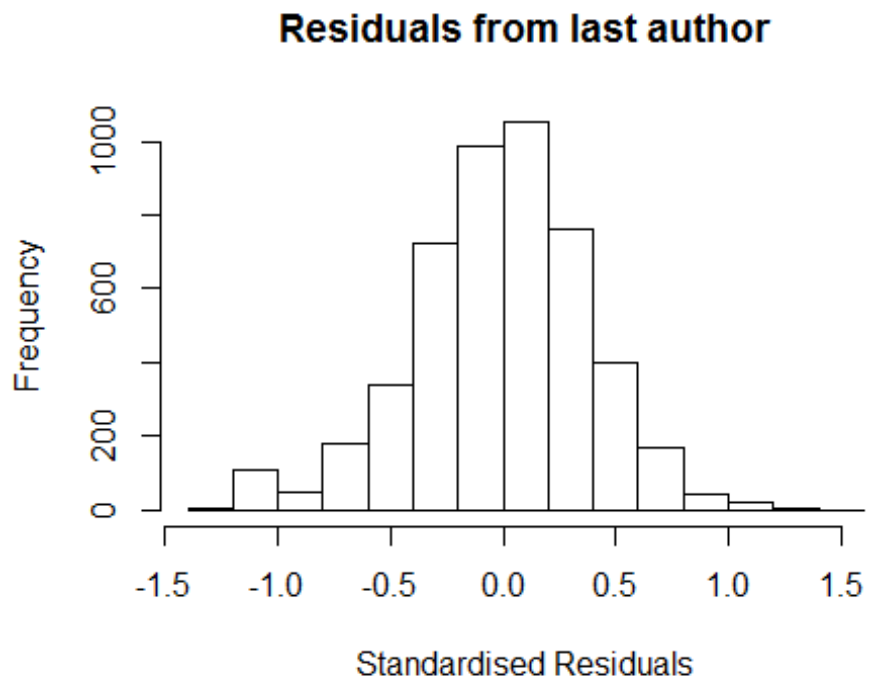

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.20575 -0.24338  0.00406  0.24198  1.44357
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)          1.10143    0.02185   50.42   <2e-16 ***
## FirstAuthorFemale1   0.01647    0.01118    1.47    0.141
## LastAuthorFemale1    0.01612    0.01175    1.37    0.170
## Year1997            -0.00388    0.03198   -0.12    0.903
## Year1998             0.01759    0.03065    0.57    0.566
## Year1999            -0.00503    0.03037   -0.17    0.869
## Year2000             0.00965    0.02985    0.32    0.746
## Year2001            -0.05747    0.03156   -1.82    0.069 .
## Year2002            -0.02394    0.03233   -0.74    0.459
## Year2003            -0.02508    0.03076   -0.82    0.415
## Year2004            -0.03028    0.03174   -0.95    0.340
## Year2005            -0.07705    0.03312   -2.33    0.020 *
```

```
## Year2006           -0.05599    0.03213   -1.74    0.081 .
## Year2007           -0.07386    0.03011   -2.45    0.014 *
## Year2008            0.01439    0.03149    0.46    0.648
## Year2009            0.05501    0.03211    1.71    0.087 .
## Year2010            0.05102    0.03063    1.67    0.096 .
## Year2011            0.07172    0.03120    2.30    0.022 *
## Year2012            0.04891    0.03010    1.63    0.104
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.36
## Multiple R-squared:  0.0158, Adjusted R-squared:  0.0121
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  413 weights are ~= 1. The remaining 4428 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0724  0.8670  0.9500  0.8940  0.9860  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          2.07e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.045  1           1.022
## Year              1.045 16           1.001
```




```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.19730 -0.24398  0.00383  0.24371  1.44002
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)          1.10498    0.02172   50.87   <2e-16 ***
## FirstAuthorFemale1  0.01896    0.01115    1.70    0.089 .
## Year1997           -0.00400    0.03202   -0.12    0.901
## Year1998            0.01719    0.03065    0.56    0.575
## Year1999           -0.00536    0.03038   -0.18    0.860
## Year2000            0.01032    0.02987    0.35    0.730
## Year2001           -0.05700    0.03163   -1.80    0.072 .
## Year2002           -0.02277    0.03235   -0.70    0.482
## Year2003           -0.02473    0.03075   -0.80    0.421
## Year2004           -0.02975    0.03180   -0.94    0.349
## Year2005           -0.07697    0.03311   -2.32    0.020 *
## Year2006           -0.05521    0.03221   -1.71    0.087 .
```

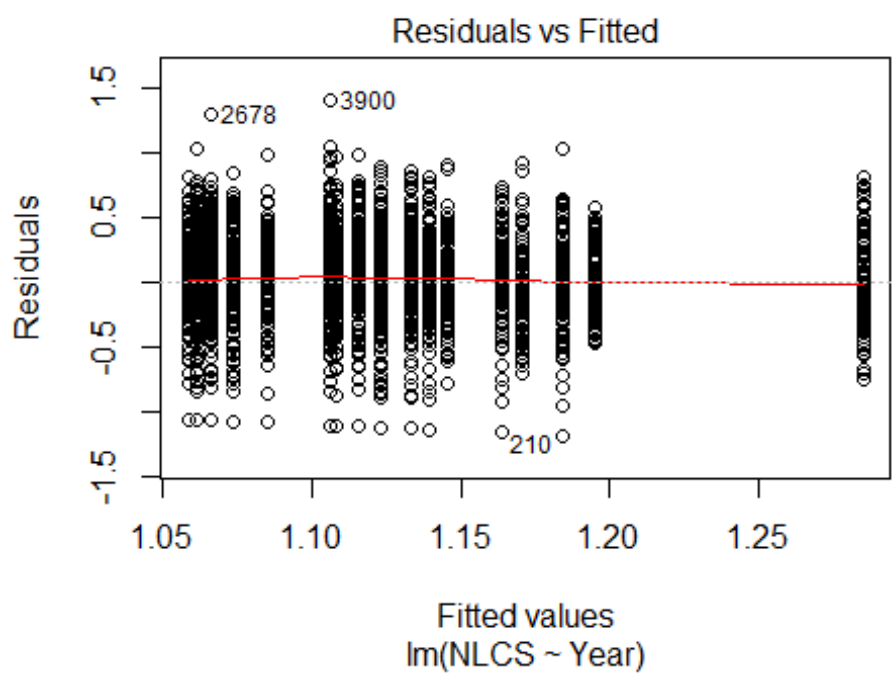
```
## Year2007           -0.07327    0.03015   -2.43    0.015 *
## Year2008            0.01568    0.03149    0.50    0.619
## Year2009            0.05592    0.03213    1.74    0.082 .
## Year2010            0.05167    0.03063    1.69    0.092 .
## Year2011            0.07336    0.03113    2.36    0.018 *
## Year2012            0.04975    0.03016    1.65    0.099 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.36
## Multiple R-squared:  0.0154, Adjusted R-squared:  0.012
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  412 weights are ~= 1. The remaining 4429 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0744  0.8670  0.9510  0.8940  0.9860  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          2.07e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##                   GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.021  1           1.010
## Year             1.021 16           1.001
```



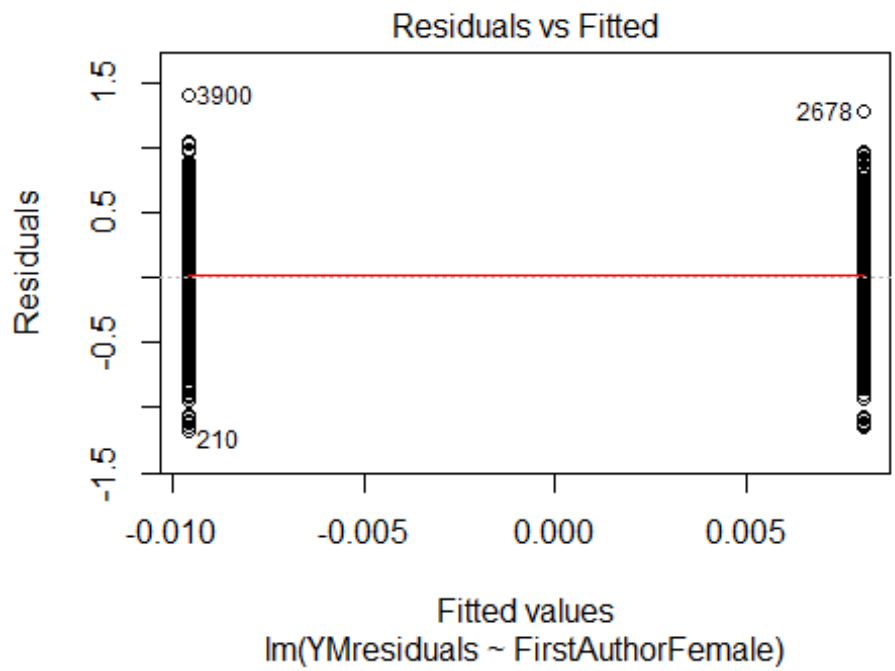

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.20013 -0.24203  0.00409  0.24354  1.43821
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.10679    0.02152   51.44   <2e-16 ***
## LastAuthorFemale1  0.01898    0.01171    1.62    0.105
## Year1997          -0.00388    0.03195   -0.12    0.903
## Year1998           0.01869    0.03055    0.61    0.541
## Year1999          -0.00412    0.03027   -0.14    0.892
## Year2000           0.00979    0.02983    0.33    0.743
## Year2001          -0.05677    0.03146   -1.80    0.071 .
## Year2002          -0.02254    0.03228   -0.70    0.485
## Year2003          -0.02428    0.03071   -0.79    0.429
## Year2004          -0.03003    0.03171   -0.95    0.344
## Year2005          -0.07657    0.03310   -2.31    0.021 *
## Year2006          -0.05563    0.03211   -1.73    0.083 .
```

```
## Year2007          -0.07198    0.02999   -2.40    0.016 *
## Year2008           0.01626    0.03140    0.52    0.604
## Year2009           0.05678    0.03196    1.78    0.076 .
## Year2010           0.05309    0.03053    1.74    0.082 .
## Year2011           0.07437    0.03102    2.40    0.017 *
## Year2012           0.05077    0.02997    1.69    0.090 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.36
## Multiple R-squared:  0.0154, Adjusted R-squared:  0.0119
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  419 weights are ~= 1. The remaining 4422 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0751  0.8660  0.9510  0.8940  0.9860  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          2.07e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Sample size for the above analysis:  4841"
## [1] ""
## [1] ""
## [1] "####################################"
## [1] "Analysis of AJSC 2807"
## [1] "####################################"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  190  173  138  157  200  164  220  183  190  188  195  256  272  321  371
## 2011 2012
##  380  346
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  114  108   83   94  124   91  153  120  131  145  143  172  207  223  250
## 2011 2012
```

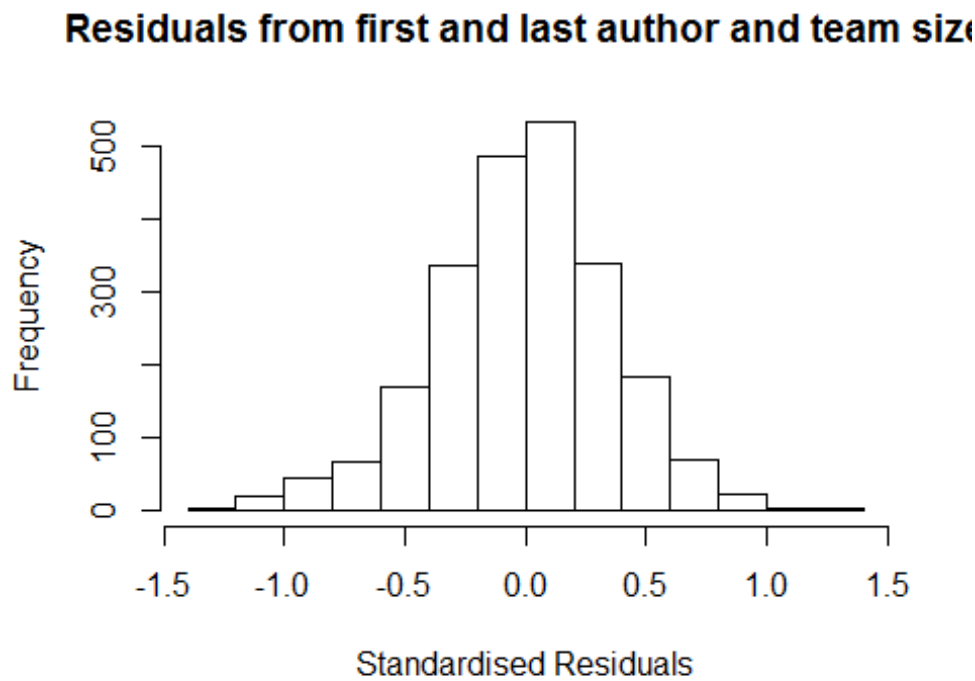
```
##  247  199
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##   98   99   72   84  105   80  134  103  114  121  126  147  184  194  226
## 2011 2012
##  221  177
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
##  Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 58, df = 16, p-value = 1e-06
```




```
##
##  Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 0.52, df = 1, p-value = 0.5
```

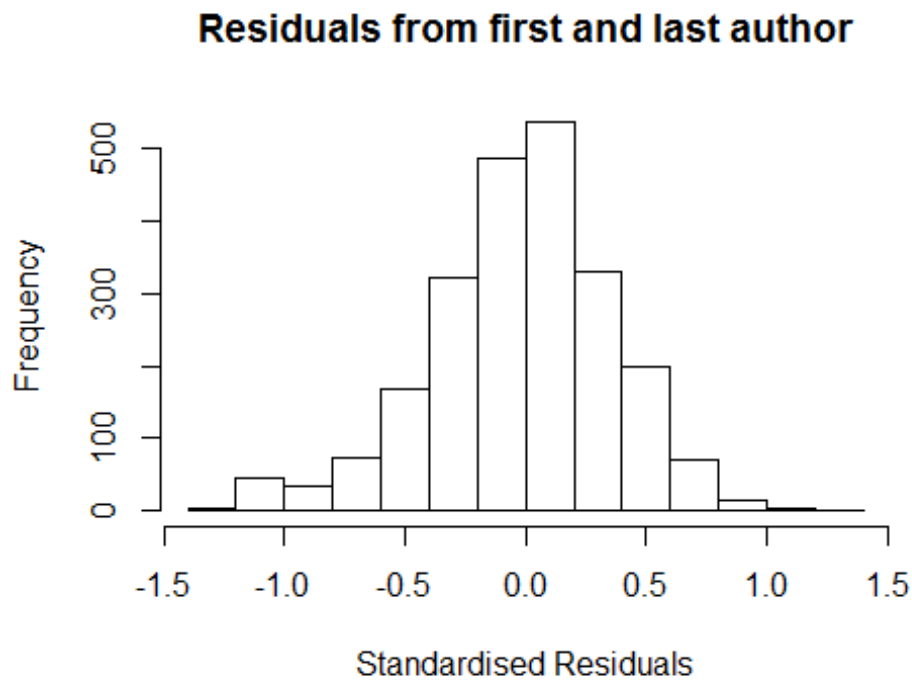



```
## [1] "Female first author team size 2018 geometric mean: 4.84659348287137"
## [1] "Male first author team size 2018 geometric mean: 5.4741245106643"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 4200, p-value = 0.1
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 4.74695458509278"
## [1] "Male last author team size 2018 geometric mean: 5.35332266208584"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 5000, p-value = 0.3
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.074  1           1.036
## LastAuthorFemale  1.090  1           1.044
## UniqueAuthors     1.214  4           1.025
## Year              1.304 16           1.008
```



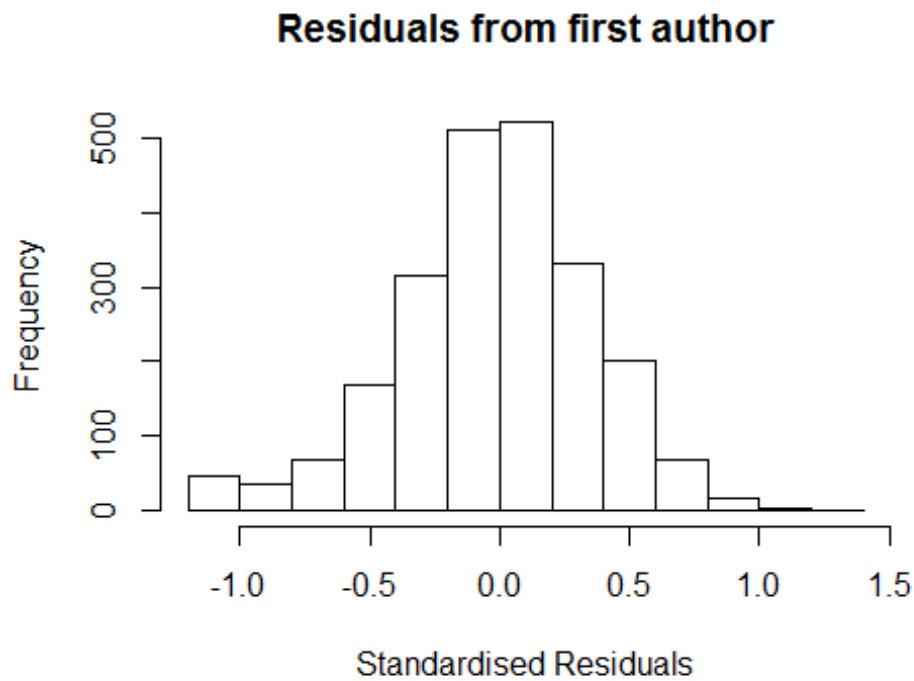

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
UniqueAuthors +
##     Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##     k.max = 1000))
##  \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.23909 -0.23174  0.00391  0.22795  1.22738
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         0.98290    0.05468   17.98  < 2e-16 ***
## FirstAuthorFemale1  0.02081    0.01573    1.32   0.1859
## LastAuthorFemale1  -0.04695    0.01585   -2.96   0.0031 **
## UniqueAuthors2      0.18279    0.04017    4.55  5.6e-06 ***
## UniqueAuthors3      0.17686    0.04048    4.37  1.3e-05 ***
## UniqueAuthors4      0.23697    0.04059    5.84  6.0e-09 ***
## UniqueAuthors5      0.22893    0.03904    5.86  5.2e-09 ***
## Year1997            0.13519    0.05522    2.45   0.0144 *
## Year1998            0.01766    0.05880    0.30   0.7639
## Year1999            0.02691    0.04832    0.56   0.5776
```

```
## Year2000           -0.00220    0.04678   -0.05   0.9625
## Year2001           -0.02108    0.05230   -0.40   0.6870
## Year2002           -0.01069    0.04962   -0.22   0.8295
## Year2003           -0.10404    0.05094   -2.04   0.0412 *
## Year2004           -0.08012    0.04859   -1.65   0.0993 .
## Year2005           -0.06694    0.04924   -1.36   0.1742
## Year2006            0.00643    0.04725    0.14   0.8918
## Year2007           -0.08425    0.04839   -1.74   0.0818 .
## Year2008           -0.09601    0.04665   -2.06   0.0397 *
## Year2009           -0.02917    0.04523   -0.64   0.5191
## Year2010           -0.00159    0.04352   -0.04   0.9708
## Year2011           -0.04897    0.04614   -1.06   0.2887
## Year2012           -0.02074    0.04950   -0.42   0.6753
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.343
## Multiple R-squared:  0.0627, Adjusted R-squared:  0.0536
## Convergence in 15 IRWLS iterations
##
## Robustness weights:
##  197 weights are ~= 1. The remaining 2088 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.166   0.864   0.951   0.896   0.986   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          4.38e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.064  1           1.031
## LastAuthorFemale  1.072  1           1.035
## Year              1.101 16           1.003
```



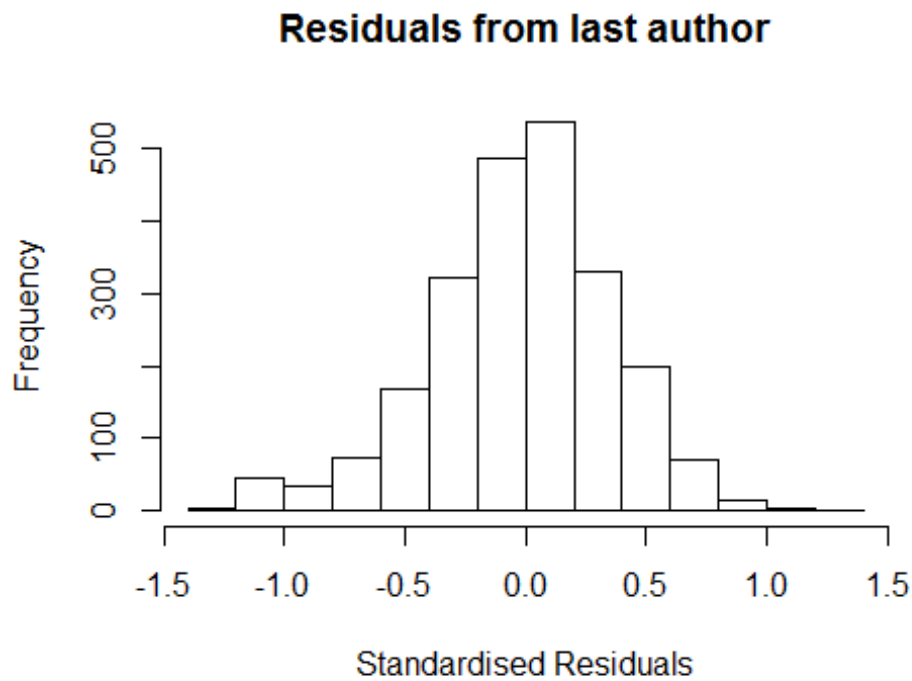

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.20616 -0.22390  0.00604  0.22388  1.25704
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.181301   0.037295   31.67   <2e-16 ***
## FirstAuthorFemale1  0.020858   0.015714    1.33    0.185
## LastAuthorFemale1  -0.062947   0.015927   -3.95    8e-05 ***
## Year1997            0.122414   0.052974    2.31    0.021 *
## Year1998            0.000151   0.055925    0.00    0.998
## Year1999            0.021943   0.046388    0.47    0.636
## Year2000           -0.012892   0.046303   -0.28    0.781
## Year2001           -0.017074   0.051467   -0.33    0.740
## Year2002           -0.027938   0.047609   -0.59    0.557
## Year2003           -0.111552   0.050543   -2.21    0.027 *
## Year2004           -0.072313   0.048410   -1.49    0.135
## Year2005           -0.068544   0.048893   -1.40    0.161
```

```
## Year2006           -0.004187   0.045680   -0.09    0.927
## Year2007           -0.088250   0.048694   -1.81    0.070 .
## Year2008           -0.095202   0.046534   -2.05    0.041 *
## Year2009           -0.026318   0.044883   -0.59    0.558
## Year2010            0.004005   0.042493    0.09    0.925
## Year2011           -0.038314   0.045492   -0.84    0.400
## Year2012            0.001910   0.048442    0.04    0.969
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.345
## Multiple R-squared:  0.0263, Adjusted R-squared:  0.0186
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  211 weights are ~= 1. The remaining 2074 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.156   0.861   0.951   0.893   0.985   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          4.38e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.047  1           1.023
## Year              1.047 16           1.001
```




```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##       Min        1Q    Median        3Q       Max
## -1.169294 -0.230842  0.000129  0.218813  1.287647
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)          1.16549    0.03719   31.34   <2e-16 ***
## FirstAuthorFemale1   0.00765    0.01569    0.49    0.626
## Year1997             0.12008    0.05349    2.25    0.025 *
## Year1998            -0.00261    0.05616   -0.05    0.963
## Year1999             0.02636    0.04657    0.57    0.571
## Year2000            -0.01099    0.04691   -0.23    0.815
## Year2001            -0.01046    0.05128   -0.20    0.838
## Year2002            -0.02781    0.04760   -0.58    0.559
## Year2003            -0.11771    0.05093   -2.31    0.021 *
## Year2004            -0.07346    0.04848   -1.52    0.130
## Year2005            -0.06988    0.04936   -1.42    0.157
## Year2006            -0.00430    0.04622   -0.09    0.926
```

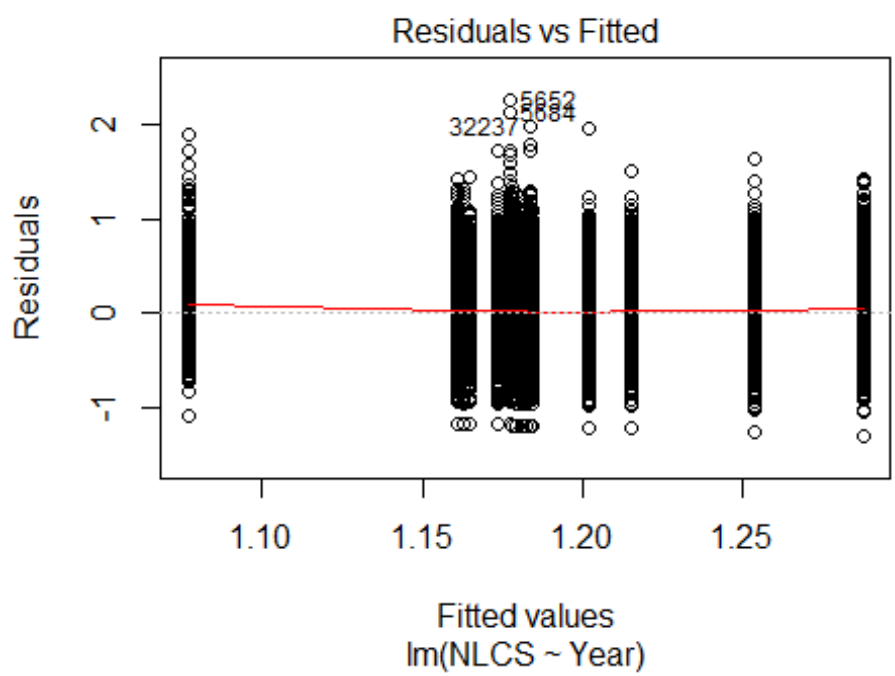
```
## Year2007           -0.09108    0.04927   -1.85    0.065 .
## Year2008           -0.09679    0.04699   -2.06    0.040 *
## Year2009           -0.02575    0.04513   -0.57    0.568
## Year2010           -0.00427    0.04301   -0.10    0.921
## Year2011           -0.04467    0.04588   -0.97    0.330
## Year2012           -0.00385    0.04905   -0.08    0.937
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.346
## Multiple R-squared:  0.0193, Adjusted R-squared:  0.0119
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  200 weights are ~= 1. The remaining 2085 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.136   0.863   0.953   0.893   0.985   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          4.38e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##                   GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.057  1           1.028
## Year             1.057 16           1.002
```



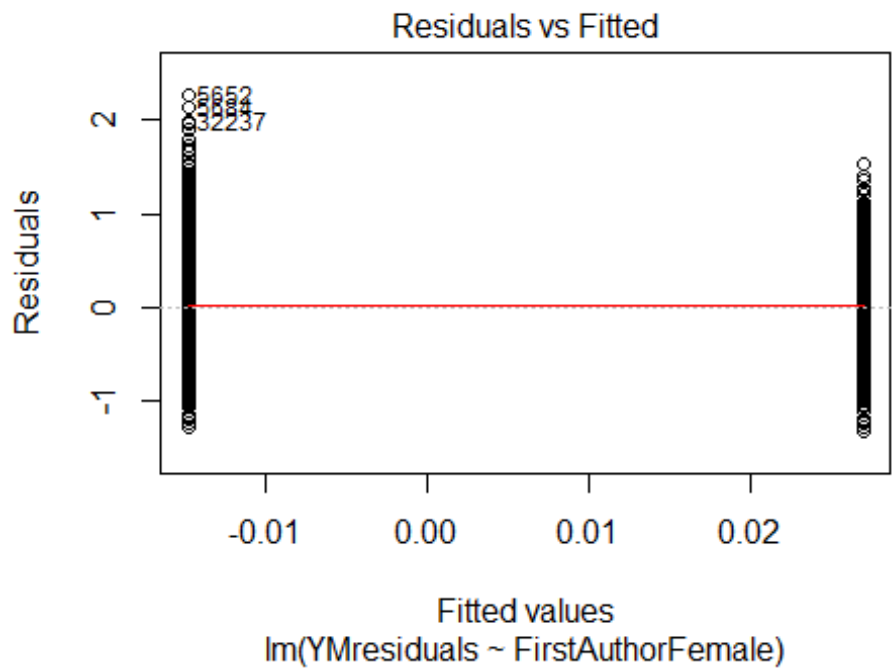

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.19615 -0.22587  0.00524  0.22405  1.26858
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)        1.189618   0.036528   32.57  < 2e-16 ***
## LastAuthorFemale1 -0.058462   0.015868   -3.68  0.00023 ***
## Year1997           0.122210   0.052812    2.31  0.02075 *
## Year1998          -0.002020   0.055593   -0.04  0.97102
## Year1999           0.021454   0.046268    0.46  0.64292
## Year2000          -0.013394   0.046323   -0.29  0.77250
## Year2001          -0.015039   0.051469   -0.29  0.77016
## Year2002          -0.026620   0.047545   -0.56  0.57560
## Year2003          -0.110996   0.050473   -2.20  0.02797 *
## Year2004          -0.072045   0.048297   -1.49  0.13591
## Year2005          -0.066426   0.048689   -1.36  0.17261
## Year2006          -0.000493   0.045665   -0.01  0.99138
```

```
## Year2007          -0.086435   0.048821   -1.77  0.07678 .
## Year2008          -0.094195   0.046551   -2.02  0.04314 *
## Year2009          -0.024239   0.044811   -0.54  0.58862
## Year2010           0.006333   0.042476    0.15  0.88149
## Year2011          -0.036743   0.045411   -0.81  0.41852
## Year2012           0.006530   0.048213    0.14  0.89227
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.345
## Multiple R-squared:  0.0257, Adjusted R-squared:  0.0184
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  208 weights are ~= 1. The remaining 2077 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.147   0.863   0.951   0.893   0.985   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          4.38e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Sample size for the above analysis:  2285"
## [1] ""
## [1] ""
## [1] "####################################"
## [1] "Analysis of AJSC 2808"
## [1] "####################################"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 1179 1394 1508 1431 1581 1413 1299 1296 1485 1659 1789 1968 2095 2423 2424
## 2011 2012
## 2794 2781
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  814  855  866  959  976  837  886  912 1087 1205 1238 1419 1473 1752 1730
## 2011 2012
```

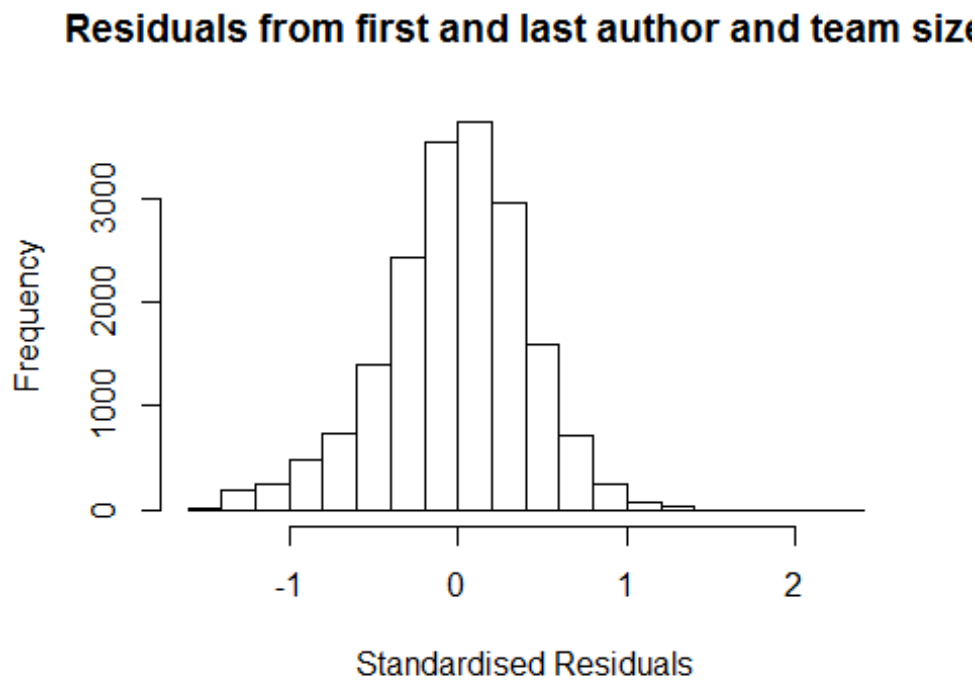
```
## 1986 2021
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  744  738  761  832  868  739  791  785  944 1034 1064 1239 1284 1512 1511
## 2011 2012
## 1750 1790
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
##  Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 620, df = 16, p-value <2e-16
```




```
##
##  Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 110, df = 1, p-value <2e-16
```

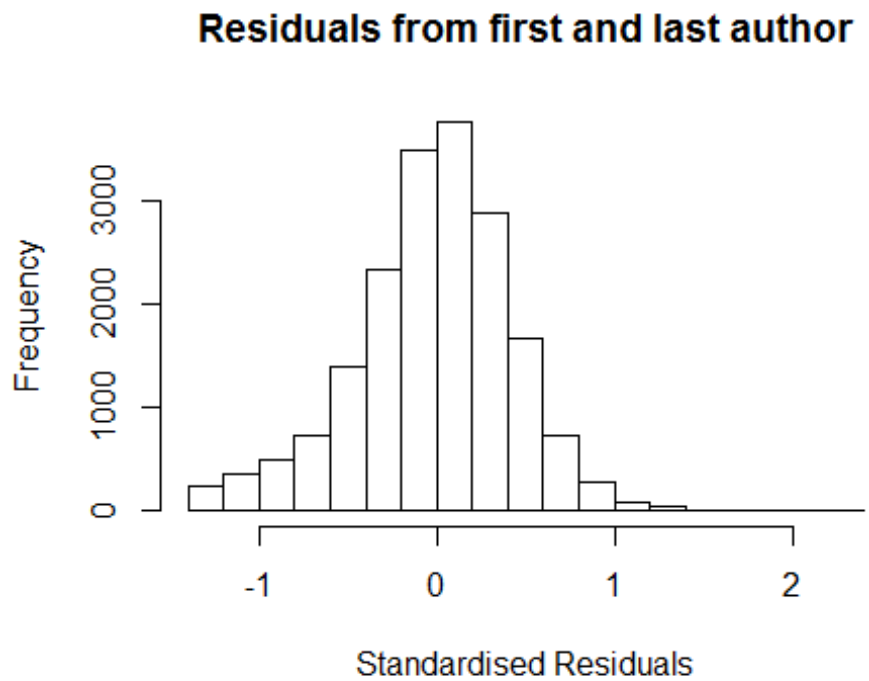



```
## [1] "Female first author team size 2018 geometric mean: 5.12493952591086"
## [1] "Male first author team size 2018 geometric mean: 4.94650585671168"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 470000, p-value = 0.6
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 5.07293197734391"
## [1] "Male last author team size 2018 geometric mean: 5.00213677615404"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 410000, p-value = 0.6
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.021  1           1.011
## LastAuthorFemale  1.029  1           1.014
## UniqueAuthors     1.076  4           1.009
## Year              1.074 16           1.002
```



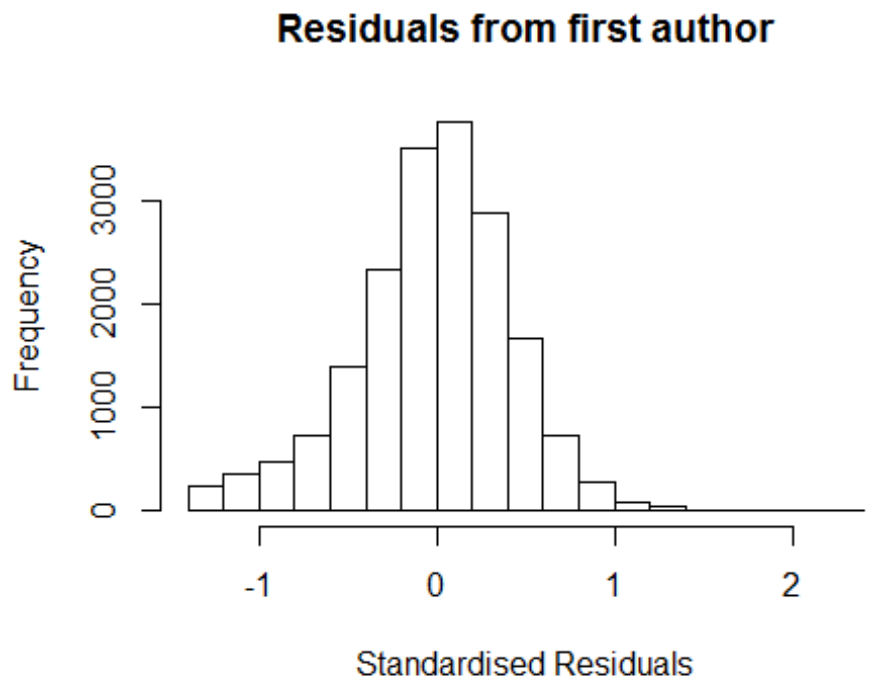

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
UniqueAuthors +
##     Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##     k.max = 1000))
##  \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.43351 -0.26070  0.00897  0.26146  2.22083
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         0.97571    0.02235   43.66  < 2e-16 ***
## FirstAuthorFemale1  0.02600    0.00629    4.13  3.6e-05 ***
## LastAuthorFemale1   0.01759    0.00704    2.50  0.01249 *
## UniqueAuthors2      0.17564    0.01810    9.71  < 2e-16 ***
## UniqueAuthors3      0.22338    0.01768   12.64  < 2e-16 ***
## UniqueAuthors4      0.26845    0.01752   15.32  < 2e-16 ***
## UniqueAuthors5      0.36551    0.01663   21.98  < 2e-16 ***
## Year1997            0.08405    0.02397    3.51  0.00046 ***
## Year1998            0.04871    0.02482    1.96  0.04970 *
## Year1999            0.00808    0.02541    0.32  0.75047
```

```
## Year2000           -0.10492    0.02642   -3.97  7.2e-05 ***
## Year2001           -0.04031    0.02169   -1.86  0.06306 .
## Year2002           -0.05505    0.02123   -2.59  0.00953 **
## Year2003           -0.07227    0.02100   -3.44  0.00058 ***
## Year2004           -0.05849    0.02041   -2.87  0.00416 **
## Year2005           -0.07773    0.02042   -3.81  0.00014 ***
## Year2006           -0.09295    0.02055   -4.52  6.1e-06 ***
## Year2007           -0.08384    0.01993   -4.21  2.6e-05 ***
## Year2008           -0.06818    0.01998   -3.41  0.00064 ***
## Year2009           -0.08049    0.01993   -4.04  5.4e-05 ***
## Year2010           -0.08896    0.01995   -4.46  8.3e-06 ***
## Year2011           -0.08240    0.01960   -4.20  2.6e-05 ***
## Year2012           -0.07992    0.01964   -4.07  4.7e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.387
## Multiple R-squared:  0.0733, Adjusted R-squared:  0.0722
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
##  4 observations c(2773,2792,16778,17259)
##   are outliers with |weight| = 0 ( < 5.4e-06);
##  1645 weights are ~= 1. The remaining 16737 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.002   0.863   0.949   0.890   0.985   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          5.44e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.017  1           1.008
## LastAuthorFemale  1.016  1           1.008
## Year              1.009 16           1.000
```



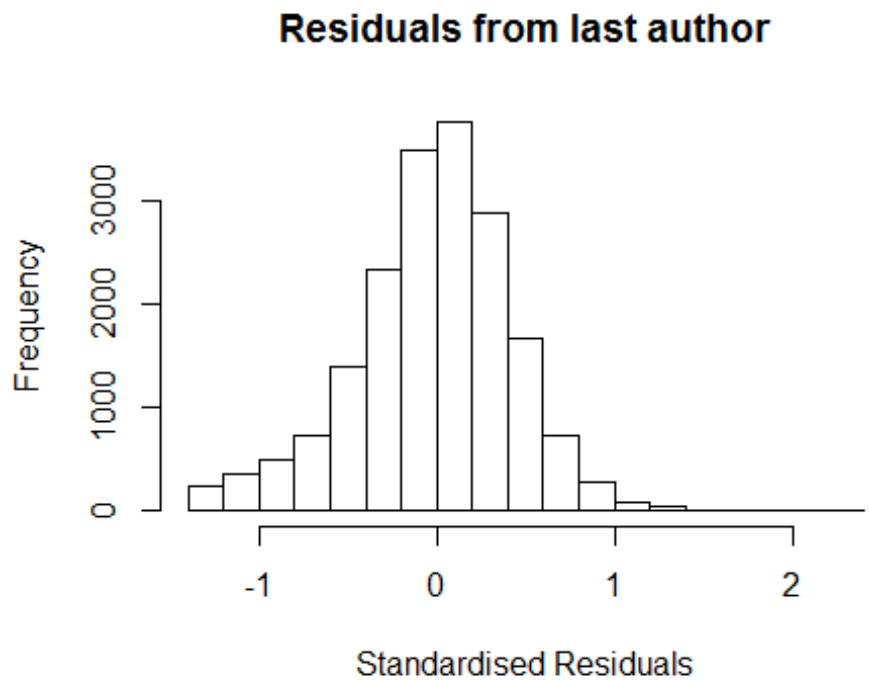

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.3455 -0.2674  0.0105  0.2652  2.2044
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)          1.21753    0.01767   68.92  < 2e-16 ***
## FirstAuthorFemale1   0.03717    0.00643    5.78  7.4e-09 ***
## LastAuthorFemale1    0.00236    0.00714    0.33  0.74087
## Year1997             0.09084    0.02475    3.67  0.00024 ***
## Year1998             0.05504    0.02562    2.15  0.03167 *
## Year1999             0.00605    0.02621    0.23  0.81741
## Year2000            -0.09999    0.02716   -3.68  0.00023 ***
## Year2001            -0.02467    0.02280   -1.08  0.27912
## Year2002            -0.02834    0.02228   -1.27  0.20332
## Year2003            -0.04099    0.02198   -1.86  0.06223 .
## Year2004            -0.03203    0.02119   -1.51  0.13061
## Year2005            -0.04331    0.02112   -2.05  0.04032 *
```

```
## Year2006           -0.05700    0.02117   -2.69  0.00711 **
## Year2007           -0.04892    0.02066   -2.37  0.01790 *
## Year2008           -0.02428    0.02076   -1.17  0.24203
## Year2009           -0.03541    0.02066   -1.71  0.08657 .
## Year2010           -0.04570    0.02076   -2.20  0.02771 *
## Year2011           -0.03324    0.02028   -1.64  0.10112
## Year2012           -0.03386    0.02029   -1.67  0.09528 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.394
## Multiple R-squared:  0.009,  Adjusted R-squared:  0.00802
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
##  4 observations c(2773,2792,3373,16778)
##   are outliers with |weight| = 0 ( < 5.4e-06);
##  1616 weights are ~= 1. The remaining 16766 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0042  0.8610  0.9490  0.8880  0.9860  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          5.44e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.005  1           1.003
## Year              1.005 16           1.000
```




```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.3464 -0.2677  0.0107  0.2647  2.2040
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)          1.21788    0.01763   69.08  < 2e-16 ***
## FirstAuthorFemale1   0.03750    0.00640    5.86  4.6e-09 ***
## Year1997             0.09098    0.02473    3.68  0.00024 ***
## Year1998             0.05504    0.02561    2.15  0.03168 *
## Year1999             0.00613    0.02620    0.23  0.81491
## Year2000            -0.09982    0.02715   -3.68  0.00024 ***
## Year2001            -0.02451    0.02278   -1.08  0.28198
## Year2002            -0.02823    0.02227   -1.27  0.20490
## Year2003            -0.04087    0.02198   -1.86  0.06293 .
## Year2004            -0.03191    0.02118   -1.51  0.13190
## Year2005            -0.04319    0.02111   -2.05  0.04081 *
## Year2006            -0.05692    0.02117   -2.69  0.00717 **
```

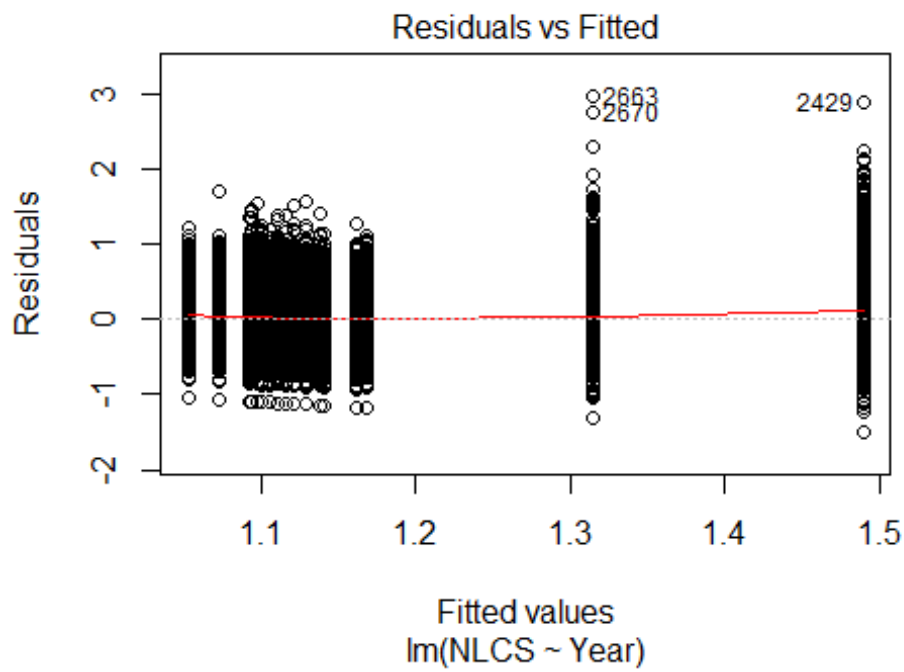
```
## Year2007           -0.04879    0.02064   -2.36  0.01813 *
## Year2008           -0.02415    0.02074   -1.16  0.24441
## Year2009           -0.03531    0.02065   -1.71  0.08731 .
## Year2010           -0.04557    0.02075   -2.20  0.02807 *
## Year2011           -0.03310    0.02026   -1.63  0.10241
## Year2012           -0.03368    0.02028   -1.66  0.09679 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.394
## Multiple R-squared:  0.00899,    Adjusted R-squared:  0.00807
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
##  4 observations c(2773,2792,3373,16778)
##   are outliers with |weight| = 0 ( < 5.4e-06);
##  1626 weights are ~= 1. The remaining 16756 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0042  0.8610  0.9490  0.8880  0.9860  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          5.44e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##                   GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.004  1           1.002
## Year             1.004 16           1.000
```



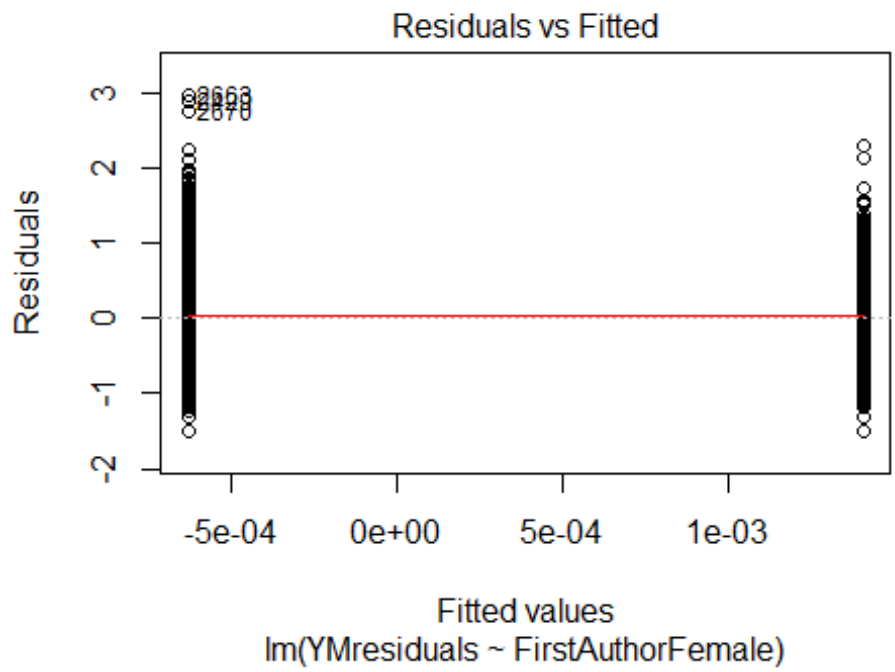

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.3275 -0.2666  0.0108  0.2656  2.1940
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.22755    0.01753   70.01  < 2e-16 ***
## LastAuthorFemale1   0.00893    0.00710    1.26  0.20855
## Year1997            0.09101    0.02474    3.68  0.00023 ***
## Year1998            0.05650    0.02559    2.21  0.02725 *
## Year1999            0.00645    0.02623    0.25  0.80579
## Year2000           -0.09934    0.02712   -3.66  0.00025 ***
## Year2001           -0.02438    0.02279   -1.07  0.28475
## Year2002           -0.02612    0.02232   -1.17  0.24179
## Year2003           -0.03939    0.02198   -1.79  0.07323 .
## Year2004           -0.03014    0.02118   -1.42  0.15481
## Year2005           -0.04110    0.02111   -1.95  0.05159 .
## Year2006           -0.05411    0.02117   -2.56  0.01060 *
```

```
## Year2007          -0.04596    0.02065   -2.23  0.02605 *
## Year2008          -0.02131    0.02075   -1.03  0.30445
## Year2009          -0.03273    0.02066   -1.58  0.11312
## Year2010          -0.04335    0.02076   -2.09  0.03681 *
## Year2011          -0.03061    0.02027   -1.51  0.13104
## Year2012          -0.03038    0.02029   -1.50  0.13423
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.394
## Multiple R-squared:  0.00724,    Adjusted R-squared:  0.00632
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
##  3 observations c(2773,2792,16778)
##   are outliers with |weight| = 0 ( < 5.4e-06);
##  1632 weights are ~= 1. The remaining 16751 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.000   0.861   0.949   0.888   0.986   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          5.44e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                  psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Sample size for the above analysis:  18386"
## [1] ""
## [1] ""
## [1] "##################################"
## [1] "Analysis of AJSC 2809"
## [1] "##################################"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 2409 1877  916  919  952  977 1023 1002  936  925 1040  992 1090 1232 1212
## 2011 2012
## 1353 1179
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
```

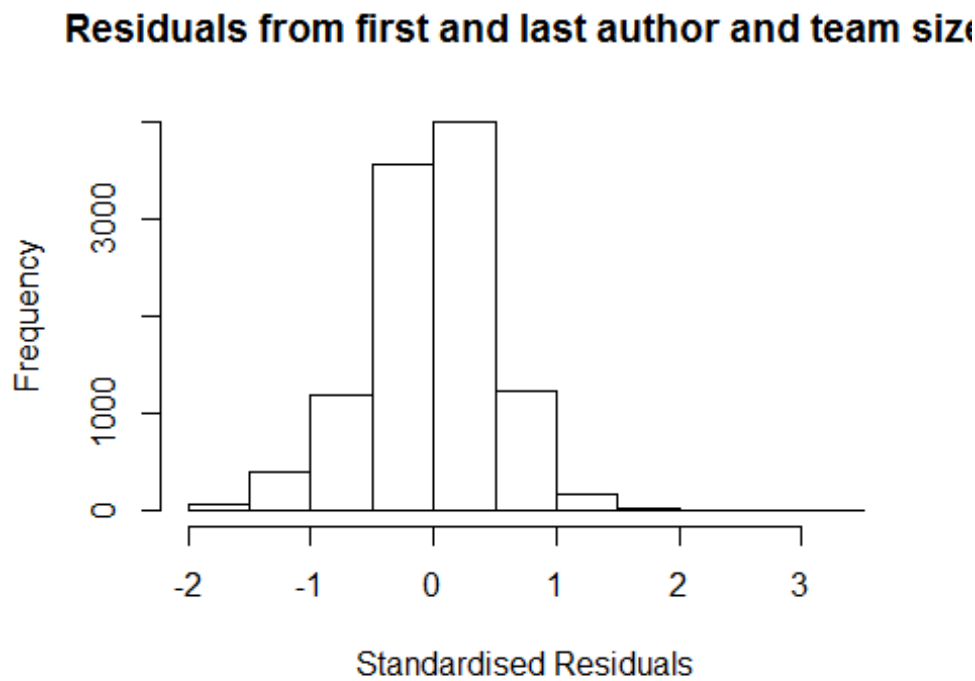
```
##  753  784  615  661  579  500  760  751  671  653  714  701  775  849  863
## 2011 2012
##  941  855
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  633  695  546  575  503  420  661  628  568  554  608  597  680  722  734
## 2011 2012
##  805  727
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
##  Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 1300, df = 16, p-value <2e-16
```




```
##
##  Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 20, df = 1, p-value = 8e-06
```

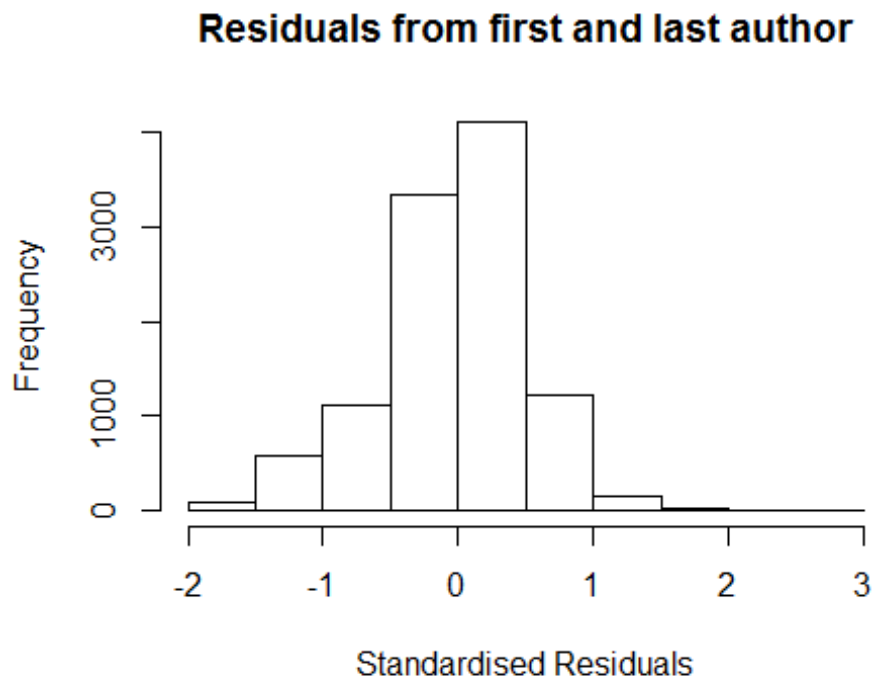



```
## [1] "Female first author team size 2018 geometric mean: 4.55546953873318"
## [1] "Male first author team size 2018 geometric mean: 3.85764355852788"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 44000, p-value = 0.005
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 4.33939906267803"
## [1] "Male last author team size 2018 geometric mean: 4.00101226789007"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 35000, p-value = 0.2
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.026  1           1.013
## LastAuthorFemale  1.017  1           1.008
## UniqueAuthors     1.097  4           1.012
## Year              1.102 16           1.003
```



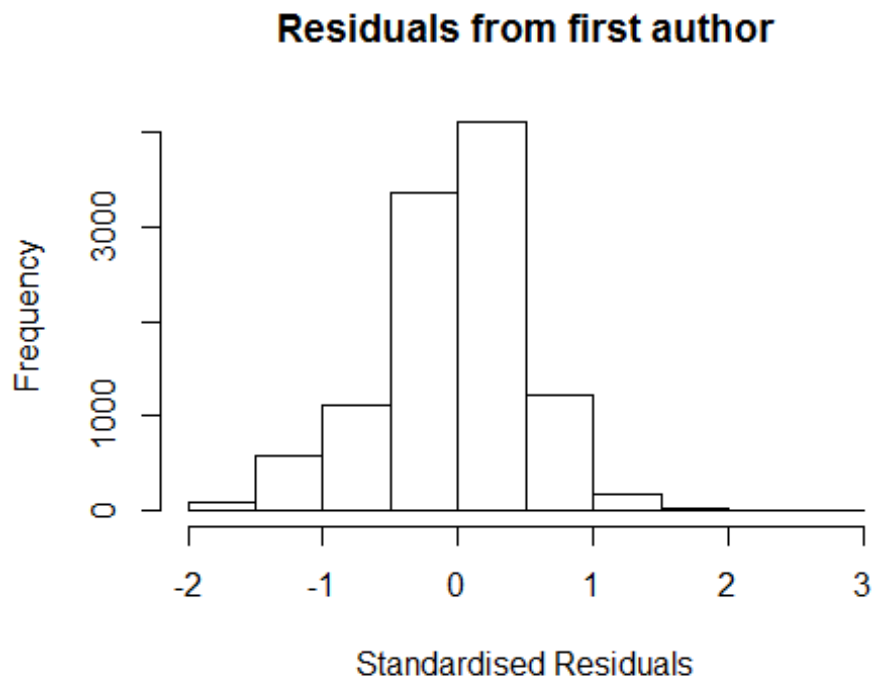

```
## [1] "List of  3 outliers with residuals above  2.5"
##        ScopusId  NLCS Year OneField Fields residuals
## 2429 0029980777 4.398 1996     2731      3     3.156
## 2663 0030779611 4.287 1997     2731      2     2.942
## 2670 0030832881 4.070 1997     2731      2     2.725
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
UniqueAuthors +
##     Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##     k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.7694 -0.2991  0.0121  0.3033  3.1562
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.24176    0.05479   22.67  < 2e-16 ***
## FirstAuthorFemale1 -0.01824    0.00999   -1.83   0.0677 .
## LastAuthorFemale1  -0.03139    0.01150   -2.73   0.0063 **
## UniqueAuthors2      0.33723    0.02260   14.92  < 2e-16 ***
## UniqueAuthors3      0.37714    0.02252   16.75  < 2e-16 ***
## UniqueAuthors4      0.42835    0.02309   18.55  < 2e-16 ***
## UniqueAuthors5      0.52769    0.02167   24.36  < 2e-16 ***
## Year1997           -0.23386    0.05854   -3.99  6.5e-05 ***
```

```
## Year1998           -0.46771    0.05453   -8.58  < 2e-16 ***
## Year1999           -0.47012    0.05456   -8.62  < 2e-16 ***
## Year2000           -0.45856    0.05554   -8.26  < 2e-16 ***
## Year2001           -0.49567    0.05617   -8.82  < 2e-16 ***
## Year2002           -0.51310    0.05556   -9.23  < 2e-16 ***
## Year2003           -0.47742    0.05549   -8.60  < 2e-16 ***
## Year2004           -0.49310    0.05425   -9.09  < 2e-16 ***
## Year2005           -0.50168    0.05428   -9.24  < 2e-16 ***
## Year2006           -0.46929    0.05403   -8.69  < 2e-16 ***
## Year2007           -0.48418    0.05399   -8.97  < 2e-16 ***
## Year2008           -0.49764    0.05383   -9.24  < 2e-16 ***
## Year2009           -0.47251    0.05371   -8.80  < 2e-16 ***
## Year2010           -0.44723    0.05367   -8.33  < 2e-16 ***
## Year2011           -0.53795    0.05347  -10.06  < 2e-16 ***
## Year2012           -0.52986    0.05404   -9.80  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.451
## Multiple R-squared:  0.133,  Adjusted R-squared:  0.131
## Convergence in 25 IRWLS iterations
##
## Robustness weights:
##  4 observations c(535,650,656,889)
##   are outliers with |weight| = 0 ( < 9.4e-06);
##  947 weights are ~= 1. The remaining 9705 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0013  0.8550  0.9500  0.8890  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          9.38e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.015  1           1.007
## LastAuthorFemale  1.008  1           1.004
## Year              1.019 16           1.001
```



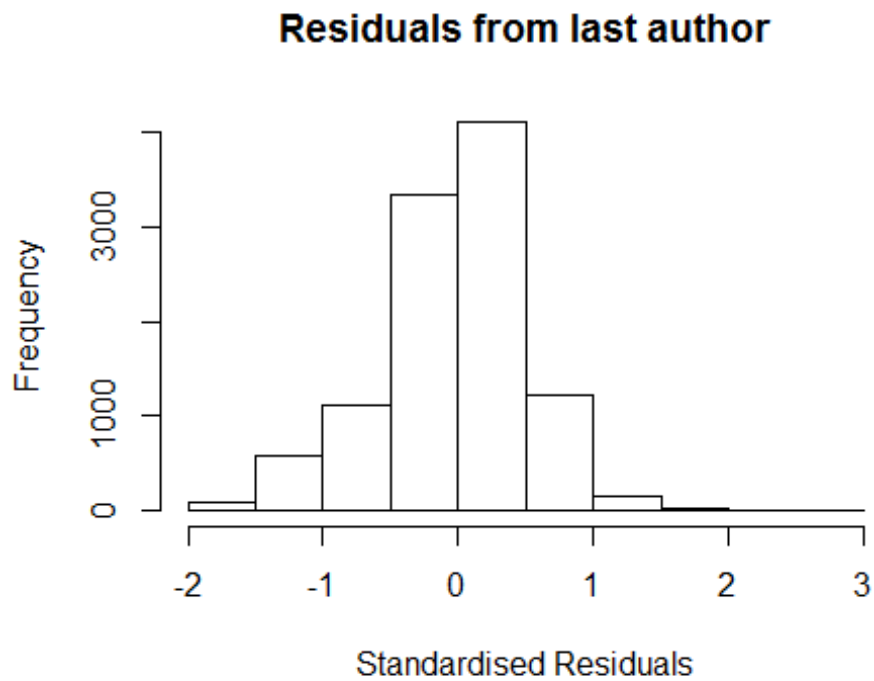

```
## [1] "List of  3 outliers with residuals above  2.5"
##        ScopusId  NLCS Year OneField Fields residuals
## 2429 0029980777 4.398 1996     2731      3     2.829
## 2663 0030779611 4.287 1997     2731      2     2.954
## 2670 0030832881 4.070 1997     2731      2     2.737
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.5820 -0.3146  0.0226  0.3159  2.9541
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)          1.5690     0.0560   28.01  < 2e-16 ***
## FirstAuthorFemale1   0.0130     0.0104    1.26  0.20924
## LastAuthorFemale1   -0.0309     0.0120   -2.57  0.01014 *
## Year1997            -0.2361     0.0639   -3.70  0.00022 ***
## Year1998            -0.4696     0.0592   -7.93  2.4e-15 ***
## Year1999            -0.4586     0.0592   -7.75  9.9e-15 ***
## Year2000            -0.4527     0.0603   -7.51  6.6e-14 ***
## Year2001            -0.4807     0.0606   -7.93  2.4e-15 ***
## Year2002            -0.4677     0.0603   -7.75  9.7e-15 ***
## Year2003            -0.4530     0.0598   -7.58  3.8e-14 ***
```

```
## Year2004            -0.4470     0.0588   -7.60  3.3e-14 ***
## Year2005            -0.4417     0.0588   -7.51  6.3e-14 ***
## Year2006            -0.3954     0.0583   -6.79  1.2e-11 ***
## Year2007            -0.4220     0.0583   -7.24  4.8e-13 ***
## Year2008            -0.4311     0.0583   -7.40  1.5e-13 ***
## Year2009            -0.4075     0.0580   -7.02  2.3e-12 ***
## Year2010            -0.3704     0.0580   -6.39  1.7e-10 ***
## Year2011            -0.4620     0.0579   -7.97  1.7e-15 ***
## Year2012            -0.4718     0.0584   -8.08  7.2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.464
## Multiple R-squared:  0.046,  Adjusted R-squared:  0.0444
## Convergence in 27 IRWLS iterations
##
## Robustness weights:
##  4 observations c(535,650,656,889)
##   are outliers with |weight| = 0 ( < 9.4e-06);
##  898 weights are ~= 1. The remaining 9754 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0001  0.8560  0.9490  0.8870  0.9860  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          9.38e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##                     GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.013  1           1.006
## Year              1.013 16           1.000
```




```
## [1] "List of  3 outliers with residuals above  2.5"
##        ScopusId  NLCS Year OneField Fields residuals
## 2429 0029980777 4.398 1996     2731      3     2.829
## 2663 0030779611 4.287 1997     2731      2     2.954
## 2670 0030832881 4.070 1997     2731      2     2.737
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.5726 -0.3128  0.0229  0.3183  2.9599
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.56287    0.05607   27.87  < 2e-16 ***
## FirstAuthorFemale1  0.00969    0.01041    0.93  0.35199
## Year1997           -0.23575    0.06398   -3.68  0.00023 ***
## Year1998           -0.46978    0.05934   -7.92  2.7e-15 ***
## Year1999           -0.45801    0.05928   -7.73  1.2e-14 ***
## Year2000           -0.45388    0.06039   -7.52  6.1e-14 ***
## Year2001           -0.47856    0.06070   -7.88  3.5e-15 ***
## Year2002           -0.46668    0.06039   -7.73  1.2e-14 ***
## Year2003           -0.45231    0.05988   -7.55  4.6e-14 ***
## Year2004           -0.44616    0.05893   -7.57  4.0e-14 ***
```

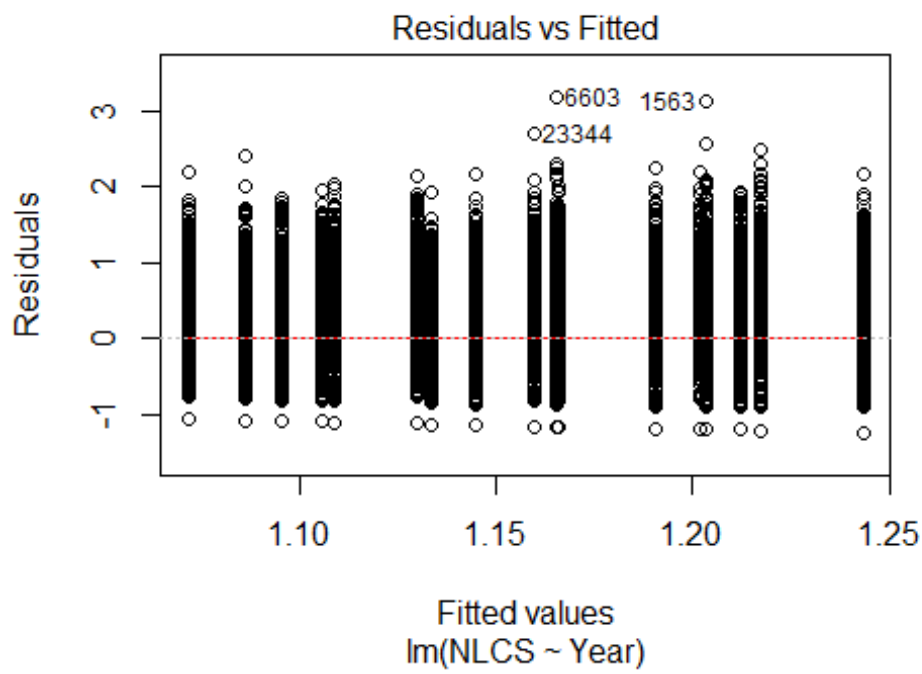
```
## Year2005           -0.44130    0.05887   -7.50  7.1e-14 ***
## Year2006           -0.39508    0.05835   -6.77  1.3e-11 ***
## Year2007           -0.42254    0.05838   -7.24  4.9e-13 ***
## Year2008           -0.43155    0.05836   -7.39  1.5e-13 ***
## Year2009           -0.40675    0.05813   -7.00  2.8e-12 ***
## Year2010           -0.37014    0.05807   -6.37  1.9e-10 ***
## Year2011           -0.46175    0.05802   -7.96  1.9e-15 ***
## Year2012           -0.47096    0.05847   -8.05  8.8e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.464
## Multiple R-squared:  0.0453, Adjusted R-squared:  0.0438
## Convergence in 27 IRWLS iterations
##
## Robustness weights:
##  4 observations c(535,650,656,889)
##   are outliers with |weight| = 0 ( < 9.4e-06);
##  896 weights are ~= 1. The remaining 9756 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.000   0.855   0.949   0.887   0.985   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          9.38e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                    psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                   "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##                   GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.007  1           1.003
## Year             1.007 16           1.000
```



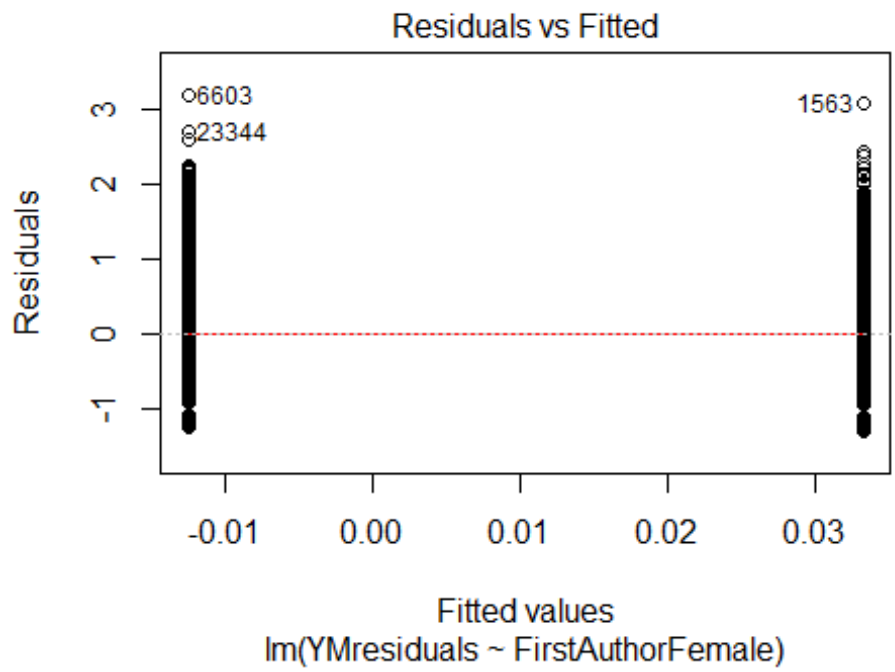

```
## [1] "List of  3 outliers with residuals above  2.5"
##        ScopusId  NLCS Year OneField Fields residuals
## 2429 0029980777 4.398 1996     2731      3     2.829
## 2663 0030779611 4.287 1997     2731      2     2.954
## 2670 0030832881 4.070 1997     2731      2     2.737
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.5729 -0.3151  0.0224  0.3148  2.9502
##
## Coefficients:
##                   Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.5729     0.0560   28.09  < 2e-16 ***
## LastAuthorFemale1  -0.0291     0.0120   -2.42  0.01550 *
## Year1997           -0.2361     0.0640   -3.69  0.00023 ***
## Year1998           -0.4703     0.0593   -7.92  2.5e-15 ***
## Year1999           -0.4591     0.0593   -7.75  1.0e-14 ***
## Year2000           -0.4535     0.0604   -7.51  6.6e-14 ***
## Year2001           -0.4807     0.0607   -7.91  2.8e-15 ***
## Year2002           -0.4685     0.0604   -7.75  9.7e-15 ***
## Year2003           -0.4533     0.0599   -7.57  4.1e-14 ***
## Year2004           -0.4471     0.0590   -7.58  3.7e-14 ***
```

```
## Year2005           -0.4417     0.0589   -7.50  7.0e-14 ***
## Year2006           -0.3949     0.0584   -6.76  1.4e-11 ***
## Year2007           -0.4221     0.0584   -7.23  5.3e-13 ***
## Year2008           -0.4307     0.0584   -7.37  1.8e-13 ***
## Year2009           -0.4073     0.0582   -7.00  2.7e-12 ***
## Year2010           -0.3703     0.0581   -6.37  2.0e-10 ***
## Year2011           -0.4613     0.0581   -7.94  2.2e-15 ***
## Year2012           -0.4715     0.0585   -8.06  8.8e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.464
## Multiple R-squared:  0.0459, Adjusted R-squared:  0.0444
## Convergence in 27 IRWLS iterations
##
## Robustness weights:
##  4 observations c(535,650,656,889)
##   are outliers with |weight| = 0 ( < 9.4e-06);
##  913 weights are ~= 1. The remaining 9739 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0002  0.8550  0.9480  0.8870  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          9.38e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Sample size for the above analysis:  10656"
## [1] ""
## [1] ""
## [1] "##################################"
## [1] "Analysis of AJSC 2900"
## [1] "##################################"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 1328 1486 1447 1556 1557 1483 1431 1085 1221 1461 1595 1815 1807 1315 1365
## 2011 2012
## 1420 1411
```

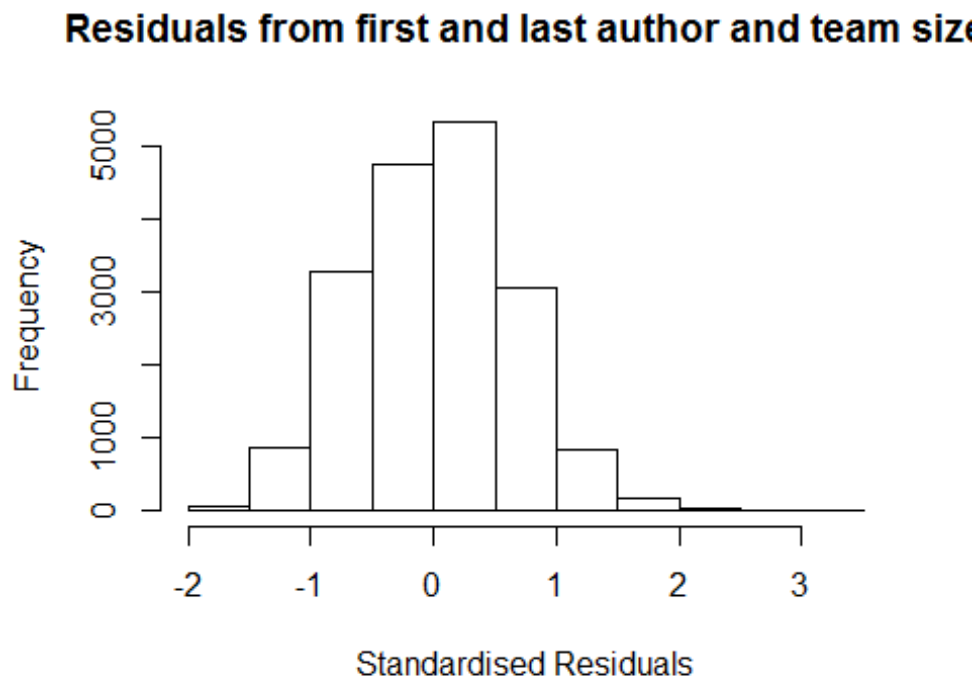
```
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  919  993  898 1086  947  943 1259  963 1059 1249 1380 1567 1603 1155 1193
## 2011 2012
## 1260 1227
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  881  932  829 1015  886  885 1175  907  986 1156 1253 1447 1487 1074 1113
## 2011 2012
## 1171 1139
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
##  Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 250, df = 16, p-value <2e-16
```




```
##
##  Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 37, df = 1, p-value = 1e-09
```



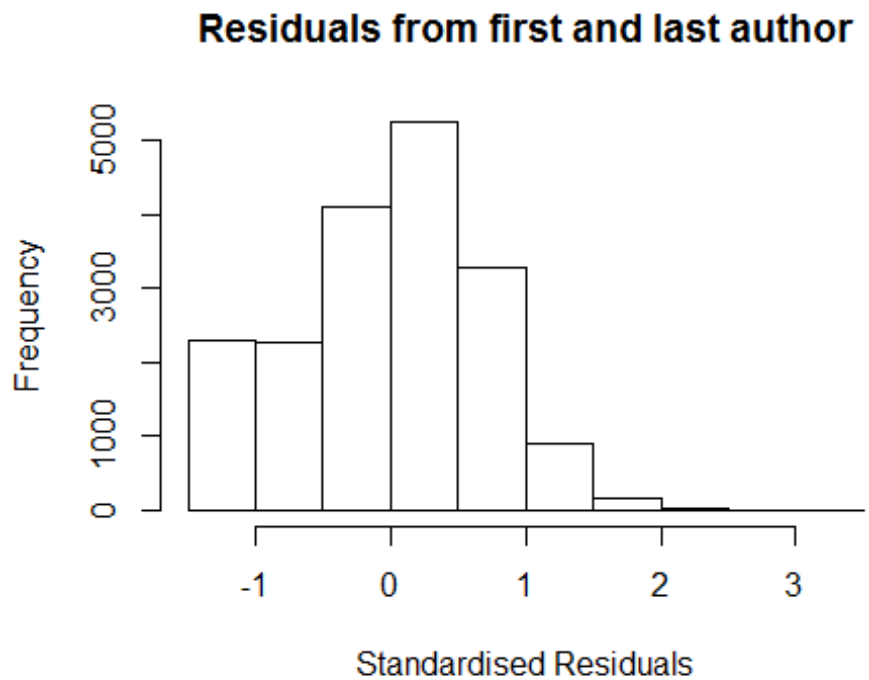

```
## [1] "Female first author team size 2018 geometric mean: 2.84540265155752"
## [1] "Male first author team size 2018 geometric mean: 3.22852871888201"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 130000, p-value = 0.005
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 2.72113681888949"
## [1] "Male last author team size 2018 geometric mean: 3.49931981044573"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 150000, p-value = 4e-09
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##                     GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.313  1           1.146
## LastAuthorFemale  1.379  1           1.174
## UniqueAuthors     1.101  4           1.012
## Year              1.074 16           1.002
```




```
## [1] "List of  5 outliers with residuals above  2.5"
##          ScopusId  NLCS Year OneField Fields residuals
## 1459   0031554094 3.775 1997     2900      1     2.826
## 1563   0030698732 4.324 1997     2900      1     2.693
## 6603   0034241415 4.339 2000     2900      1     3.451
## 9035   0035460243 3.430 2001     2900      1     2.510
## 23344 74349101245 3.859 2010     2900      1     3.024
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
UniqueAuthors +
##     Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##     k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.7105 -0.4556  0.0243  0.4477  3.4508
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)           1.0179     0.0297   34.27  < 2e-16 ***
## FirstAuthorFemale1    0.0356     0.0128    2.78  0.00536 **
## LastAuthorFemale1    -0.0302     0.0122   -2.48  0.01314 *
## UniqueAuthors2        0.2807     0.0138   20.39  < 2e-16 ***
## UniqueAuthors3        0.4381     0.0153   28.57  < 2e-16 ***
## UniqueAuthors4        0.5339     0.0165   32.38  < 2e-16 ***
```

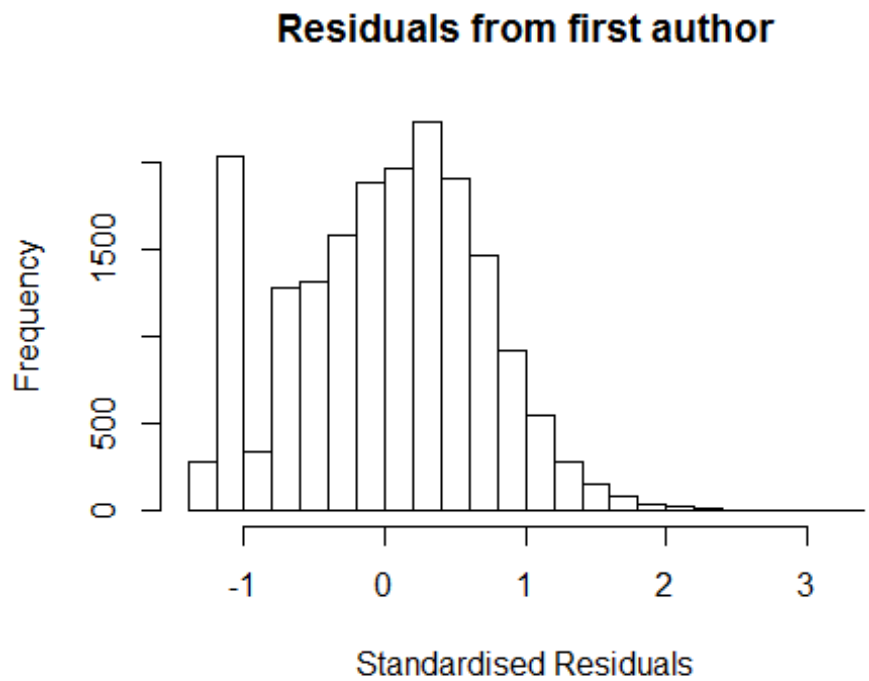
```
## UniqueAuthors5       0.6871     0.0142   48.26  < 2e-16 ***
## Year1997            -0.0744     0.0363   -2.05  0.04053 *
## Year1998            -0.0713     0.0395   -1.81  0.07073 .
## Year1999            -0.0699     0.0363   -1.92  0.05432 .
## Year2000            -0.1351     0.0368   -3.67  0.00024 ***
## Year2001            -0.1028     0.0356   -2.89  0.00391 **
## Year2002            -0.2142     0.0331   -6.48  9.3e-11 ***
## Year2003            -0.2054     0.0341   -6.02  1.8e-09 ***
## Year2004            -0.2137     0.0329   -6.49  8.7e-11 ***
## Year2005            -0.2538     0.0322   -7.88  3.4e-15 ***
## Year2006            -0.2595     0.0314   -8.27  < 2e-16 ***
## Year2007            -0.2378     0.0313   -7.60  3.2e-14 ***
## Year2008            -0.2714     0.0313   -8.67  < 2e-16 ***
## Year2009            -0.1564     0.0339   -4.61  4.1e-06 ***
## Year2010            -0.1881     0.0330   -5.69  1.3e-08 ***
## Year2011            -0.2203     0.0338   -6.51  7.5e-11 ***
## Year2012            -0.1717     0.0338   -5.09  3.7e-07 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.664
## Multiple R-squared:  0.141,  Adjusted R-squared:  0.14
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  observation 3859 is an outlier with |weight| = 0 ( < 5.5e-06);
##  1552 weights are ~= 1. The remaining 16783 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0032  0.8710  0.9510  0.9140  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          5.45e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.263  1           1.124
```

```
## LastAuthorFemale  1.267  1           1.125
## Year              1.036 16           1.001
```



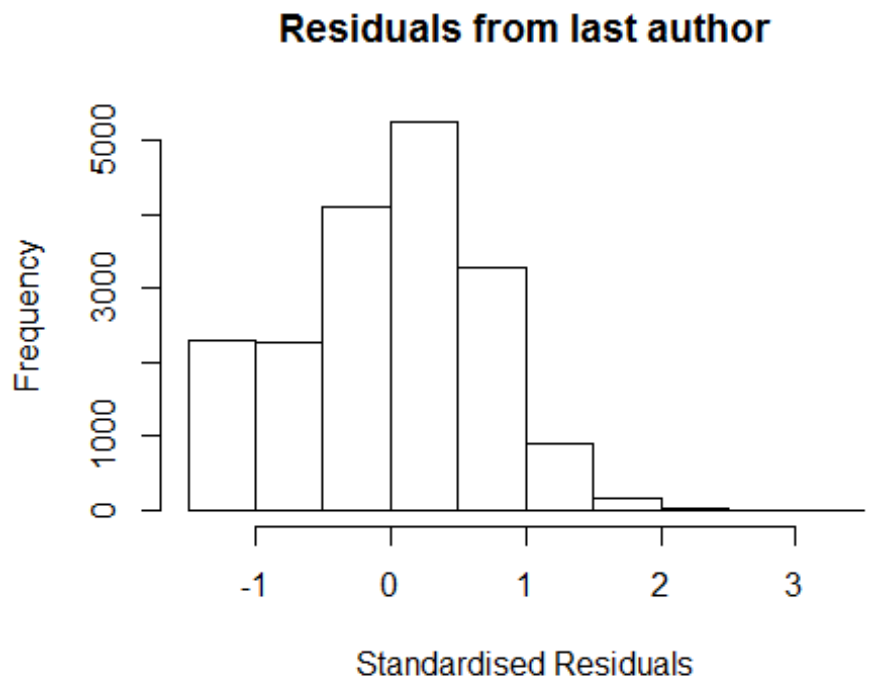

```
## [1] "List of  6 outliers with residuals above  2.5"
##          ScopusId  NLCS Year OneField Fields residuals
## 1459   0031554094 3.775 1997     2900      1     2.628
## 1563   0030698732 4.324 1997     2900      1     3.077
## 3259   0031686381 3.699 1998     2900      1     2.595
## 6603   0034241415 4.339 2000     2900      1     3.241
## 15768 33745233024 3.494 2006     2900      1     2.512
## 23344 74349101245 3.859 2010     2900      1     2.738
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.3609 -0.4913  0.0511  0.4808  3.2410
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)           1.3093     0.0294   44.53  < 2e-16 ***
## FirstAuthorFemale1    0.0516     0.0141    3.67  0.00024 ***
## LastAuthorFemale1    -0.1521     0.0128  -11.86  < 2e-16 ***
## Year1997             -0.0620     0.0374   -1.66  0.09739 .
```

```
## Year1998            -0.0529     0.0411   -1.29  0.19785
## Year1999            -0.0597     0.0378   -1.58  0.11380
## Year2000            -0.1108     0.0378   -2.93  0.00336 **
## Year2001            -0.0619     0.0369   -1.68  0.09311 .
## Year2002            -0.1557     0.0337   -4.62  3.9e-06 ***
## Year2003            -0.1135     0.0351   -3.24  0.00121 **
## Year2004            -0.1131     0.0339   -3.34  0.00084 ***
## Year2005            -0.1593     0.0334   -4.77  1.8e-06 ***
## Year2006            -0.1752     0.0321   -5.45  5.0e-08 ***
## Year2007            -0.1533     0.0319   -4.81  1.5e-06 ***
## Year2008            -0.1896     0.0323   -5.87  4.4e-09 ***
## Year2009            -0.0865     0.0352   -2.46  0.01391 *
## Year2010            -0.0879     0.0345   -2.55  0.01078 *
## Year2011            -0.1291     0.0348   -3.71  0.00021 ***
## Year2012            -0.0502     0.0347   -1.44  0.14872
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.701
## Multiple R-squared:  0.0124, Adjusted R-squared:  0.0114
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  1526 weights are ~= 1. The remaining 16810 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0007  0.8660  0.9470  0.9120  0.9840  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          5.45e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.022  1           1.011
## Year              1.022 16           1.001
```




```
## [1] "List of  6 outliers with residuals above  2.5"
##          ScopusId  NLCS Year OneField Fields residuals
## 1459   0031554094 3.775 1997     2900      1     2.628
## 1563   0030698732 4.324 1997     2900      1     3.077
## 3259   0031686381 3.699 1998     2900      1     2.595
## 6603   0034241415 4.339 2000     2900      1     3.241
## 15768 33745233024 3.494 2006     2900      1     2.512
## 23344 74349101245 3.859 2010     2900      1     2.738
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.2646 -0.4990  0.0453  0.4842  3.2142
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)           1.2646     0.0286   44.23  < 2e-16 ***
## FirstAuthorFemale1   -0.0346     0.0126   -2.75  0.00605 **
## Year1997             -0.0590     0.0372   -1.58  0.11347
## Year1998             -0.0498     0.0409   -1.22  0.22365
## Year1999             -0.0553     0.0377   -1.47  0.14193
## Year2000             -0.1052     0.0375   -2.80  0.00507 **
## Year2001             -0.0581     0.0367   -1.58  0.11334
```

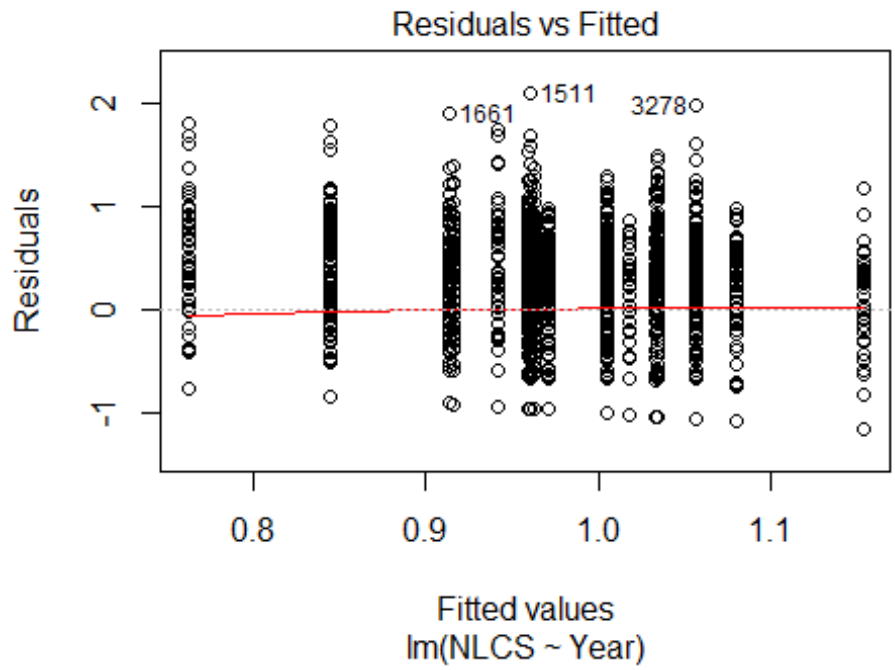
```
## Year2002            -0.1462     0.0335   -4.37  1.3e-05 ***
## Year2003            -0.1032     0.0348   -2.96  0.00307 **
## Year2004            -0.1049     0.0336   -3.12  0.00180 **
## Year2005            -0.1497     0.0331   -4.52  6.3e-06 ***
## Year2006            -0.1673     0.0319   -5.24  1.6e-07 ***
## Year2007            -0.1416     0.0317   -4.47  8.0e-06 ***
## Year2008            -0.1832     0.0322   -5.70  1.2e-08 ***
## Year2009            -0.0878     0.0350   -2.50  0.01227 *
## Year2010            -0.0910     0.0343   -2.66  0.00793 **
## Year2011            -0.1254     0.0346   -3.62  0.00029 ***
## Year2012            -0.0492     0.0346   -1.42  0.15496
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.704
## Multiple R-squared:  0.00511,    Adjusted R-squared:  0.00418
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  1412 weights are ~= 1. The remaining 16924 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0026  0.8730  0.9480  0.9130  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          5.45e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##                   GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.025  1           1.013
## Year             1.025 16           1.001
```



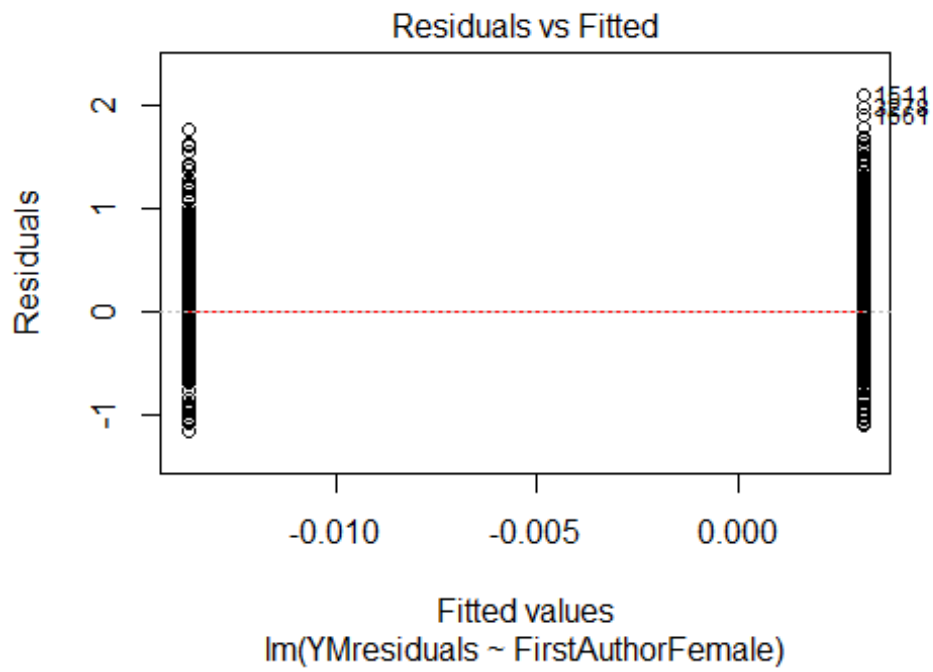

```
## [1] "List of  6 outliers with residuals above  2.5"
##          ScopusId  NLCS Year OneField Fields residuals
## 1459   0031554094 3.775 1997     2900      1     2.628
## 1563   0030698732 4.324 1997     2900      1     3.077
## 3259   0031686381 3.699 1998     2900      1     2.595
## 6603   0034241415 4.339 2000     2900      1     3.241
## 15768 33745233024 3.494 2006     2900      1     2.512
## 23344 74349101245 3.859 2010     2900      1     2.738
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.3307 -0.4982  0.0477  0.4809  3.2473
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)          1.3307     0.0287   46.41  < 2e-16 ***
## LastAuthorFemale1   -0.1264     0.0117  -10.84  < 2e-16 ***
## Year1997            -0.0631     0.0375   -1.68  0.09241 .
## Year1998            -0.0532     0.0411   -1.29  0.19636
## Year1999            -0.0604     0.0378   -1.60  0.11039
## Year2000            -0.1127     0.0378   -2.98  0.00289 **
## Year2001            -0.0624     0.0369   -1.69  0.09108 .
```

```
## Year2002           -0.1571     0.0338   -4.65  3.3e-06 ***
## Year2003           -0.1152     0.0351   -3.28  0.00104 **
## Year2004           -0.1153     0.0339   -3.40  0.00068 ***
## Year2005           -0.1602     0.0334   -4.79  1.7e-06 ***
## Year2006           -0.1781     0.0322   -5.54  3.1e-08 ***
## Year2007           -0.1551     0.0319   -4.86  1.2e-06 ***
## Year2008           -0.1904     0.0324   -5.88  4.2e-09 ***
## Year2009           -0.0852     0.0352   -2.42  0.01567 *
## Year2010           -0.0872     0.0345   -2.53  0.01157 *
## Year2011           -0.1264     0.0348   -3.63  0.00029 ***
## Year2012           -0.0485     0.0348   -1.39  0.16318
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.701
## Multiple R-squared:  0.0117, Adjusted R-squared:  0.0108
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  1531 weights are ~= 1. The remaining 16805 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0005  0.8660  0.9480  0.9120  0.9840  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          5.45e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Sample size for the above analysis:  18336"
## [1] ""
## [1] ""
## [1] "####################################"
## [1] "Analysis of AJSC 2901"
## [1] "####################################"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##   57  100   68   88  142  146  144  197  181  175  208  173  235  303  323
## 2011 2012
```

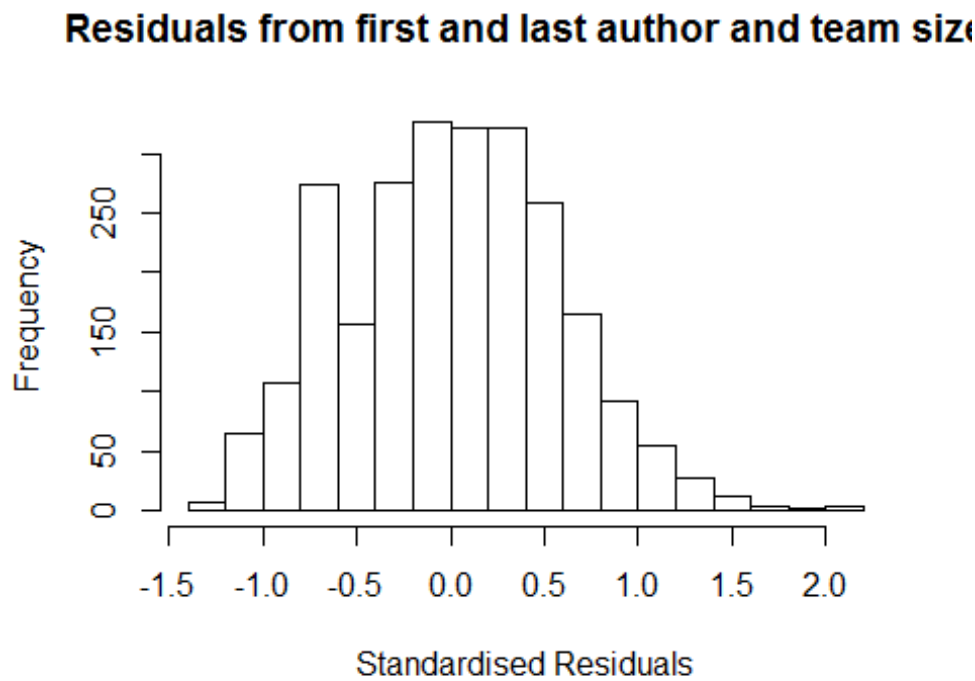
```
##  332  330
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##   36   71   47   55   95   97  121  178  155  161  181  148  208  260  276
## 2011 2012
##  284  283
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##   35   68   45   50   90   93  108  164  145  145  168  139  197  242  258
## 2011 2012
##  260  263
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
##  Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 32, df = 16, p-value = 0.01
```




```
##
##  Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 0.55, df = 1, p-value = 0.5
```

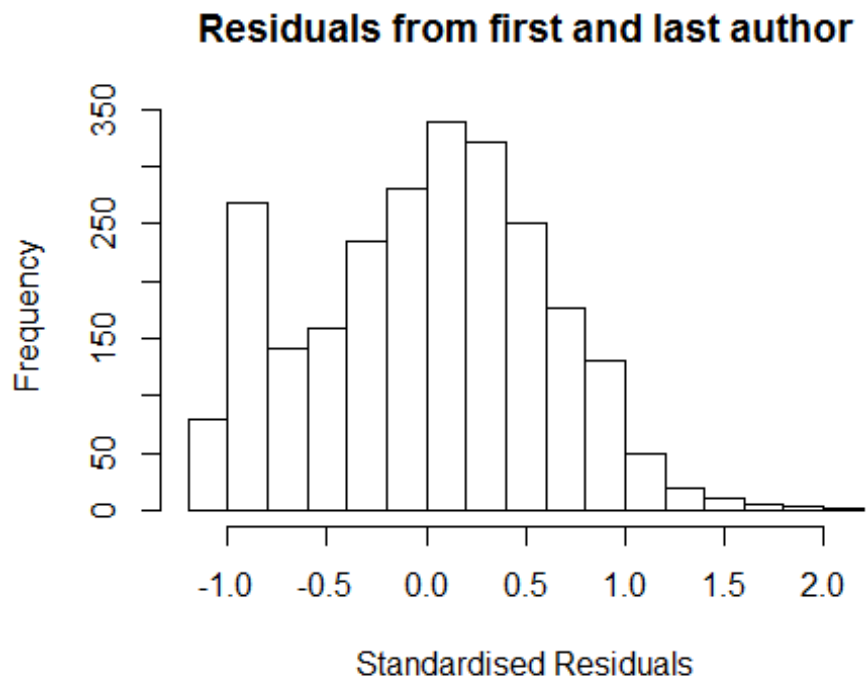



```
## [1] "Female first author team size 2018 geometric mean: 3.13702524642977"
## [1] "Male first author team size 2018 geometric mean: 3.95375742815858"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 5000, p-value = 0.03
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 2.95084307724201"
## [1] "Male last author team size 2018 geometric mean: 4.50907327640181"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 5000, p-value = 2e-06
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.177  1           1.085
## LastAuthorFemale  1.159  1           1.077
## UniqueAuthors     1.209  4           1.024
## Year              1.259 16           1.007
```



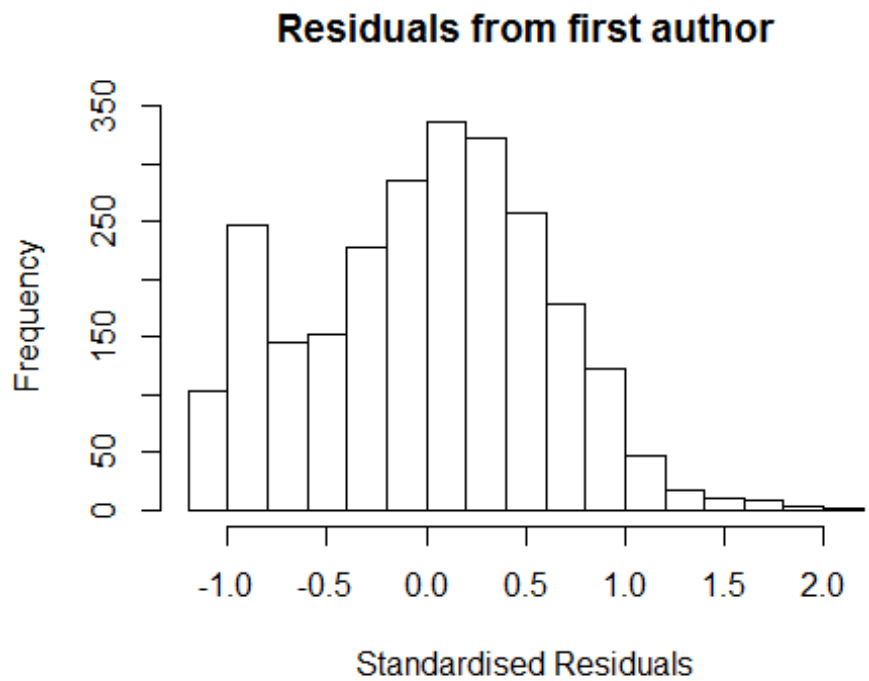

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
UniqueAuthors +
##     Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##     k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.3843 -0.3940  0.0151  0.3991  2.1029
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)          0.83053    0.09898    8.39  < 2e-16 ***
## FirstAuthorFemale1   0.03685    0.03238    1.14   0.2552
## LastAuthorFemale1    0.00902    0.02833    0.32   0.7503
## UniqueAuthors2       0.22673    0.03503    6.47  1.2e-10 ***
## UniqueAuthors3       0.39931    0.03668   10.89  < 2e-16 ***
## UniqueAuthors4       0.42636    0.04181   10.20  < 2e-16 ***
## UniqueAuthors5       0.44327    0.03565   12.43  < 2e-16 ***
## Year1997             0.14539    0.12021    1.21   0.2266
## Year1998             0.04952    0.13326    0.37   0.7102
## Year1999             0.14744    0.11683    1.26   0.2071
```

```
## Year2000           -0.17931    0.12504   -1.43   0.1517
## Year2001           -0.35806    0.11618   -3.08   0.0021 **
## Year2002           -0.18840    0.11105   -1.70   0.0899 .
## Year2003           -0.10584    0.10258   -1.03   0.3023
## Year2004           -0.13589    0.10975   -1.24   0.2158
## Year2005           -0.17221    0.10500   -1.64   0.1011
## Year2006           -0.14907    0.10347   -1.44   0.1498
## Year2007           -0.23689    0.10886   -2.18   0.0296 *
## Year2008           -0.24224    0.10510   -2.30   0.0213 *
## Year2009           -0.10357    0.10143   -1.02   0.3073
## Year2010           -0.16066    0.09978   -1.61   0.1075
## Year2011           -0.10508    0.10100   -1.04   0.2983
## Year2012           -0.11575    0.10070   -1.15   0.2505
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.589
## Multiple R-squared:  0.107,  Adjusted R-squared:  0.0988
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  197 weights are ~= 1. The remaining 2273 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.175   0.871   0.950   0.912   0.986   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          4.05e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.145  1           1.070
## LastAuthorFemale  1.131  1           1.063
## Year              1.068 16           1.002
```



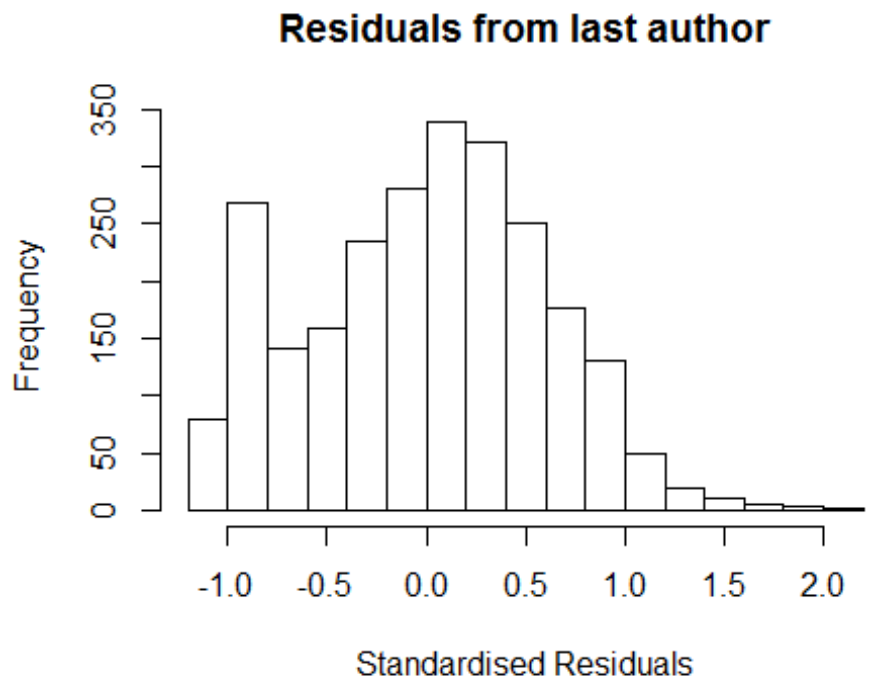

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.1860 -0.4392  0.0395  0.4239  2.0969
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)           1.0358     0.1021   10.14   <2e-16 ***
## FirstAuthorFemale1    0.0414     0.0344    1.20   0.2294
## LastAuthorFemale1    -0.0582     0.0294   -1.98   0.0482 *
## Year1997              0.0870     0.1186    0.73   0.4632
## Year1998              0.0215     0.1346    0.16   0.8731
## Year1999              0.1502     0.1169    1.28   0.1991
## Year2000             -0.1184     0.1258   -0.94   0.3466
## Year2001             -0.3301     0.1219   -2.71   0.0068 **
## Year2002             -0.1302     0.1139   -1.14   0.2531
## Year2003             -0.0457     0.1058   -0.43   0.6660
## Year2004             -0.0626     0.1135   -0.55   0.5814
## Year2005             -0.0901     0.1098   -0.82   0.4120
```

```
## Year2006            -0.0589     0.1073   -0.55   0.5830
## Year2007            -0.1414     0.1141   -1.24   0.2153
## Year2008            -0.1935     0.1101   -1.76   0.0788 .
## Year2009            -0.0243     0.1047   -0.23   0.8164
## Year2010            -0.0693     0.1032   -0.67   0.5021
## Year2011             0.0040     0.1043    0.04   0.9694
## Year2012             0.0191     0.1040    0.18   0.8540
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.615
## Multiple R-squared:  0.022,  Adjusted R-squared:  0.0148
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  229 weights are ~= 1. The remaining 2241 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.221   0.849   0.946   0.911   0.985   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          4.05e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.045  1           1.022
## Year              1.045 16           1.001
```




```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.1597 -0.4341  0.0425  0.4202  2.0849
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)          1.00989    0.10037   10.06   <2e-16 ***
## FirstAuthorFemale1   0.01769    0.03263    0.54   0.5878
## Year1997             0.08803    0.11815    0.75   0.4563
## Year1998             0.02295    0.13482    0.17   0.8648
## Year1999             0.14977    0.11649    1.29   0.1987
## Year2000            -0.11571    0.12500   -0.93   0.3547
## Year2001            -0.32914    0.12109   -2.72   0.0066 **
## Year2002            -0.12565    0.11332   -1.11   0.2676
## Year2003            -0.04778    0.10516   -0.45   0.6497
## Year2004            -0.06050    0.11289   -0.54   0.5920
## Year2005            -0.08689    0.10914   -0.80   0.4260
## Year2006            -0.05548    0.10658   -0.52   0.6027
```

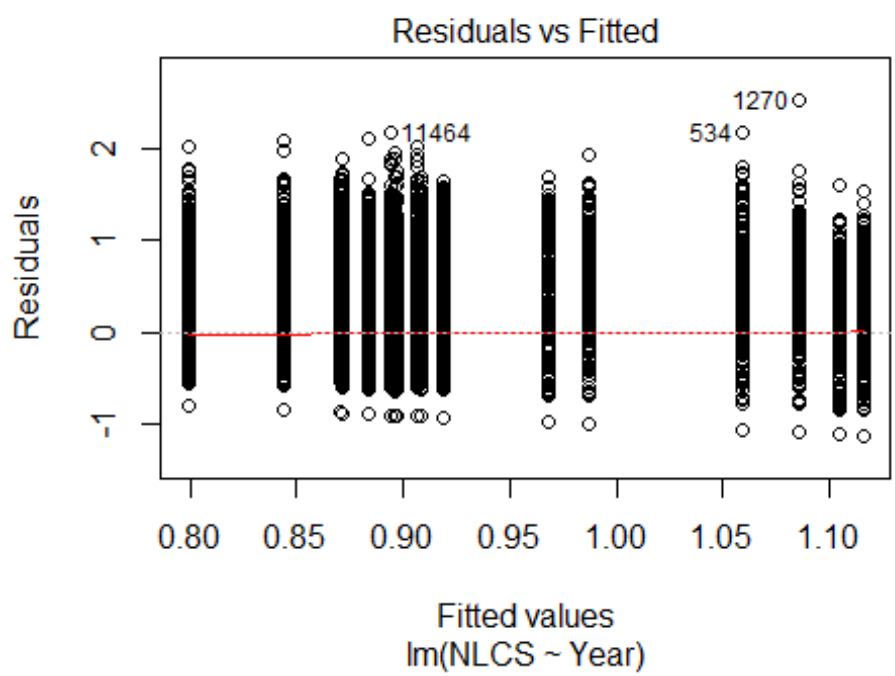
```
## Year2007           -0.13955    0.11364   -1.23   0.2196
## Year2008           -0.19358    0.10945   -1.77   0.0771 .
## Year2009           -0.01976    0.10407   -0.19   0.8494
## Year2010           -0.06638    0.10259   -0.65   0.5177
## Year2011            0.00325    0.10361    0.03   0.9750
## Year2012            0.02125    0.10328    0.21   0.8370
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.615
## Multiple R-squared:  0.0207, Adjusted R-squared:  0.0139
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  224 weights are ~= 1. The remaining 2246 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.226   0.850   0.947   0.911   0.986   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          4.05e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##                   GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.029  1           1.014
## Year             1.029 16           1.001
```



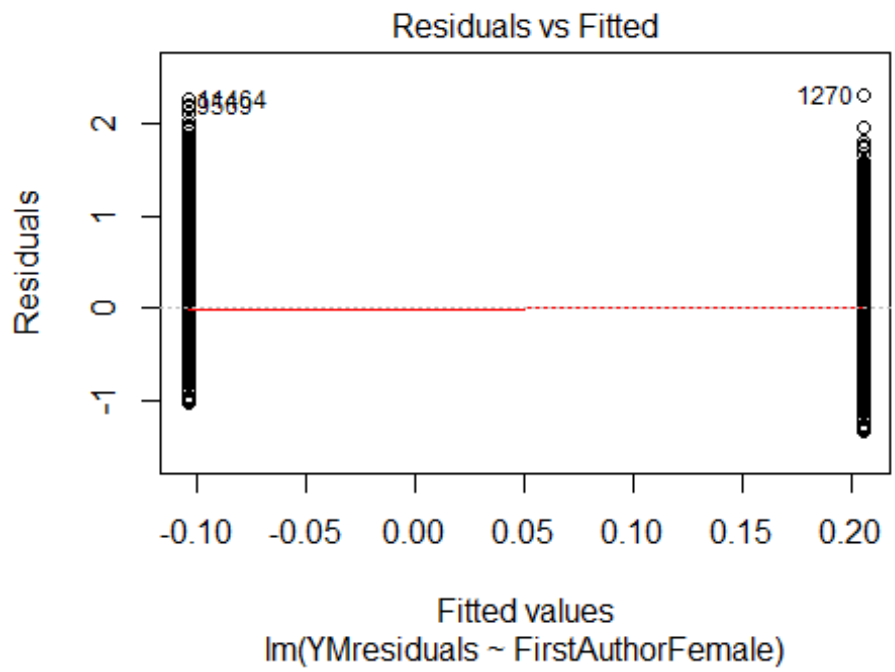

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.2142 -0.4398  0.0399  0.4291  2.1042
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.06096    0.09874   10.75   <2e-16 ***
## LastAuthorFemale1 -0.04500    0.02830   -1.59   0.1119
## Year1997            0.08481    0.11820    0.72   0.4731
## Year1998            0.01892    0.13448    0.14   0.8881
## Year1999            0.15323    0.11657    1.31   0.1888
## Year2000           -0.12001    0.12518   -0.96   0.3378
## Year2001           -0.33564    0.12121   -2.77   0.0057 **
## Year2002           -0.12912    0.11349   -1.14   0.2554
## Year2003           -0.04937    0.10514   -0.47   0.6387
## Year2004           -0.06216    0.11298   -0.55   0.5823
## Year2005           -0.08871    0.10919   -0.81   0.4167
## Year2006           -0.06316    0.10662   -0.59   0.5536
```

```
## Year2007          -0.14210    0.11359   -1.25   0.2111
## Year2008          -0.19509    0.10939   -1.78   0.0746 .
## Year2009          -0.02597    0.10412   -0.25   0.8031
## Year2010          -0.07085    0.10257   -0.69   0.4898
## Year2011           0.00365    0.10368    0.04   0.9719
## Year2012           0.01950    0.10342    0.19   0.8505
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.615
## Multiple R-squared:  0.0215, Adjusted R-squared:  0.0147
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  219 weights are ~= 1. The remaining 2251 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.218   0.850   0.946   0.912   0.985   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          4.05e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Sample size for the above analysis:  2470"
## [1] ""
## [1] ""
## [1] "####################################"
## [1] "Analysis of AJSC 2902"
## [1] "####################################"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  522  761  734  709  797  716  976  941  841  809  653  695  917 1099 1033
## 2011 2012
## 1060  983
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  365  392  429  373  374  393  659  696  708  676  545  605  784  947  869
## 2011 2012
```

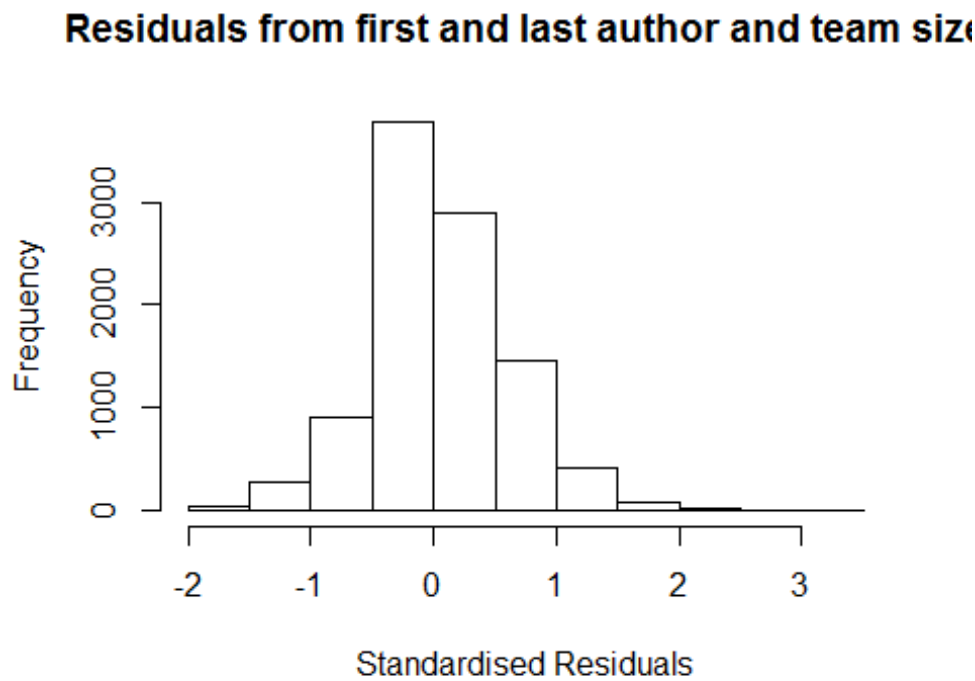
```
##  902  831
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  337  358  394  348  349  366  622  655  679  636  520  570  726  883  804
## 2011 2012
##  827  764
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
##  Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 53, df = 16, p-value = 9e-06
```




```
##
##  Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 18, df = 1, p-value = 2e-05
```

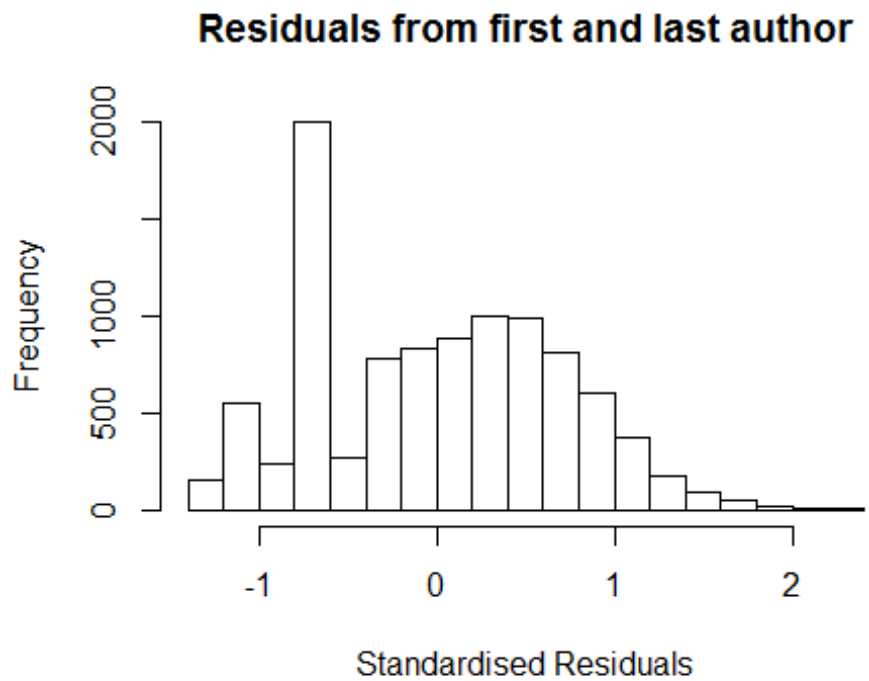



```
## [1] "Female first author team size 2018 geometric mean: 2.41997084350374"
## [1] "Male first author team size 2018 geometric mean: 3.11307903261214"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 36000, p-value = 2e-04
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 2.17761626911899"
## [1] "Male last author team size 2018 geometric mean: 3.5797839356303"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 32000, p-value = 2e-14
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.627  1           1.276
## LastAuthorFemale  1.889  1           1.375
## UniqueAuthors     1.460  4           1.048
## Year              1.103 16           1.003
```



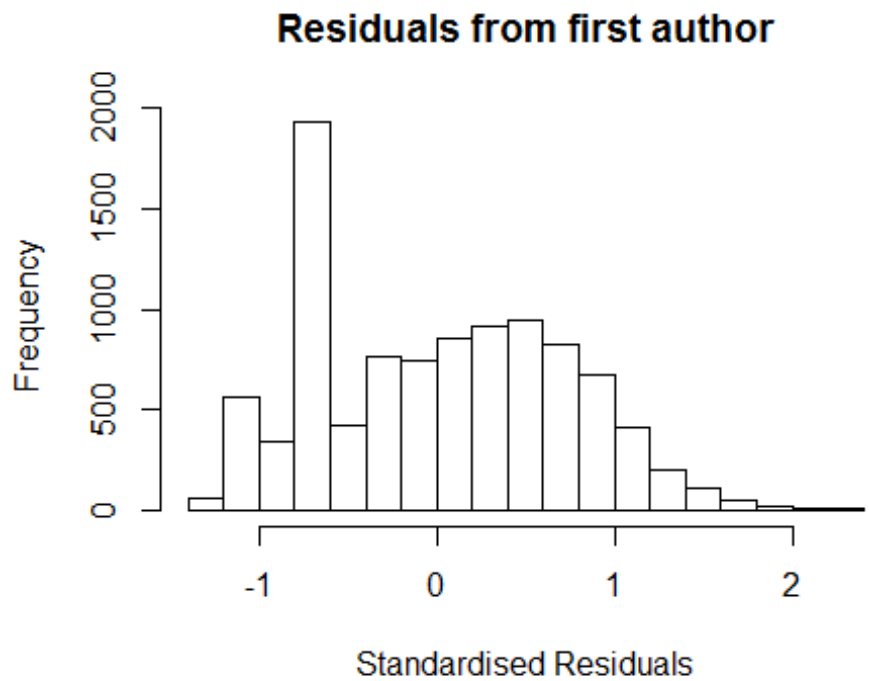

```
## [1] "List of  2 outliers with residuals above  2.5"
##          ScopusId  NLCS Year OneField Fields residuals
## 1270   1842334456 3.603 1997     2712      3     3.083
## 11464 67650084651 3.058 2009     2902      4     2.685
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
UniqueAuthors +
##     Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##     k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.6812 -0.3783 -0.0109  0.3959  3.0834
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         0.53688    0.03907   13.74  < 2e-16 ***
## FirstAuthorFemale1 -0.01868    0.01564   -1.19  0.23233
## LastAuthorFemale1  -0.02213    0.01633   -1.36  0.17540
## UniqueAuthors2      0.32534    0.02081   15.64  < 2e-16 ***
## UniqueAuthors3      0.70188    0.02710   25.90  < 2e-16 ***
## UniqueAuthors4      0.92658    0.02651   34.96  < 2e-16 ***
## UniqueAuthors5      1.15359    0.01690   68.25  < 2e-16 ***
## Year1997           -0.01730    0.05028   -0.34  0.73076
## Year1998           -0.06761    0.04840   -1.40  0.16245
```

```
## Year1999           -0.09505    0.05007   -1.90  0.05768 .
## Year2000           -0.00925    0.04954   -0.19  0.85189
## Year2001           -0.03931    0.04605   -0.85  0.39333
## Year2002           -0.06446    0.04418   -1.46  0.14463
## Year2003           -0.10462    0.04263   -2.45  0.01414 *
## Year2004           -0.15207    0.04232   -3.59  0.00033 ***
## Year2005           -0.11779    0.04298   -2.74  0.00614 **
## Year2006           -0.13577    0.04370   -3.11  0.00190 **
## Year2007           -0.16996    0.04348   -3.91  9.3e-05 ***
## Year2008           -0.13201    0.04195   -3.15  0.00165 **
## Year2009           -0.12298    0.04121   -2.98  0.00285 **
## Year2010           -0.11702    0.04223   -2.77  0.00559 **
## Year2011           -0.14914    0.04091   -3.65  0.00027 ***
## Year2012           -0.17238    0.04221   -4.08  4.5e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.496
## Multiple R-squared:  0.456,  Adjusted R-squared:  0.455
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
##  3 observations c(633,7069,7314) are outliers with |weight| = 0 ( < 1e-
05);
##  613 weights are ~= 1. The remaining 9222 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0098  0.8500  0.9440  0.8860  0.9730  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.02e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.442  1           1.201
## LastAuthorFemale  1.456  1           1.207
## Year              1.041 16           1.001
```



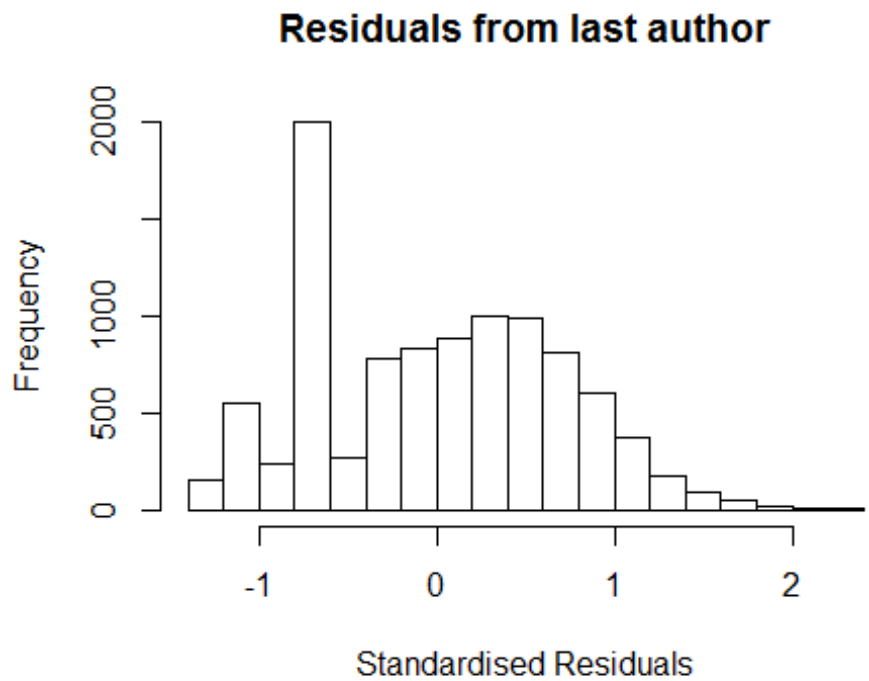

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.3520 -0.6568  0.0292  0.5363  2.3954
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)           1.2640     0.0481   26.30   <2e-16 ***
## FirstAuthorFemale1   -0.0717     0.0221   -3.25   0.0012 **
## LastAuthorFemale1    -0.4194     0.0205  -20.46   <2e-16 ***
## Year1997              0.0317     0.0630    0.50   0.6148
## Year1998             -0.0563     0.0614   -0.92   0.3594
## Year1999             -0.0746     0.0641   -1.16   0.2446
## Year2000              0.0880     0.0603    1.46   0.1448
## Year2001              0.0783     0.0563    1.39   0.1641
## Year2002             -0.0901     0.0564   -1.60   0.1102
## Year2003             -0.1550     0.0551   -2.81   0.0049 **
## Year2004             -0.1671     0.0539   -3.10   0.0019 **
## Year2005             -0.0887     0.0539   -1.64   0.1002
```

```
## Year2006            -0.0796     0.0563   -1.42   0.1570
## Year2007            -0.1330     0.0563   -2.36   0.0182 *
## Year2008            -0.1161     0.0542   -2.14   0.0321 *
## Year2009            -0.1104     0.0532   -2.08   0.0380 *
## Year2010            -0.0656     0.0535   -1.23   0.2199
## Year2011            -0.0616     0.0532   -1.16   0.2468
## Year2012            -0.1150     0.0541   -2.12   0.0336 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.704
## Multiple R-squared:  0.103,  Adjusted R-squared:  0.101
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
##  644 weights are ~= 1. The remaining 9194 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.223   0.887   0.926   0.911   0.979   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.02e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.025  1           1.012
## Year              1.025 16           1.001
```




```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.2944 -0.7247  0.0216  0.5651  2.3624
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)           1.1915     0.0480   24.80  < 2e-16 ***
## FirstAuthorFemale1   -0.3298     0.0180  -18.34  < 2e-16 ***
## Year1997              0.0491     0.0641    0.77  0.44349
## Year1998             -0.0588     0.0624   -0.94  0.34638
## Year1999             -0.0762     0.0645   -1.18  0.23698
## Year2000              0.0894     0.0614    1.46  0.14526
## Year2001              0.1029     0.0573    1.80  0.07227 .
## Year2002             -0.0948     0.0568   -1.67  0.09540 .
## Year2003             -0.1702     0.0553   -3.08  0.00210 **
## Year2004             -0.1965     0.0546   -3.60  0.00032 ***
## Year2005             -0.1027     0.0547   -1.88  0.06040 .
## Year2006             -0.0793     0.0574   -1.38  0.16726
```

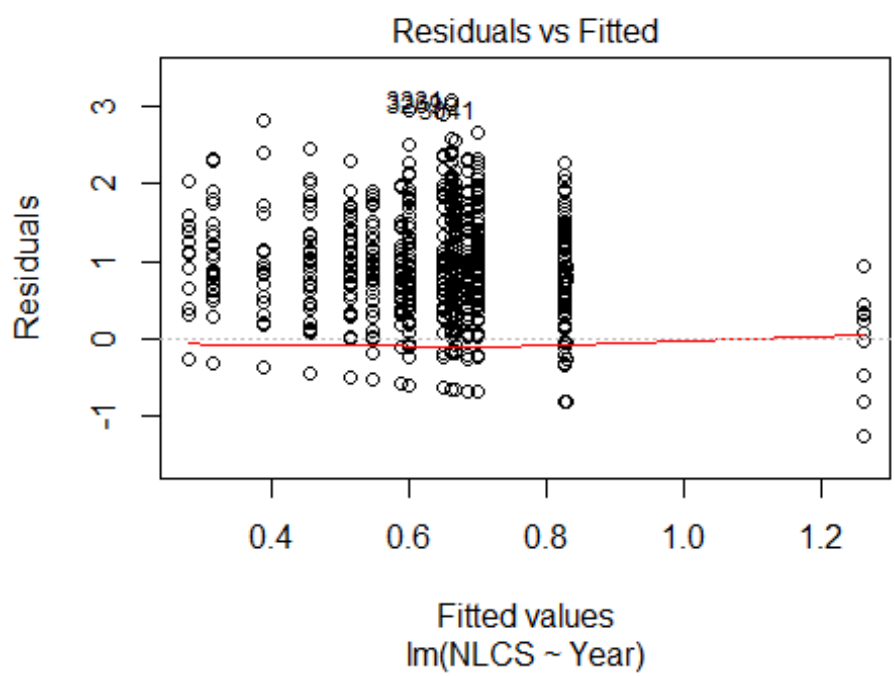
```
## Year2007            -0.1194     0.0573   -2.09  0.03709 *
## Year2008            -0.1175     0.0552   -2.13  0.03320 *
## Year2009            -0.1146     0.0541   -2.12  0.03421 *
## Year2010            -0.0764     0.0545   -1.40  0.16072
## Year2011            -0.0772     0.0541   -1.43  0.15357
## Year2012            -0.1284     0.0547   -2.35  0.01883 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.772
## Multiple R-squared:  0.0537, Adjusted R-squared:  0.052
## Convergence in 10 IRWLS iterations
##
## Robustness weights:
##  651 weights are ~= 1. The remaining 9187 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.329   0.901   0.930   0.922   0.980   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.02e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##                   GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.032  1           1.016
## Year             1.032 16           1.001
```



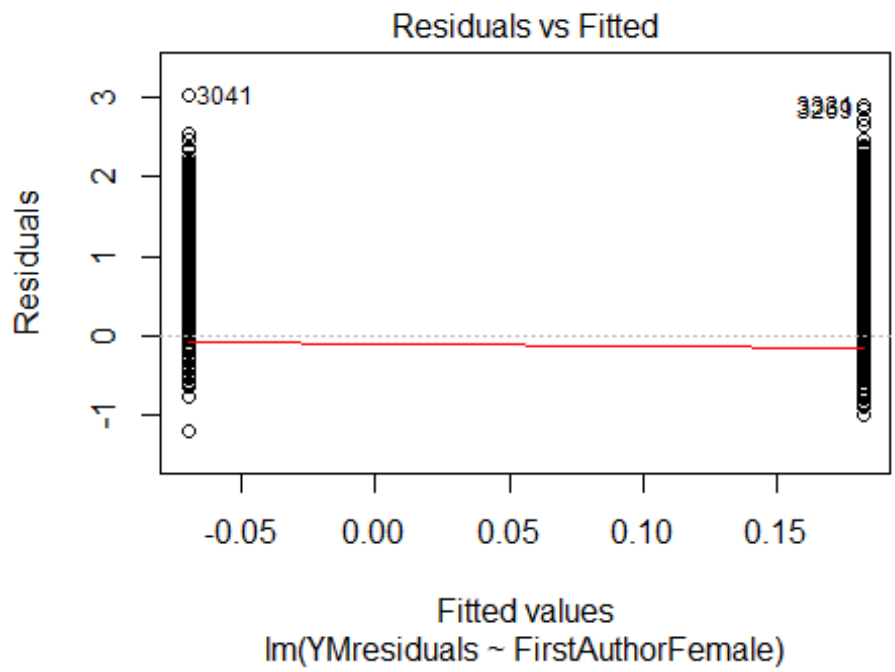

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.3308 -0.6627  0.0244  0.5356  2.3911
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.2449     0.0475   26.21   <2e-16 ***
## LastAuthorFemale1  -0.4594     0.0173  -26.62   <2e-16 ***
## Year1997            0.0275     0.0632    0.43   0.6640
## Year1998           -0.0589     0.0616   -0.96   0.3387
## Year1999           -0.0783     0.0644   -1.21   0.2244
## Year2000            0.0859     0.0604    1.42   0.1549
## Year2001            0.0740     0.0564    1.31   0.1897
## Year2002           -0.0974     0.0565   -1.72   0.0848 .
## Year2003           -0.1602     0.0550   -2.91   0.0036 **
## Year2004           -0.1738     0.0539   -3.22   0.0013 **
## Year2005           -0.0974     0.0539   -1.81   0.0709 .
## Year2006           -0.0871     0.0562   -1.55   0.1212
```

```
## Year2007           -0.1406     0.0563   -2.50   0.0125 *
## Year2008           -0.1228     0.0541   -2.27   0.0234 *
## Year2009           -0.1186     0.0531   -2.23   0.0257 *
## Year2010           -0.0739     0.0535   -1.38   0.1667
## Year2011           -0.0689     0.0532   -1.30   0.1951
## Year2012           -0.1224     0.0541   -2.26   0.0238 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.707
## Multiple R-squared:  0.101,  Adjusted R-squared:  0.0995
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
##  645 weights are ~= 1. The remaining 9193 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.230   0.887   0.927   0.912   0.979   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.02e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Sample size for the above analysis:  9838"
## [1] ""
## [1] ""
## [1] "####################################"
## [1] "Analysis of AJSC 2903"
## [1] "####################################"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  123  234  248  217  214   93  378  364  280  222  161  208  320  364  358
## 2011 2012
##  297  310
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##   93   87  110   81   11   11  162  174  213  162  144  179  291  337  320
## 2011 2012
```

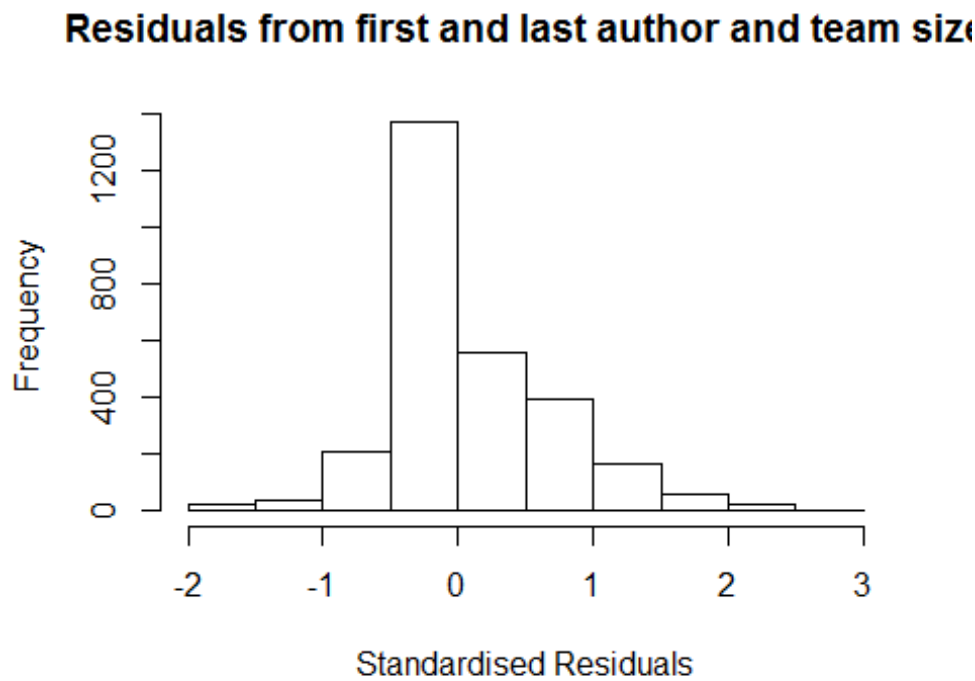
```
##  271  275
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##   92   83  107   81   11   10  160  169  211  156  141  174  286  316  305
## 2011 2012
##  256  265
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
##  Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 87, df = 16, p-value = 9e-12
```




```
##
##  Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 84, df = 1, p-value <2e-16
```



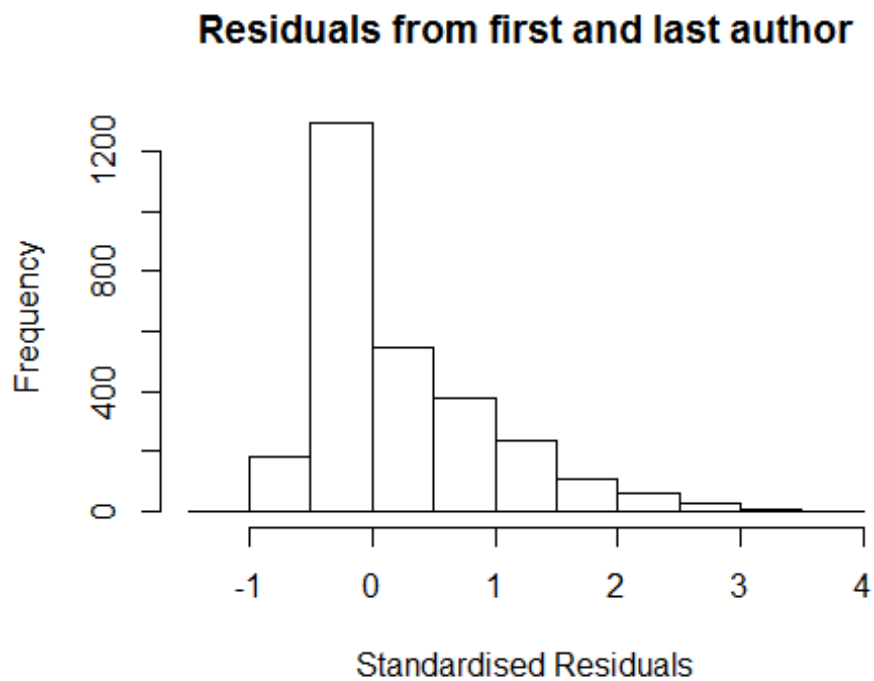

```
## [1] "Female first author team size 2018 geometric mean: 2.12017289766673"
## [1] "Male first author team size 2018 geometric mean: 3.1134521520173"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 3900, p-value = 0.001
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 1.96914853266314"
## [1] "Male last author team size 2018 geometric mean: 3.4028488489672"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 3700, p-value = 5e-07
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 8.749  1           2.958
## LastAuthorFemale  8.768  1           2.961
## UniqueAuthors     2.856  4           1.140
## Year              3.038 16           1.035
```




```
## [1] "List of  2 outliers with residuals above  2.5"
##         ScopusId  NLCS Year OneField Fields residuals
## 3385 67650084651 3.058 2009     2902      4     2.839
## 3436 60749130072 2.769 2009     2902      4     2.550
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
UniqueAuthors +
##     Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##     k.max = 1000))
##  \--> method = "MM"
## Residuals:
##    Min     1Q Median     3Q    Max
## -1.849 -0.230 -0.110  0.448  2.839
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)          0.1478     0.0456    3.24   0.0012 **
## FirstAuthorFemale1   0.0462     0.0658    0.70   0.4825
## LastAuthorFemale1   -0.0577     0.0653   -0.88   0.3767
## UniqueAuthors2       0.2883     0.0528    5.46  5.2e-08 ***
## UniqueAuthors3       0.5667     0.1034    5.48  4.6e-08 ***
## UniqueAuthors4       0.9807     0.1628    6.02  1.9e-09 ***
## UniqueAuthors5       1.5016     0.0779   19.27  < 2e-16 ***
## Year1997            -0.0964     0.0497   -1.94   0.0527 .
## Year1998            -0.0806     0.0483   -1.67   0.0950 .
```

```
## Year1999            -0.0379     0.0600   -0.63   0.5274
## Year2000            -0.0266     0.1316   -0.20   0.8401
## Year2001             0.6603     0.3533    1.87   0.0618 .
## Year2002             0.0620     0.0698    0.89   0.3745
## Year2003             0.0349     0.0488    0.72   0.4744
## Year2004             0.0641     0.0540    1.19   0.2356
## Year2005             0.1993     0.0628    3.18   0.0015 **
## Year2006             0.0473     0.0705    0.67   0.5021
## Year2007             0.0891     0.0586    1.52   0.1289
## Year2008             0.0576     0.0502    1.15   0.2514
## Year2009             0.0824     0.0502    1.64   0.1012
## Year2010             0.1531     0.0551    2.78   0.0055 **
## Year2011             0.1045     0.0534    1.96   0.0505 .
## Year2012             0.0971     0.0559    1.74   0.0826 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.365
## Multiple R-squared:  0.517,  Adjusted R-squared:  0.513
## Convergence in 38 IRWLS iterations
##
## Robustness weights:
##  64 observations
c(131,148,243,310,386,390,391,462,561,590,597,637,729,797,811,820,867,1021,10
23,1026,1093,1096,1109,1142,1143,1165,1207,1208,1246,1400,1401,1412,1424,1464
,1537,1591,1784,1799,1808,1829,1843,1856,1859,1865,1873,1890,1912,1974,1999,2
031,2072,2076,2115,2150,2276,2435,2485,2567,2581,2596,2628,2704,2734,2749)
##   are outliers with |weight| <= 2e-05 ( < 3.5e-05);
##  44 weights are ~= 1. The remaining 2715 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0002  0.7940  0.9500  0.8340  0.9730  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          3.54e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                    psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##                      GVIF Df GVIF^(1/(2*Df))
```

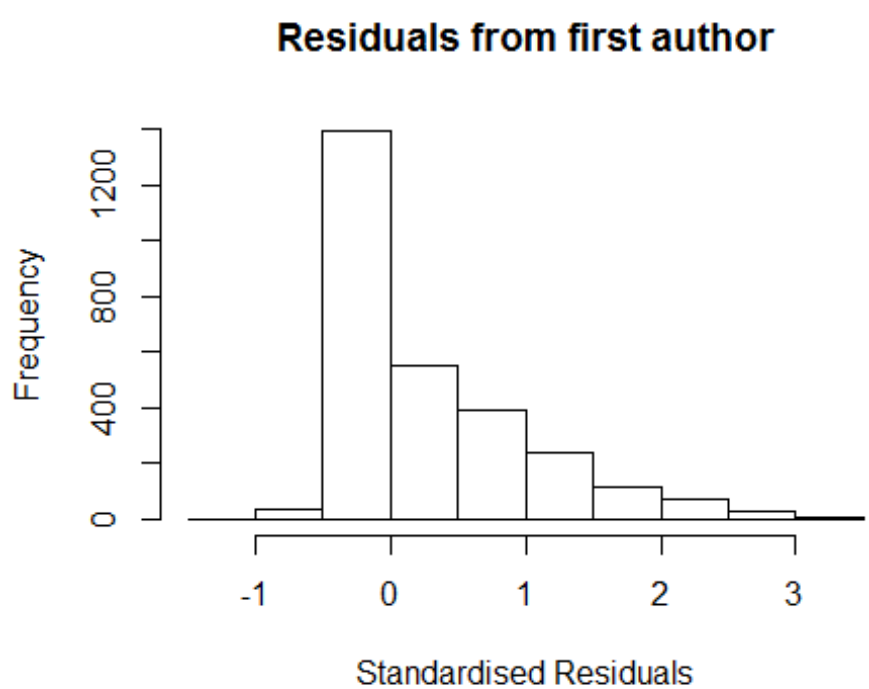
```
## FirstAuthorFemale 11.513  1           3.393
## LastAuthorFemale  11.359  1           3.370
## Year               1.952 16           1.021
```




```
## [1] "List of  27 outliers with residuals above  2.5"
##         ScopusId  NLCS Year OneField Fields residuals
## 108   0242505996 3.218 1996     2902      5     2.917
## 312   0031181060 2.610 1997     2902      5     2.515
## 335   1542755229 2.650 1997     2902      5     2.774
## 1558  0141687084 2.925 2003     2902      4     2.723
## 1579  0037563869 2.925 2003     2902      4     2.723
## 2450 33845478961 2.649 2006     2902      4     2.605
## 2502 33846683368 3.015 2006     2719      5     2.971
## 2504 33947510199 3.556 2006     2719      5     3.183
## 2655 34548496059 3.024 2007     2719      5     2.548
## 2710 33846119150 2.977 2007     2902      4     2.611
## 2915 51549115042 2.714 2008     2719      5     2.638
## 2990 44249122833 3.115 2008     2719      5     2.711
## 3041 40249115349 3.558 2008     2719      5     2.935
## 3183 73849118275 3.246 2009     2719      5     3.144
## 3269 67449130309 3.706 2009     2719      5     3.604
## 3284 70449724751 3.106 2009     2719      5     2.675
## 3287 70450003685 2.760 2009     2719      5     2.658
## 3328 67649460937 3.078 2009     2719      5     2.647
## 3331 67649482684 3.748 2009     2719      5     3.317
## 3363 70449640403 2.647 2009     2719      5     2.545
## 3376 66049128443 3.049 2009     2719      5     2.947
```

```
## 3378 66049138872 2.930 2009     2719      5     2.828
## 3385 67650084651 3.058 2009     2902      4     2.737
## 3619 77956631044 3.365 2010     2719      5     2.834
## 3625 77956713424 2.985 2010     2719      5     2.564
## 3714 77953517631 2.763 2010     2719      5     2.561
## 4107 79952469323 3.228 2011     2719      5     2.700
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.0855 -0.3211 -0.0935  0.6157  3.6040
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)           0.3014     0.0658    4.58  4.9e-06 ***
## FirstAuthorFemale1    0.2190     0.1117    1.96  0.04996 *
## LastAuthorFemale1    -0.3288     0.1058   -3.11  0.00191 **
## Year1997             -0.0969     0.0568   -1.71  0.08820 .
## Year1998             -0.0959     0.0549   -1.75  0.08071 .
## Year1999             -0.0330     0.0676   -0.49  0.62599
## Year2000              0.4519     0.3374    1.34  0.18064
## Year2001              0.8938     0.2456    3.64  0.00028 ***
## Year2002              0.0996     0.0757    1.32  0.18857
## Year2003              0.0104     0.0592    0.18  0.85989
## Year2004              0.1110     0.0615    1.81  0.07118 .
## Year2005              0.2751     0.0695    3.96  7.7e-05 ***
## Year2006              0.0713     0.0727    0.98  0.32701
## Year2007              0.1748     0.0808    2.16  0.03052 *
## Year2008              0.1029     0.0586    1.76  0.07896 .
## Year2009              0.1294     0.0565    2.29  0.02212 *
## Year2010              0.2295     0.0660    3.48  0.00051 ***
## Year2011              0.2262     0.0679    3.33  0.00087 ***
## Year2012              0.3336     0.0853    3.91  9.5e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.419
## Multiple R-squared:  0.0883, Adjusted R-squared:  0.0825
## Convergence in 40 IRWLS iterations
##
## Robustness weights:
##  94 observations
c(79,80,131,148,310,385,386,456,549,556,561,565,574,590,792,820,842,937,1093,
1096,1109,1112,1140,1141,1143,1144,1165,1207,1208,1246,1249,1250,1258,1262,12
68,1311,1400,1401,1402,1405,1424,1454,1480,1506,1514,1541,1542,1587,1591,1592
,1716,1771,1784,1795,1796,1799,1808,1816,1826,1829,1843,1852,1857,1859,1861,1
```

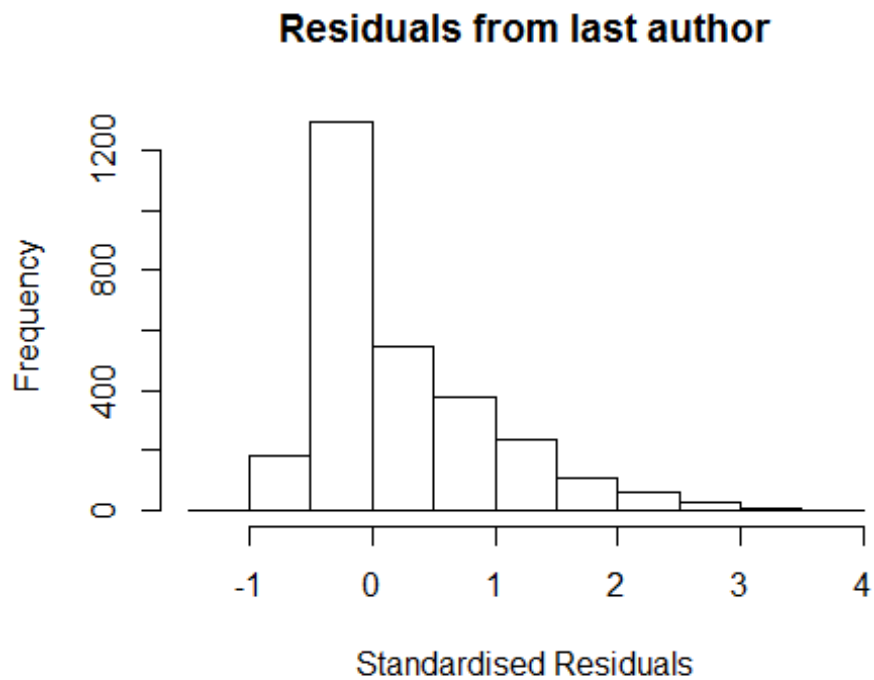
```
865,1873,1890,1912,2042,2062,2072,2076,2112,2117,2147,2149,2259,2264,2317,233
4,2335,2436,2447,2462,2472,2477,2589,2702,2704,2706,2735,2740,2749)
##   are outliers with |weight| <= 3.3e-07 ( < 3.5e-05);
##  78 weights are ~= 1. The remaining 2651 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0004  0.8130  0.9170  0.8280  0.9590  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          3.54e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.997  1           1.413
## Year              1.997 16           1.022
```



```
## [1] "List of  27 outliers with residuals above  2.5"
##         ScopusId  NLCS Year OneField Fields residuals
## 108   0242505996 3.218 1996     2902      5     2.917
## 312   0031181060 2.610 1997     2902      5     2.515
## 335   1542755229 2.650 1997     2902      5     2.774
## 1558  0141687084 2.925 2003     2902      4     2.723
## 1579  0037563869 2.925 2003     2902      4     2.723
## 2450 33845478961 2.649 2006     2902      4     2.605
## 2502 33846683368 3.015 2006     2719      5     2.971
## 2504 33947510199 3.556 2006     2719      5     3.183
## 2655 34548496059 3.024 2007     2719      5     2.548
## 2710 33846119150 2.977 2007     2902      4     2.611
## 2915 51549115042 2.714 2008     2719      5     2.638
## 2990 44249122833 3.115 2008     2719      5     2.711
## 3041 40249115349 3.558 2008     2719      5     2.935
## 3183 73849118275 3.246 2009     2719      5     3.144
## 3269 67449130309 3.706 2009     2719      5     3.604
## 3284 70449724751 3.106 2009     2719      5     2.675
## 3287 70450003685 2.760 2009     2719      5     2.658
## 3328 67649460937 3.078 2009     2719      5     2.647
## 3331 67649482684 3.748 2009     2719      5     3.317
## 3363 70449640403 2.647 2009     2719      5     2.545
## 3376 66049128443 3.049 2009     2719      5     2.947
## 3378 66049138872 2.930 2009     2719      5     2.828
## 3385 67650084651 3.058 2009     2902      4     2.737
## 3619 77956631044 3.365 2010     2719      5     2.834
## 3625 77956713424 2.985 2010     2719      5     2.564
## 3714 77953517631 2.763 2010     2719      5     2.561
## 4107 79952469323 3.228 2011     2719      5     2.700
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.3120 -0.3078 -0.0488  0.6723  3.3703
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)          0.2509     0.0737    3.40  0.00067 ***
## FirstAuthorFemale1  -0.0574     0.0434   -1.32  0.18632
## Year1997            -0.1017     0.0533   -1.91  0.05619 .
## Year1998            -0.1049     0.0516   -2.03  0.04210 *
## Year1999            -0.0225     0.0650   -0.35  0.72873
## Year2000             0.5039     0.3838    1.31  0.18931
## Year2001             1.1185     0.2267    4.93  8.6e-07 ***
## Year2002             0.0699     0.0816    0.86  0.39202
## Year2003             0.0175     0.0572    0.31  0.75941
```

```
## Year2004             0.1039     0.0620    1.68  0.09394 .
## Year2005             0.2809     0.0695    4.04  5.4e-05 ***
## Year2006             0.0569     0.0747    0.76  0.44649
## Year2007             0.1458     0.0870    1.68  0.09401 .
## Year2008             0.0875     0.0586    1.49  0.13544
## Year2009             0.1269     0.0568    2.23  0.02554 *
## Year2010             0.2117     0.0698    3.03  0.00243 **
## Year2011             0.1903     0.0662    2.88  0.00406 **
## Year2012             0.3000     0.1030    2.91  0.00362 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.387
## Multiple R-squared:  0.0658, Adjusted R-squared:  0.0601
## Convergence in 24 IRWLS iterations
##
## Robustness weights:
##  130 observations
c(79,80,95,131,148,177,310,385,386,390,456,466,549,556,561,565,574,590,597,63
7,729,792,797,811,820,833,937,1023,1093,1096,1109,1112,1124,1140,1141,1143,11
44,1165,1207,1208,1239,1246,1249,1250,1258,1262,1268,1295,1311,1334,1400,1401
,1402,1405,1412,1424,1454,1480,1506,1509,1541,1542,1555,1582,1586,1587,1589,1
591,1592,1716,1769,1771,1784,1795,1796,1799,1808,1816,1821,1824,1826,1829,184
3,1852,1857,1859,1861,1865,1873,1890,1912,2042,2062,2071,2072,2076,2111,2112,
2117,2147,2148,2149,2164,2165,2259,2264,2317,2334,2335,2398,2401,2447,2462,24
64,2472,2476,2477,2485,2511,2542,2589,2607,2628,2702,2704,2706,2735,2737,2740
,2749)
##   are outliers with |weight| <= 2.2e-05 ( < 3.5e-05);
##  67 weights are ~= 1. The remaining 2626 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0002  0.8200  0.9250  0.8140  0.9580  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          3.54e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
```

```
## LastAuthorFemale 2.684  1           1.638
## Year             2.684 16           1.031
```

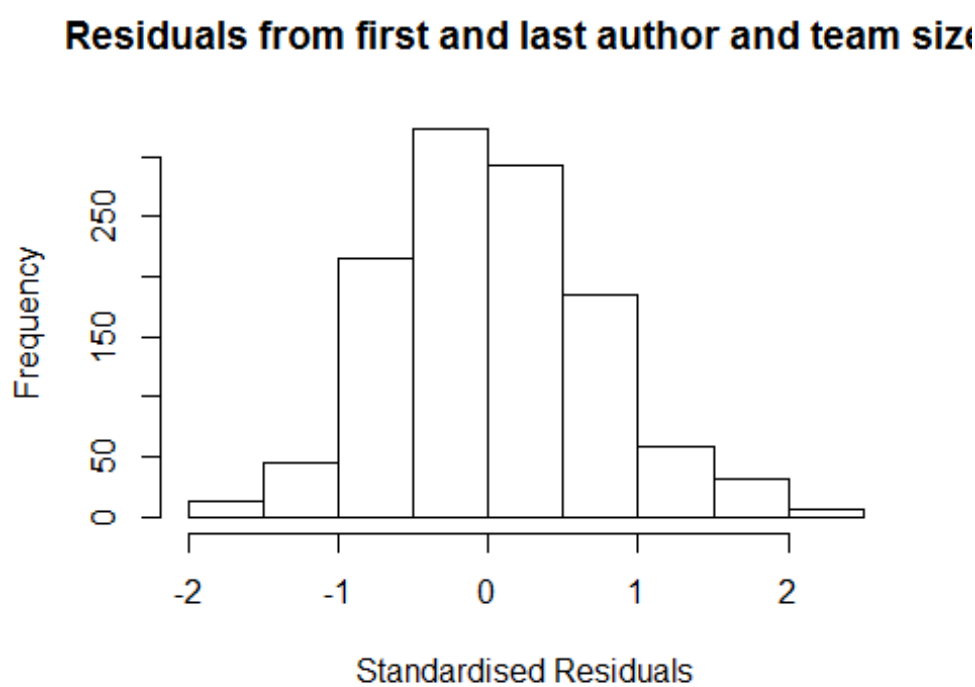



```
## [1] "List of  27 outliers with residuals above  2.5"
##         ScopusId  NLCS Year OneField Fields residuals
## 108   0242505996 3.218 1996     2902      5     2.917
## 312   0031181060 2.610 1997     2902      5     2.515
## 335   1542755229 2.650 1997     2902      5     2.774
## 1558  0141687084 2.925 2003     2902      4     2.723
## 1579  0037563869 2.925 2003     2902      4     2.723
## 2450 33845478961 2.649 2006     2902      4     2.605
## 2502 33846683368 3.015 2006     2719      5     2.971
## 2504 33947510199 3.556 2006     2719      5     3.183
## 2655 34548496059 3.024 2007     2719      5     2.548
## 2710 33846119150 2.977 2007     2902      4     2.611
## 2915 51549115042 2.714 2008     2719      5     2.638
## 2990 44249122833 3.115 2008     2719      5     2.711
## 3041 40249115349 3.558 2008     2719      5     2.935
## 3183 73849118275 3.246 2009     2719      5     3.144
## 3269 67449130309 3.706 2009     2719      5     3.604
## 3284 70449724751 3.106 2009     2719      5     2.675
## 3287 70450003685 2.760 2009     2719      5     2.658
## 3328 67649460937 3.078 2009     2719      5     2.647
## 3331 67649482684 3.748 2009     2719      5     3.317
## 3363 70449640403 2.647 2009     2719      5     2.545
## 3376 66049128443 3.049 2009     2719      5     2.947
## 3378 66049138872 2.930 2009     2719      5     2.828
```

```
## 3385 67650084651 3.058 2009     2902      4     2.737
## 3619 77956631044 3.365 2010     2719      5     2.834
## 3625 77956713424 2.985 2010     2719      5     2.564
## 3714 77953517631 2.763 2010     2719      5     2.561
## 4107 79952469323 3.228 2011     2719      5     2.700
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.2051 -0.3074 -0.0785  0.6516  3.3986
##
## Coefficients:
##                   Estimate Std. Error t value Pr(>|t|)
## (Intercept)        0.29986    0.08059    3.72   0.0002 ***
## LastAuthorFemale1 -0.11983    0.05584   -2.15   0.0320 *
## Year1997          -0.09542    0.05456   -1.75   0.0804 .
## Year1998          -0.10152    0.05300   -1.92   0.0555 .
## Year1999          -0.02942    0.06624   -0.44   0.6570
## Year2000           0.47035    0.34603    1.36   0.1742
## Year2001           1.02505    0.23556    4.35  1.4e-05 ***
## Year2002           0.08445    0.08082    1.04   0.2961
## Year2003           0.00981    0.05865    0.17   0.8672
## Year2004           0.10838    0.06261    1.73   0.0836 .
## Year2005           0.28338    0.06974    4.06  5.0e-05 ***
## Year2006           0.05782    0.07476    0.77   0.4393
## Year2007           0.15797    0.08696    1.82   0.0694 .
## Year2008           0.08939    0.05880    1.52   0.1286
## Year2009           0.12734    0.05696    2.24   0.0255 *
## Year2010           0.21826    0.07020    3.11   0.0019 **
## Year2011           0.20412    0.06986    2.92   0.0035 **
## Year2012           0.31781    0.10192    3.12   0.0018 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.393
## Multiple R-squared:  0.0731, Adjusted R-squared:  0.0675
## Convergence in 25 IRWLS iterations
##
## Robustness weights:
##  120 observations
c(79,80,95,131,148,177,310,385,386,390,456,466,549,556,561,565,574,590,597,63
7,729,792,811,820,833,937,1023,1093,1096,1109,1112,1140,1141,1143,1144,1165,1
207,1208,1246,1249,1250,1258,1262,1268,1295,1311,1334,1400,1401,1402,1405,141
2,1424,1454,1480,1506,1541,1542,1555,1582,1586,1587,1589,1591,1592,1716,1769,
1771,1784,1795,1796,1799,1808,1816,1821,1826,1829,1843,1852,1857,1859,1861,18
65,1873,1890,1912,2042,2062,2071,2072,2076,2111,2112,2117,2147,2148,2149,2165
```

```
,2259,2264,2317,2334,2335,2398,2447,2462,2464,2472,2477,2485,2511,2589,2607,2
628,2702,2704,2706,2735,2740,2749)
##   are outliers with |weight| <= 1.5e-05 ( < 3.5e-05);
##  76 weights are ~= 1. The remaining 2627 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0007  0.7980  0.9150  0.8170  0.9590  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          3.54e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Sample size for the above analysis:  2823"
## [1] ""
## [1] ""
## [1] "####################################"
## [1] "Analysis of AJSC 2904"
## [1] "####################################"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##   52   59   83   61   68   82   89   86   81   79   61   95  152  194  201
## 2011 2012
##  135  151
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##   23   18   46   30    3    1   85   80   70   67   60   83  131  170  166
## 2011 2012
##  114  126
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##   23   18   42   30    3    1   81   72   67   64   56   74  125  151  150
## 2011 2012
##  103  113
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
## [1] "Female first author team size 2018 geometric mean: 4.02697641304799"
## [1] "Male first author team size 2018 geometric mean: 4.05019479433859"
##
```

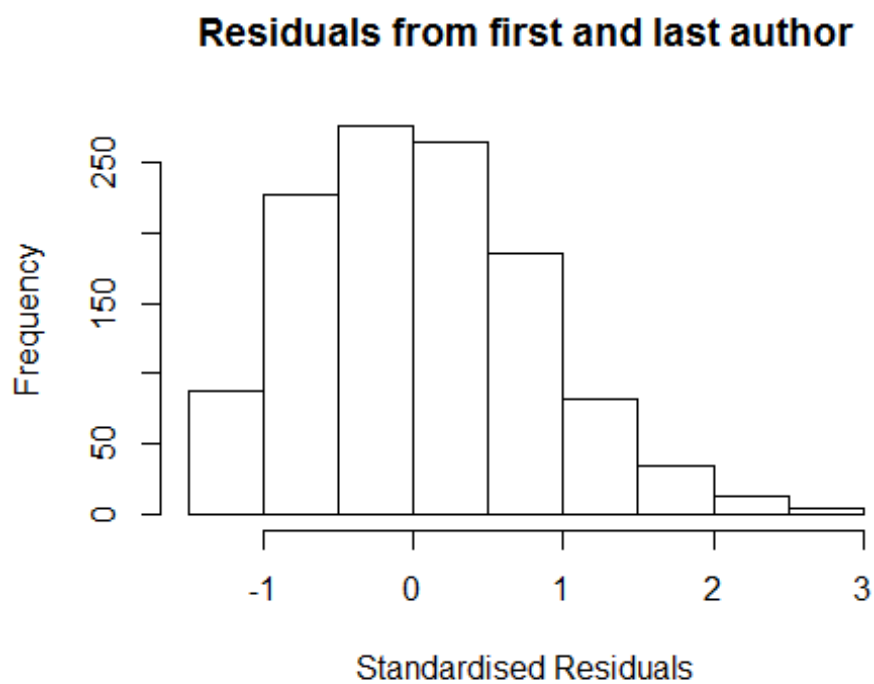
```
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 2200, p-value = 1
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 3.77221284518514"
## [1] "Male last author team size 2018 geometric mean: 4.29040044316832"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 2000, p-value = 0.2
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##                     GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 11.388  1           3.375
## LastAuthorFemale  12.515  1           3.538
## UniqueAuthors      5.018  4           1.223
## Year              42.488 16           1.124
```




```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
```

```
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
UniqueAuthors +
##     Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##     k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.8401 -0.4845 -0.0236  0.4649  2.4649
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)          0.0370     0.0433    0.85   0.3936
## FirstAuthorFemale1  -0.0104     0.0487   -0.21   0.8301
## LastAuthorFemale1   -0.0367     0.0502   -0.73   0.4646
## UniqueAuthors2       0.1838     0.0585    3.14   0.0017 **
## UniqueAuthors3       0.4096     0.0706    5.80  8.6e-09 ***
## UniqueAuthors4       0.6191     0.0817    7.58  7.3e-14 ***
## UniqueAuthors5       1.0824     0.0710   15.25  < 2e-16 ***
## Year1997            -0.0235     0.0357   -0.66   0.5108
## Year1998             0.0928     0.0564    1.65   0.1002
## Year1999             0.2263     0.0874    2.59   0.0098 **
## Year2000             0.1119     0.1319    0.85   0.3963
## Year2001            -0.0370     0.0433   -0.85   0.3936
## Year2002             0.7576     0.0987    7.67  3.6e-14 ***
## Year2003             0.7599     0.1152    6.60  6.4e-11 ***
## Year2004             0.8530     0.1066    8.00  3.0e-15 ***
## Year2005             0.8199     0.0907    9.04  < 2e-16 ***
## Year2006             0.5766     0.1334    4.32  1.7e-05 ***
## Year2007             0.6551     0.1072    6.11  1.4e-09 ***
## Year2008             0.5323     0.0812    6.55  8.5e-11 ***
## Year2009             0.4947     0.0735    6.73  2.7e-11 ***
## Year2010             0.4535     0.0757    5.99  2.8e-09 ***
## Year2011             0.5701     0.0785    7.26  6.9e-13 ***
## Year2012             0.6390     0.0822    7.78  1.6e-14 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.654
## Multiple R-squared:  0.343,  Adjusted R-squared:  0.33
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
##  117 weights are ~= 1. The remaining 1056 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.124   0.862   0.941   0.895   0.982   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
```

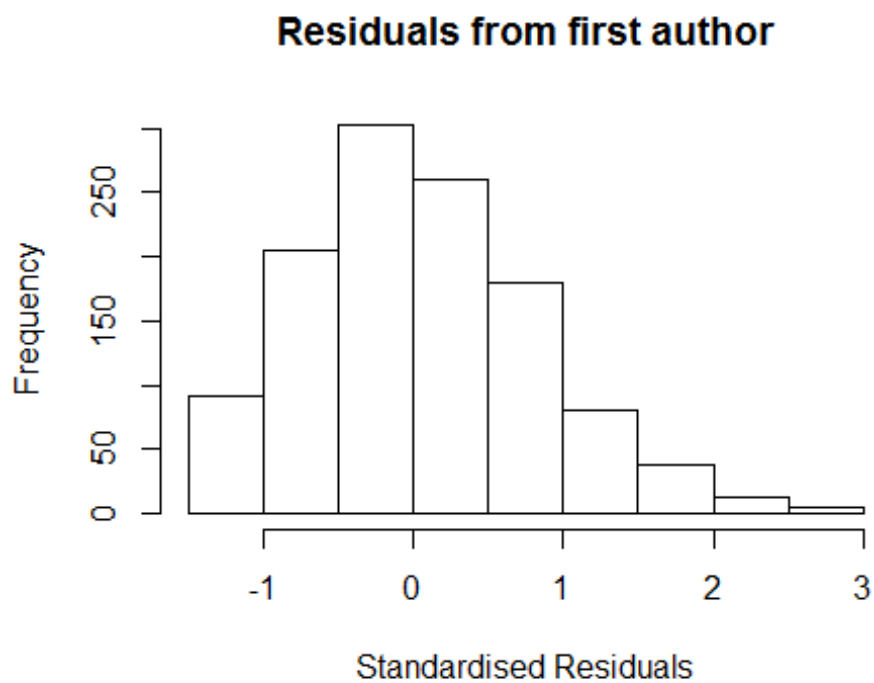
```
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          8.53e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 40.05  1           6.328
## LastAuthorFemale  46.26  1           6.802
## Year              51.98 16           1.131
```




```
## [1] "List of  4 outliers with residuals above  2.5"
##         ScopusId  NLCS Year OneField Fields residuals
## 817  33947510199 3.556 2006     2719      5     2.577
## 1118 40249115349 3.558 2008     2719      5     2.530
## 1165 67449130309 3.706 2009     2719      5     2.947
## 1215 67649482684 3.748 2009     2719      5     2.814
##
```

```
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.39618 -0.53175 -0.00195  0.54738  2.94749
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         0.12952    0.05138    2.52   0.0118 *
## FirstAuthorFemale1  0.03005    0.05428    0.55   0.5800
## LastAuthorFemale1  -0.17544    0.05550   -3.16   0.0016 **
## Year1997            0.00948    0.03661    0.26   0.7957
## Year1998            0.11667    0.05929    1.97   0.0493 *
## Year1999            0.28472    0.09132    3.12   0.0019 **
## Year2000            0.01976    0.13424    0.15   0.8830
## Year2001           -0.12952    0.05138   -2.52   0.0118 *
## Year2002            1.00127    0.09393   10.66   <2e-16 ***
## Year2003            1.03725    0.10941    9.48   <2e-16 ***
## Year2004            1.02789    0.10153   10.12   <2e-16 ***
## Year2005            1.07105    0.09938   10.78   <2e-16 ***
## Year2006            0.84911    0.15214    5.58    3e-08 ***
## Year2007            1.14522    0.10398   11.01   <2e-16 ***
## Year2008            0.86881    0.08513   10.21   <2e-16 ***
## Year2009            0.80443    0.08305    9.69   <2e-16 ***
## Year2010            0.84765    0.08504    9.97   <2e-16 ***
## Year2011            1.04922    0.09675   10.84   <2e-16 ***
## Year2012            1.26666    0.08631   14.68   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.772
## Multiple R-squared:  0.148,  Adjusted R-squared:  0.134
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
##  116 weights are ~= 1. The remaining 1057 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.113   0.860   0.945   0.904   0.984   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          8.53e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
```

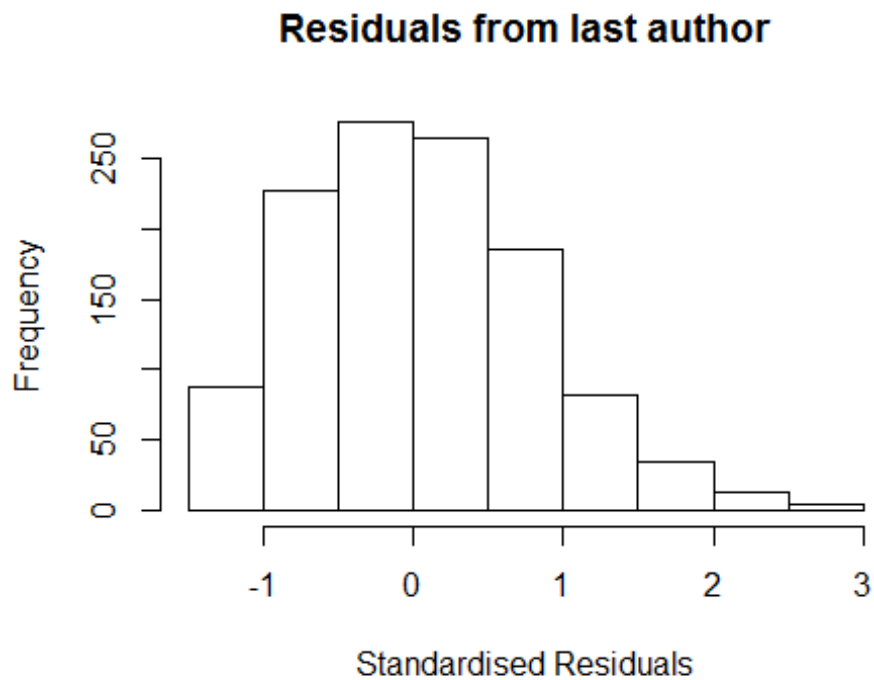
```
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                    cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 176.6  1          13.291
## Year              176.6 16           1.175
```




```
## [1] "List of  4 outliers with residuals above  2.5"
##         ScopusId  NLCS Year OneField Fields residuals
## 817  33947510199 3.556 2006     2719      5     2.577
## 1118 40249115349 3.558 2008     2719      5     2.530
## 1165 67449130309 3.706 2009     2719      5     2.947
## 1215 67649482684 3.748 2009     2719      5     2.814
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.3744 -0.5030 -0.0151  0.5530  2.8303
```

```
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)          0.0594     0.0442    1.34   0.1793
## FirstAuthorFemale1  -0.0442     0.0485   -0.91   0.3616
## Year1997            -0.0200     0.0310   -0.65   0.5175
## Year1998             0.1232     0.0571    2.16   0.0311 *
## Year1999             0.2993     0.0932    3.21   0.0014 **
## Year2000             0.0899     0.1316    0.68   0.4947
## Year2001            -0.0594     0.0442   -1.34   0.1793
## Year2002             1.0235     0.0932   10.99   <2e-16 ***
## Year2003             1.0451     0.1093    9.56   <2e-16 ***
## Year2004             1.0592     0.0972   10.90   <2e-16 ***
## Year2005             1.0964     0.0960   11.42   <2e-16 ***
## Year2006             0.8897     0.1500    5.93    4e-09 ***
## Year2007             1.1820     0.1039   11.38   <2e-16 ***
## Year2008             0.9010     0.0847   10.64   <2e-16 ***
## Year2009             0.8584     0.0798   10.76   <2e-16 ***
## Year2010             0.8749     0.0846   10.35   <2e-16 ***
## Year2011             1.0906     0.0946   11.53   <2e-16 ***
## Year2012             1.3150     0.0846   15.54   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.772
## Multiple R-squared:  0.14,   Adjusted R-squared:  0.127
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
##  105 weights are ~= 1. The remaining 1068 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.151   0.867   0.946   0.904   0.985   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          8.53e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
```

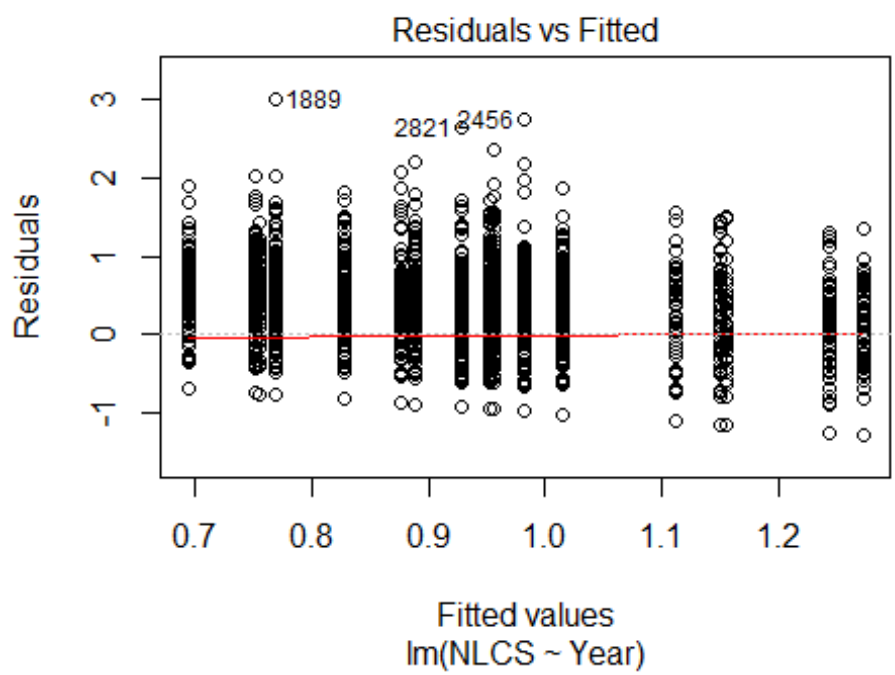
```
## LastAuthorFemale 31.74  1           5.634
## Year             31.74 16           1.114
```



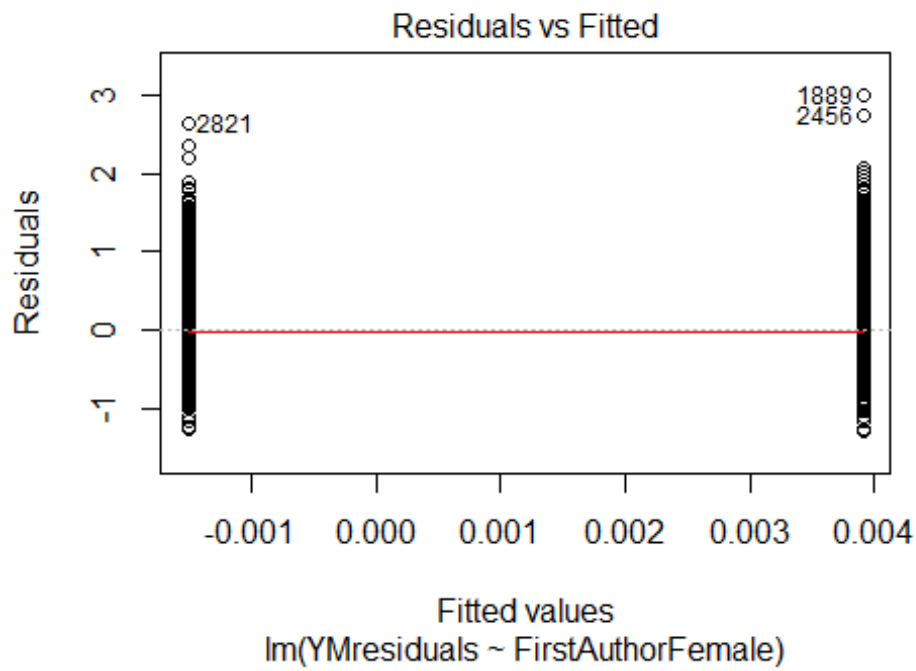

```
## [1] "List of  4 outliers with residuals above  2.5"
##         ScopusId  NLCS Year OneField Fields residuals
## 817  33947510199 3.556 2006     2719      5     2.577
## 1118 40249115349 3.558 2008     2719      5     2.530
## 1165 67449130309 3.706 2009     2719      5     2.947
## 1215 67649482684 3.748 2009     2719      5     2.814
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##       Min        1Q    Median        3Q       Max
## -1.406109 -0.530746 -0.000936  0.547226  2.922064
##
## Coefficients:
##                   Estimate Std. Error t value Pr(>|t|)
## (Intercept)        0.14115    0.04807    2.94   0.0034 **
## LastAuthorFemale1 -0.16214    0.04926   -3.29   0.0010 **
## Year1997           0.01202    0.03729    0.32   0.7474
## Year1998           0.11414    0.05983    1.91   0.0567 .
## Year1999           0.28156    0.09091    3.10   0.0020 **
## Year2000           0.00813    0.13302    0.06   0.9513
```

```
## Year2001          -0.14115    0.04807   -2.94   0.0034 **
## Year2002           1.00033    0.09430   10.61  < 2e-16 ***
## Year2003           1.03297    0.10973    9.41  < 2e-16 ***
## Year2004           1.02719    0.10184   10.09  < 2e-16 ***
## Year2005           1.06926    0.09945   10.75  < 2e-16 ***
## Year2006           0.84748    0.15219    5.57  3.2e-08 ***
## Year2007           1.14074    0.10424   10.94  < 2e-16 ***
## Year2008           0.86477    0.08544   10.12  < 2e-16 ***
## Year2009           0.80493    0.08336    9.66  < 2e-16 ***
## Year2010           0.84602    0.08537    9.91  < 2e-16 ***
## Year2011           1.04650    0.09686   10.80  < 2e-16 ***
## Year2012           1.26496    0.08690   14.56  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.772
## Multiple R-squared:  0.147,  Adjusted R-squared:  0.135
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
##  121 weights are ~= 1. The remaining 1052 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.120   0.859   0.944   0.903   0.983   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          8.53e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Sample size for the above analysis:  1173"
## [1] ""
## [1] ""
## [1] "####################################"
## [1] "Analysis of AJSC 2905"
## [1] "####################################"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  109  117  143  146  188  228  272  211  245  254  227  245  285  297  376
```

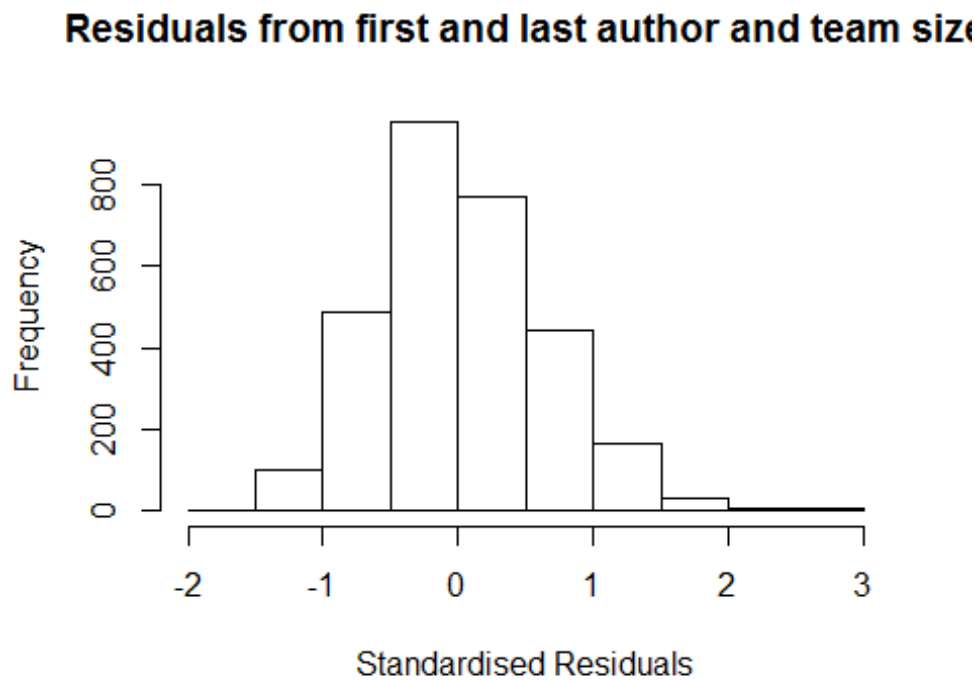
```
## 2011 2012
##  367  333
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##   75   49   70   71   84  131  243  193  219  224  195  208  239  258  322
## 2011 2012
##  313  279
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##   72   43   65   67   82  125  229  182  210  209  181  194  225  245  306
## 2011 2012
##  273  251
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
##  Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 22, df = 16, p-value = 0.2
```




```
##
##  Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 17, df = 1, p-value = 3e-05
```

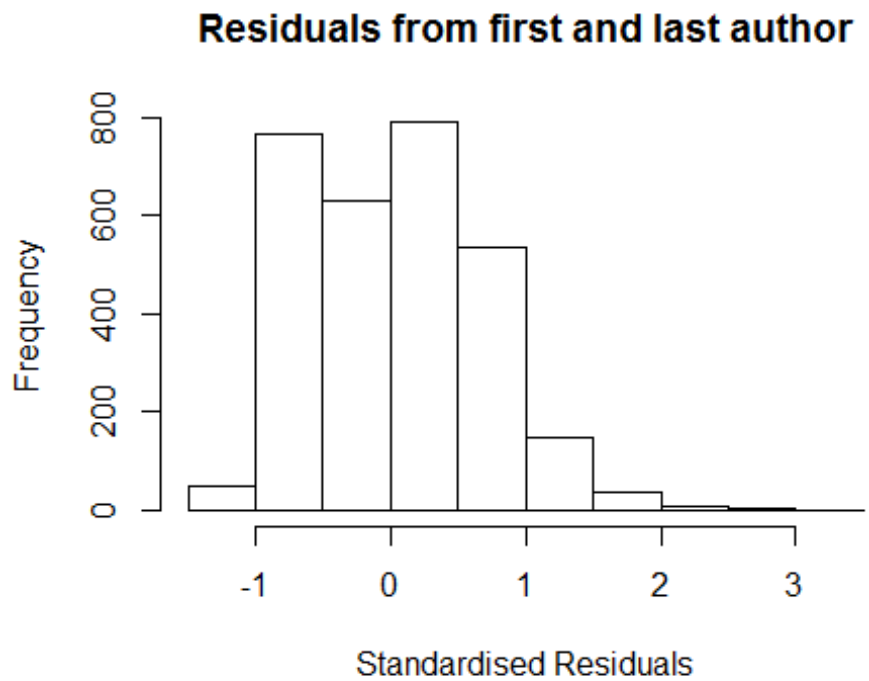



```
## [1] "Female first author team size 2018 geometric mean: 2.75376393007459"
## [1] "Male first author team size 2018 geometric mean: 2.76230410300434"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 4200, p-value = 0.7
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 2.52027700418133"
## [1] "Male last author team size 2018 geometric mean: 3.48737482969889"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 3600, p-value = 0.002
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.334  1           1.155
## LastAuthorFemale  1.431  1           1.196
## UniqueAuthors     1.282  4           1.032
## Year              1.309 16           1.008
```



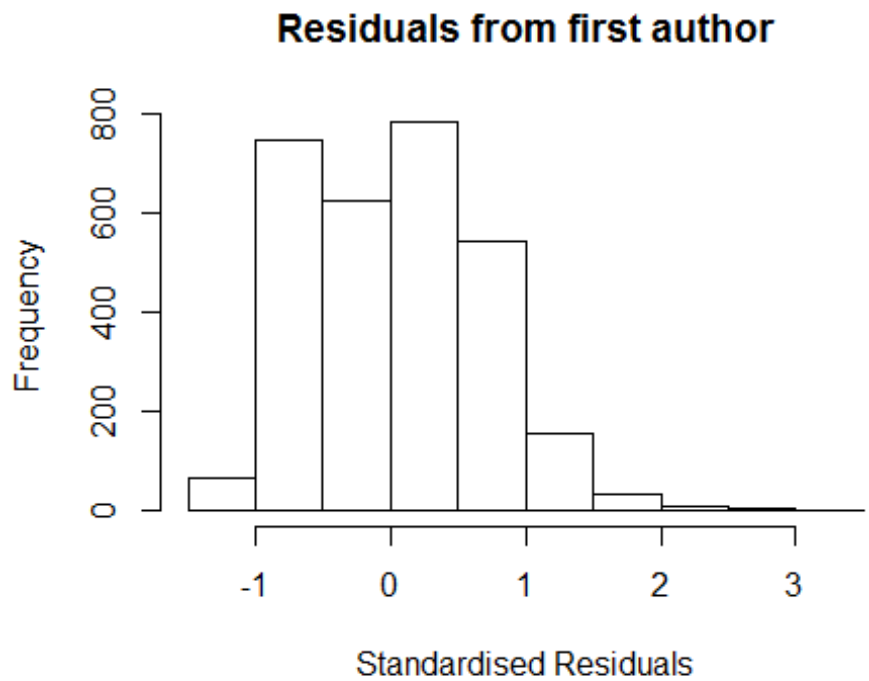

```
## [1] "List of  3 outliers with residuals above  2.5"
##         ScopusId  NLCS Year OneField Fields residuals
## 1889 27744485315 3.768 2005     2905      1     2.676
## 2821 43549086424 3.565 2008     2714      2     2.538
## 3205 78149400477 3.305 2010     2714      2     2.736
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
UniqueAuthors +
##     Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##     k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.5475 -0.4301 -0.0351  0.4303  2.7361
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)          0.8543     0.0904    9.45  < 2e-16 ***
## FirstAuthorFemale1   0.0784     0.0307    2.55  0.01075 *
## LastAuthorFemale1   -0.0352     0.0292   -1.20  0.22906
## UniqueAuthors2       0.4780     0.0325   14.71  < 2e-16 ***
## UniqueAuthors3       0.5987     0.0382   15.67  < 2e-16 ***
## UniqueAuthors4       0.6082     0.0429   14.16  < 2e-16 ***
## UniqueAuthors5       0.7277     0.0377   19.29  < 2e-16 ***
## Year1997            -0.1747     0.1284   -1.36  0.17379
```

```
## Year1998            -0.1126     0.1222   -0.92  0.35696
## Year1999            -0.1139     0.1368   -0.83  0.40515
## Year2000             0.0161     0.1154    0.14  0.88919
## Year2001            -0.4170     0.0993   -4.20  2.7e-05 ***
## Year2002            -0.4853     0.0929   -5.22  1.9e-07 ***
## Year2003            -0.3512     0.0996   -3.52  0.00043 ***
## Year2004            -0.4674     0.0956   -4.89  1.1e-06 ***
## Year2005            -0.4899     0.0946   -5.18  2.4e-07 ***
## Year2006            -0.4036     0.0954   -4.23  2.4e-05 ***
## Year2007            -0.3374     0.0960   -3.51  0.00045 ***
## Year2008            -0.3483     0.0927   -3.76  0.00018 ***
## Year2009            -0.3617     0.0945   -3.83  0.00013 ***
## Year2010            -0.3286     0.0920   -3.57  0.00036 ***
## Year2011            -0.3522     0.0915   -3.85  0.00012 ***
## Year2012            -0.3100     0.0930   -3.33  0.00087 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.55
## Multiple R-squared:  0.241,  Adjusted R-squared:  0.235
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
##  2 observations c(1166,2159) are outliers with |weight| = 0 ( < 3.4e-05);
##  208 weights are ~= 1. The remaining 2749 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0009  0.8700  0.9350  0.8940  0.9810  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          3.38e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.232  1           1.110
## LastAuthorFemale  1.262  1           1.124
## Year              1.074 16           1.002
```



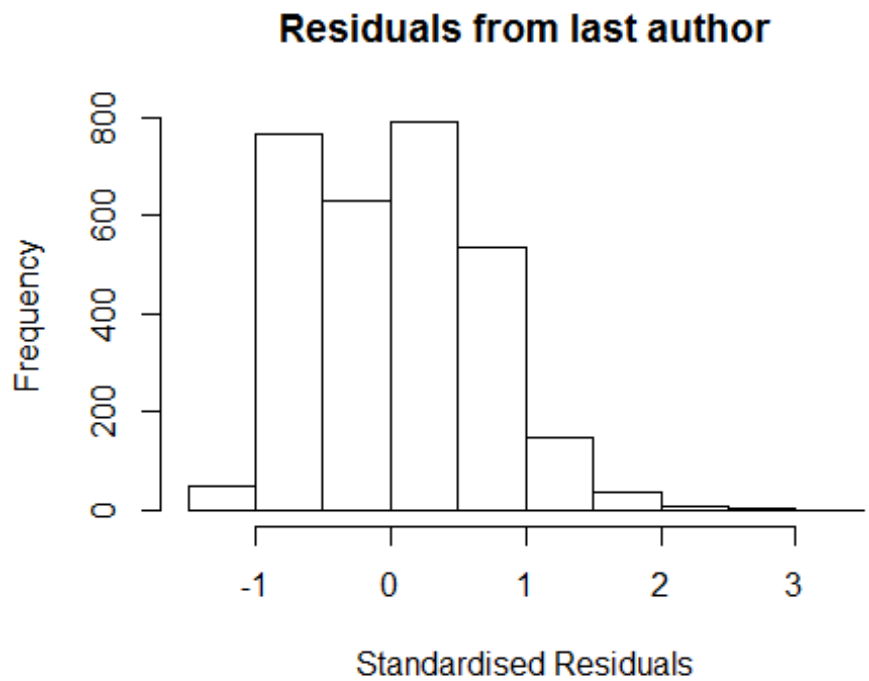

```
## [1] "List of  3 outliers with residuals above  2.5"
##         ScopusId  NLCS Year OneField Fields residuals
## 1889 27744485315 3.768 2005     2905      1     3.042
## 2456 34547138943 3.721 2007     2700      2     2.757
## 2821 43549086424 3.565 2008     2714      2     2.689
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.3817 -0.5677  0.0221  0.4860  3.0418
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.28699    0.07096   18.14  < 2e-16 ***
## FirstAuthorFemale1  0.09784    0.03361    2.91   0.0036 **
## LastAuthorFemale1  -0.14015    0.03098   -4.52  6.3e-06 ***
## Year1997           -0.14670    0.12543   -1.17   0.2423
## Year1998           -0.16263    0.10907   -1.49   0.1360
## Year1999           -0.12603    0.12002   -1.05   0.2938
## Year2000           -0.00311    0.09692   -0.03   0.9744
## Year2001           -0.57073    0.09477   -6.02  1.9e-09 ***
## Year2002           -0.60164    0.08041   -7.48  9.6e-14 ***
## Year2003           -0.44014    0.08708   -5.05  4.6e-07 ***
```

```
## Year2004           -0.55142    0.08454   -6.52  8.1e-11 ***
## Year2005           -0.56077    0.08301   -6.76  1.7e-11 ***
## Year2006           -0.44452    0.08484   -5.24  1.7e-07 ***
## Year2007           -0.32254    0.08649   -3.73   0.0002 ***
## Year2008           -0.36895    0.07951   -4.64  3.6e-06 ***
## Year2009           -0.42351    0.08206   -5.16  2.6e-07 ***
## Year2010           -0.31574    0.07805   -4.05  5.4e-05 ***
## Year2011           -0.33988    0.07746   -4.39  1.2e-05 ***
## Year2012           -0.24786    0.07832   -3.16   0.0016 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.706
## Multiple R-squared:  0.0557, Adjusted R-squared:  0.0499
## Convergence in 10 IRWLS iterations
##
## Robustness weights:
##  234 weights are ~= 1. The remaining 2725 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0241  0.8840  0.9410  0.9190  0.9840  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          3.38e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.026  1           1.013
## Year              1.026 16           1.001
```




```
## [1] "List of  3 outliers with residuals above  2.5"
##         ScopusId  NLCS Year OneField Fields residuals
## 1889 27744485315 3.768 2005     2905      1     3.042
## 2456 34547138943 3.721 2007     2700      2     2.757
## 2821 43549086424 3.565 2008     2714      2     2.689
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.2718 -0.5644  0.0187  0.5002  3.0985
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)          1.2455     0.0689   18.08  < 2e-16 ***
## FirstAuthorFemale1   0.0263     0.0303    0.87  0.38433
## Year1997            -0.1523     0.1234   -1.23  0.21732
## Year1998            -0.1509     0.1084   -1.39  0.16403
## Year1999            -0.1392     0.1184   -1.18  0.23972
## Year2000            -0.0197     0.0949   -0.21  0.83575
## Year2001            -0.5679     0.0949   -5.98  2.5e-09 ***
## Year2002            -0.6102     0.0792   -7.71  1.7e-14 ***
## Year2003            -0.4552     0.0862   -5.28  1.4e-07 ***
## Year2004            -0.5579     0.0841   -6.63  3.8e-11 ***
```

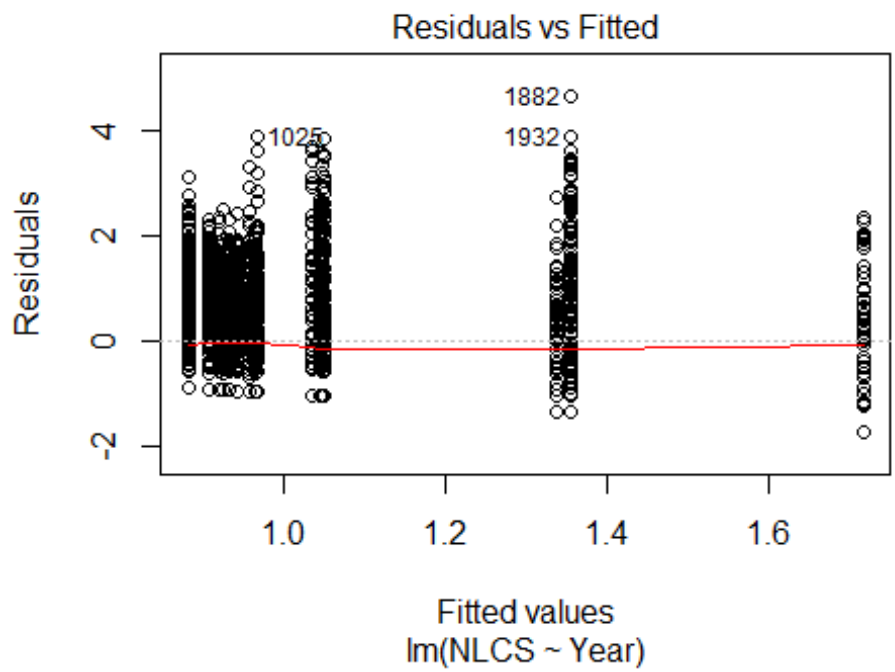
```
## Year2005            -0.5760     0.0821   -7.01  2.9e-12 ***
## Year2006            -0.4425     0.0844   -5.24  1.7e-07 ***
## Year2007            -0.3268     0.0853   -3.83  0.00013 ***
## Year2008            -0.3731     0.0786   -4.75  2.2e-06 ***
## Year2009            -0.4165     0.0816   -5.11  3.5e-07 ***
## Year2010            -0.3194     0.0770   -4.15  3.4e-05 ***
## Year2011            -0.3342     0.0762   -4.38  1.2e-05 ***
## Year2012            -0.2443     0.0774   -3.16  0.00161 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.708
## Multiple R-squared:  0.0496, Adjusted R-squared:  0.0441
## Convergence in 10 IRWLS iterations
##
## Robustness weights:
##  230 weights are ~= 1. The remaining 2729 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.016   0.880   0.941   0.919   0.983   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          3.38e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##                  GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.05  1           1.025
## Year             1.05 16           1.002
```



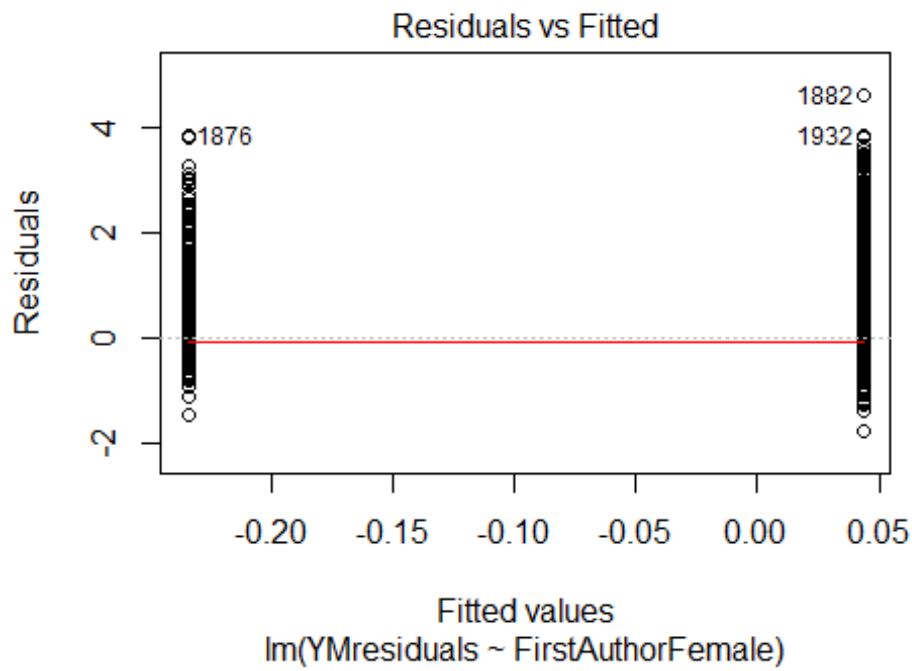

```
## [1] "List of  3 outliers with residuals above  2.5"
##         ScopusId  NLCS Year OneField Fields residuals
## 1889 27744485315 3.768 2005     2905      1     3.042
## 2456 34547138943 3.721 2007     2700      2     2.757
## 2821 43549086424 3.565 2008     2714      2     2.689
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.3240 -0.5623  0.0319  0.4954  3.0023
##
## Coefficients:
##                   Estimate Std. Error t value Pr(>|t|)
## (Intercept)        1.32404    0.06940   19.08  < 2e-16 ***
## LastAuthorFemale1 -0.09442    0.02885   -3.27   0.0011 **
## Year1997          -0.14841    0.12496   -1.19   0.2351
## Year1998          -0.16538    0.11132   -1.49   0.1375
## Year1999          -0.12265    0.11962   -1.03   0.3053
## Year2000          -0.00276    0.09735   -0.03   0.9774
## Year2001          -0.56866    0.09523   -5.97  2.6e-09 ***
## Year2002          -0.59966    0.08029   -7.47  1.1e-13 ***
## Year2003          -0.43739    0.08671   -5.04  4.8e-07 ***
## Year2004          -0.54143    0.08431   -6.42  1.6e-10 ***
```

```
## Year2005          -0.55833    0.08287   -6.74  1.9e-11 ***
## Year2006          -0.43253    0.08482   -5.10  3.6e-07 ***
## Year2007          -0.32096    0.08617   -3.72   0.0002 ***
## Year2008          -0.36757    0.07943   -4.63  3.9e-06 ***
## Year2009          -0.41818    0.08203   -5.10  3.6e-07 ***
## Year2010          -0.31648    0.07797   -4.06  5.1e-05 ***
## Year2011          -0.33814    0.07721   -4.38  1.2e-05 ***
## Year2012          -0.24071    0.07813   -3.08   0.0021 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.707
## Multiple R-squared:  0.0529, Adjusted R-squared:  0.0474
## Convergence in 10 IRWLS iterations
##
## Robustness weights:
##  232 weights are ~= 1. The remaining 2727 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0319  0.8840  0.9420  0.9190  0.9830  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          3.38e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                  psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Sample size for the above analysis:  2959"
## [1] ""
## [1] ""
## [1] "##################################"
## [1] "Analysis of AJSC 2906"
## [1] "##################################"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  218  345  365  350  345  218  511  454  369  326  303  297  410  438  408
## 2011 2012
##  428  454
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
```

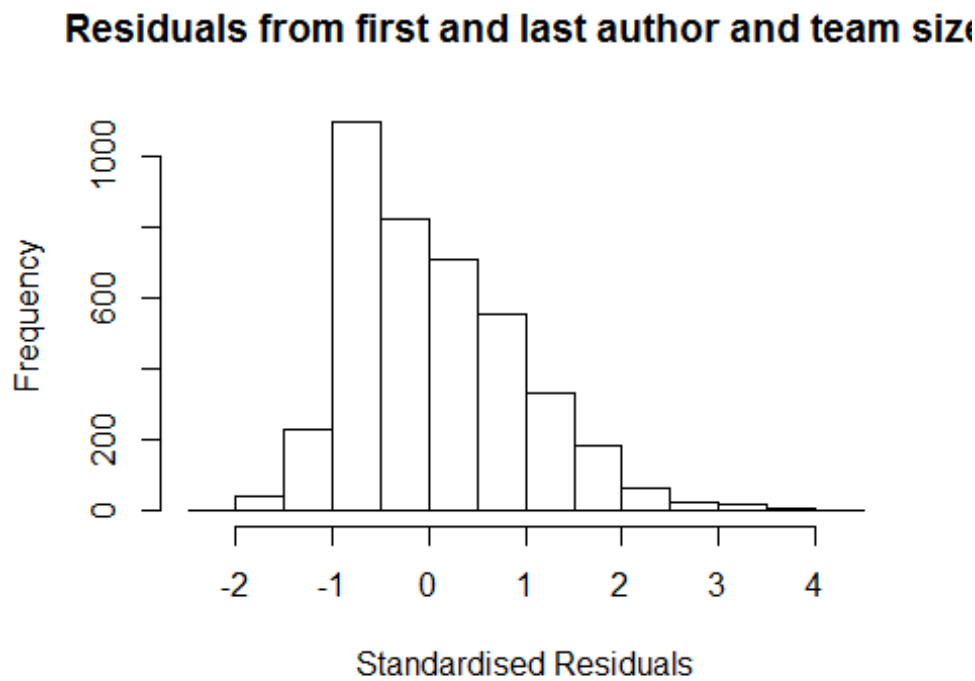
```
##  116  120  144  123   71   91  286  253  296  263  277  265  375  401  363
## 2011 2012
##  393  403
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  112  117  142  121   67   88  275  242  285  245  268  256  362  390  338
## 2011 2012
##  376  382
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
##  Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 300, df = 16, p-value <2e-16
```




```
##
##  Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 3.1, df = 1, p-value = 0.08
```



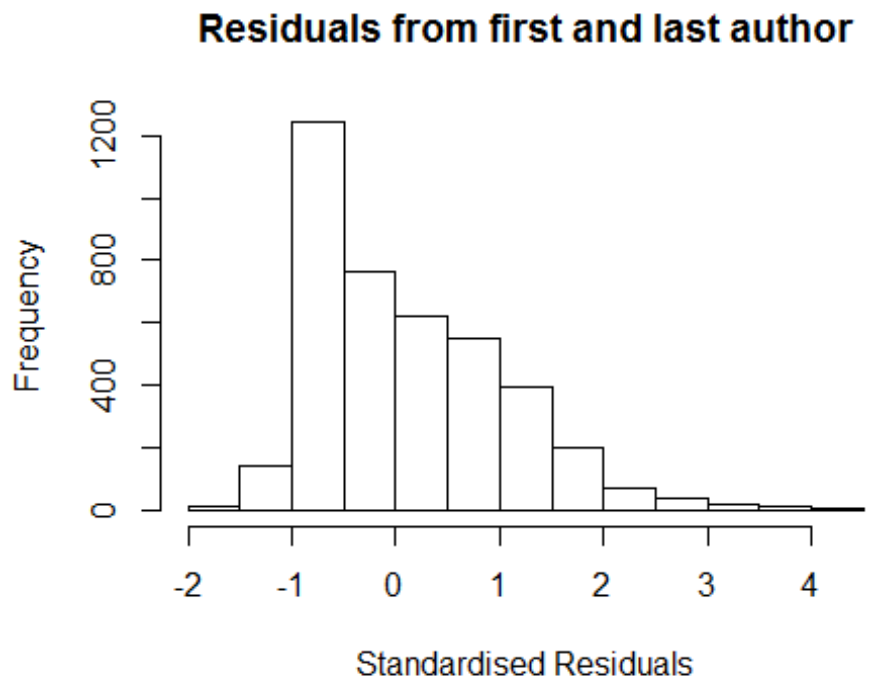

```
## [1] "Female first author team size 2018 geometric mean: 2.08487046641704"
## [1] "Male first author team size 2018 geometric mean: 2.39081082880515"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 8000, p-value = 0.2
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 1.97423704200591"
## [1] "Male last author team size 2018 geometric mean: 2.78662114114098"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 8000, p-value = 2e-04
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##                     GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 3.270  1           1.808
## LastAuthorFemale  3.115  1           1.765
## UniqueAuthors     1.258  4           1.029
## Year              1.275 16           1.008
```




```
## [1] "List of  46 outliers with residuals above  2.5"
##         ScopusId  NLCS Year OneField Fields residuals
## 8     0030280711 2.980 1996     2906      1     2.653
## 21    0030185218 3.233 1996     2906      1     2.643
## 273   0031229882 4.173 1997     2906      1     2.822
## 274   0031229883 4.722 1997     2906      1     2.968
## 315   0031132148 4.483 1997     2906      1     3.466
## 318   0031135082 4.667 1997     2906      1     3.316
## 403   0031180540 3.068 1997     2906      2     2.520
## 638   0032110648 3.871 1998     2906      1     3.619
## 643   0032113631 4.552 1998     2906      1     4.037
## 675   0032065176 4.592 1998     2906      1     3.607
## 676   0032065459 4.912 1998     2906      1     3.191
## 727   0031605853 3.997 1998     2906      1     3.012
## 728   0031606773 4.095 1998     2906      1     2.628
## 730   0031613267 4.082 1998     2906      1     3.567
## 943   0032605184 3.621 1999     2906      1     3.359
## 970   0033227121 3.809 1999     2906      1     3.283
## 998   0033192424 4.157 1999     2906      1     2.642
## 1000  0033192728 3.119 1999     2906      1     2.857
## 1025  0033160572 4.842 1999     2906      1     4.316
## 1309  0040756608 3.986 2000     2906      1     3.014
## 1321  0034329793 3.575 2000     2906      1     2.603
## 1870  0036357147 4.632 2002     2906      1     3.879
## 1871  0036365051 4.514 2002     2906      1     2.546
## 1872  0036559999 4.698 2002     2906      1     2.730
## 1873  0036560012 4.740 2002     2906      1     3.184
```

```
## 1876  0036633995 4.964 2002     2906      1     3.662
## 1884  0036834950 4.839 2002     2906      1     2.880
## 1885  0036981338 3.820 2002     2906      1     3.067
## 1890  0040675271 3.948 2002     2906      1     3.195
## 1931  0036730953 4.099 2002     2906      1     2.886
## 1932  0040081731 5.249 2002     2906      1     3.281
## 1936  0036675310 3.948 2002     2906      1     2.726
## 1962  0036517418 4.839 2002     2906      1     3.106
## 1964  0036519032 4.781 2002     2906      1     2.822
## 1986  0036593110 3.624 2002     2906      1     2.871
## 1999  0036678690 3.624 2002     2906      1     2.871
## 2011  0036886230 4.463 2002     2906      1     2.721
## 2016  0036887074 4.168 2002     2906      1     2.621
## 2021 85026135285 3.643 2002     2906      2     2.890
## 2448  0037362753 3.671 2003     2906      1     2.955
## 2453  1442310011 3.294 2003     2906      1     2.578
## 2872  7644225814 3.466 2004     2906      1     2.912
## 2900  7644230128 3.299 2004     2906      1     3.009
## 2919  4344577382 4.012 2004     2906      1     3.458
## 2955  1842539567 3.994 2004     2906      1     2.637
## 3223 27644548535 4.281 2005     2906      1     2.857
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
UniqueAuthors +
##     Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##     k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -2.0295 -0.5537 -0.0919  0.6279  4.3161
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         0.32705    0.10871    3.01   0.0026 **
## FirstAuthorFemale1  0.25426    0.05708    4.45  8.6e-06 ***
## LastAuthorFemale1   0.00916    0.05332    0.17   0.8636
## UniqueAuthors2      0.46977    0.03918   11.99  < 2e-16 ***
## UniqueAuthors3      0.80317    0.05507   14.58  < 2e-16 ***
## UniqueAuthors4      0.98901    0.06453   15.33  < 2e-16 ***
## UniqueAuthors5      1.21515    0.07153   16.99  < 2e-16 ***
## Year1997           -0.04277    0.14466   -0.30   0.7675
## Year1998           -0.07533    0.13103   -0.57   0.5654
## Year1999           -0.06459    0.13164   -0.49   0.6237
## Year2000            0.64504    0.19974    3.23   0.0013 **
## Year2001            0.37403    0.15359    2.44   0.0149 *
## Year2002            0.16222    0.13104    1.24   0.2158
## Year2003            0.12539    0.12242    1.02   0.3058
## Year2004           -0.03677    0.11664   -0.32   0.7526
```

```
## Year2005            0.03053    0.11710    0.26   0.7943
## Year2006           -0.03894    0.11631   -0.33   0.7378
## Year2007           -0.08130    0.11859   -0.69   0.4930
## Year2008            0.00456    0.11278    0.04   0.9677
## Year2009            0.03490    0.11293    0.31   0.7573
## Year2010           -0.00982    0.11300   -0.09   0.9307
## Year2011           -0.00958    0.11379   -0.08   0.9329
## Year2012           -0.08390    0.11343   -0.74   0.4595
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.728
## Multiple R-squared:  0.233,  Adjusted R-squared:  0.229
## Convergence in 17 IRWLS iterations
##
## Robustness weights:
##  9 observations c(129,239,244,248,264,385,649,655,1232)
##   are outliers with |weight| = 0 ( < 2.5e-05);
##  231 weights are ~= 1. The remaining 3826 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0009  0.8460  0.9430  0.8820  0.9780  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          2.46e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                    psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 3.029  1           1.741
## LastAuthorFemale  3.035  1           1.742
## Year              1.070 16           1.002
```

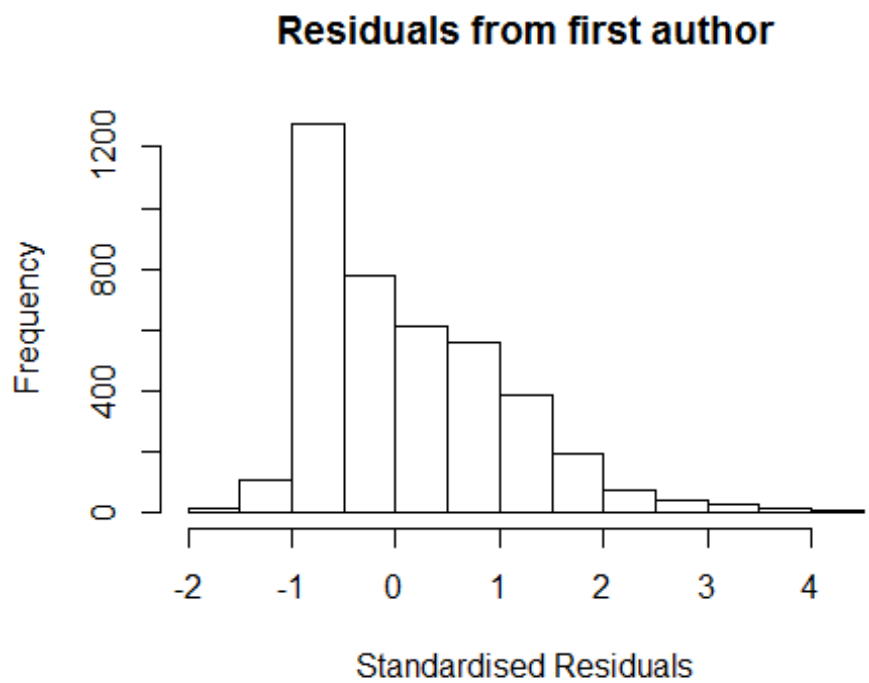



```
## [1] "List of  77 outliers with residuals above  2.5"
##         ScopusId  NLCS Year OneField Fields residuals
## 6     0030276739 3.458 1996     2906      1     2.642
## 16    0030228453 3.680 1996     2906      1     2.864
## 41    0029689219 3.737 1996     2906      1     3.451
## 52    0029692337 3.344 1996     2906      2     2.796
## 204   0242505996 3.218 1996     2902      5     2.670
## 242   0031279020 3.439 1997     2906      1     2.700
## 273   0031229882 4.173 1997     2906      1     3.434
## 274   0031229883 4.722 1997     2906      1     3.721
## 315   0031132148 4.483 1997     2906      1     3.744
## 317   0031132946 3.952 1997     2906      1     3.213
## 318   0031135082 4.667 1997     2906      1     3.928
## 387   0030638711 3.280 1997     2906      1     2.541
## 591   0032199161 2.868 1998     2906      1     2.688
## 616   0032162185 3.348 1998     2906      1     2.638
## 617   0032163901 3.532 1998     2906      1     2.560
## 638   0032110648 3.871 1998     2906      1     3.429
## 639   0032110728 3.224 1998     2906      1     2.514
## 643   0032113631 4.552 1998     2906      1     3.842
## 674   0032063376 4.172 1998     2906      1     3.462
## 675   0032065176 4.592 1998     2906      1     3.882
## 676   0032065459 4.912 1998     2906      1     3.940
## 679   0032067286 3.376 1998     2906      1     2.666
## 707   0032013132 3.376 1998     2906      1     2.666
## 727   0031605853 3.997 1998     2906      1     3.287
## 728   0031606773 4.095 1998     2906      1     3.653
```

```
## 729   0031608014 3.666 1998     2906      1     2.694
## 730   0031613267 4.082 1998     2906      1     3.372
## 943   0032605184 3.621 1999     2906      1     3.140
## 970   0033227121 3.809 1999     2906      1     3.059
## 998   0033192424 4.157 1999     2906      1     3.407
## 1000  0033192728 3.119 1999     2906      1     2.638
## 1025  0033160572 4.842 1999     2906      1     4.092
## 1062  0033088424 4.183 1999     2906      1     3.171
## 1309  0040756608 3.986 2000     2906      1     2.618
## 1673  0035514591 3.553 2001     2906      1     2.793
## 1689  0035409423 4.066 2001     2906      1     2.776
## 1870  0036357147 4.632 2002     2906      1     3.618
## 1871  0036365051 4.514 2002     2906      1     3.500
## 1872  0036559999 4.698 2002     2906      1     3.684
## 1873  0036560012 4.740 2002     2906      1     3.726
## 1876  0036633995 4.964 2002     2906      1     4.481
## 1883  0036834614 3.864 2002     2906      1     2.850
## 1884  0036834950 4.839 2002     2906      1     3.563
## 1885  0036981338 3.820 2002     2906      1     2.806
## 1890  0040675271 3.948 2002     2906      1     2.934
## 1931  0036730953 4.099 2002     2906      1     2.823
## 1932  0040081731 5.249 2002     2906      1     4.235
## 1936  0036675310 3.948 2002     2906      1     2.934
## 1962  0036517418 4.839 2002     2906      1     3.563
## 1964  0036519032 4.781 2002     2906      1     3.505
## 1965  0036519092 3.907 2002     2906      1     2.631
## 1977  0036155135 4.026 2002     2906      1     2.750
## 1980  0036157312 3.322 2002     2906      1     2.577
## 1986  0036593110 3.624 2002     2906      1     2.610
## 1999  0036678690 3.624 2002     2906      1     2.610
## 2011  0036886230 4.463 2002     2906      1     3.449
## 2016  0036887074 4.168 2002     2906      1     2.892
## 2021 85026135285 3.643 2002     2906      2     2.629
## 2028 85026149648 3.558 2002     2906      2     2.544
## 2389  2142856146 4.241 2003     2906      1     3.322
## 2398  1542649646 3.456 2003     2906      1     2.537
## 2406  0642313730 3.924 2003     2906      1     3.536
## 2410  0038486780 3.621 2003     2906      1     2.702
## 2411  0038486782 4.391 2003     2906      1     3.210
## 2448  0037362753 3.671 2003     2906      1     2.752
## 2449  0038742958 3.621 2003     2906      1     2.702
## 2490  0038175222 3.521 2003     2906      3     2.602
## 2872  7644225814 3.466 2004     2906      1     2.709
## 2874  7644237468 3.648 2004     2906      1     3.159
## 2900  7644230128 3.299 2004     2906      1     2.810
## 2919  4344577382 4.012 2004     2906      1     3.255
## 2940  3042778188 3.339 2004     2906      1     2.582
## 2955  1842539567 3.994 2004     2906      1     3.237
## 3223 27644548535 4.281 2005     2906      1     3.422
## 3291 12744281198 3.159 2005     2906      1     2.830
```

```
## 3303 12144271596 3.905 2005     2906      1     3.046
## 5069 77956697132 2.978 2010     2906      1     2.638
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##    Min     1Q Median     3Q    Max
## -1.899 -0.739 -0.102  0.715  4.481
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         0.54757    0.11979    4.57  5.0e-06 ***
## FirstAuthorFemale1  0.53050    0.06432    8.25  < 2e-16 ***
## LastAuthorFemale1  -0.26196    0.05969   -4.39  1.2e-05 ***
## Year1997           -0.07754    0.15650   -0.50   0.6203
## Year1998           -0.10580    0.14409   -0.73   0.4628
## Year1999           -0.06611    0.14250   -0.46   0.6427
## Year2000            0.82058    0.20491    4.00  6.3e-05 ***
## Year2001            0.47415    0.16710    2.84   0.0046 **
## Year2002            0.19778    0.14023    1.41   0.1585
## Year2003            0.10283    0.13217    0.78   0.4366
## Year2004           -0.05887    0.12560   -0.47   0.6393
## Year2005            0.04325    0.12547    0.34   0.7303
## Year2006            0.09835    0.12496    0.79   0.4313
## Year2007            0.00772    0.12839    0.06   0.9520
## Year2008            0.02888    0.12181    0.24   0.8126
## Year2009            0.14237    0.12109    1.18   0.2398
## Year2010            0.05482    0.12151    0.45   0.6519
## Year2011            0.12398    0.12067    1.03   0.3042
## Year2012            0.08847    0.12085    0.73   0.4642
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.855
## Multiple R-squared:  0.0473, Adjusted R-squared:  0.0431
## Convergence in 15 IRWLS iterations
##
## Robustness weights:
##  3 observations c(385,655,704) are outliers with |weight| = 0 ( < 2.5e-
05);
##  206 weights are ~= 1. The remaining 3857 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0011  0.8890  0.9300  0.8950  0.9760  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
```

```
##          1.00e-07          1.00e-07          2.46e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.032  1           1.016
## Year              1.032 16           1.001
```

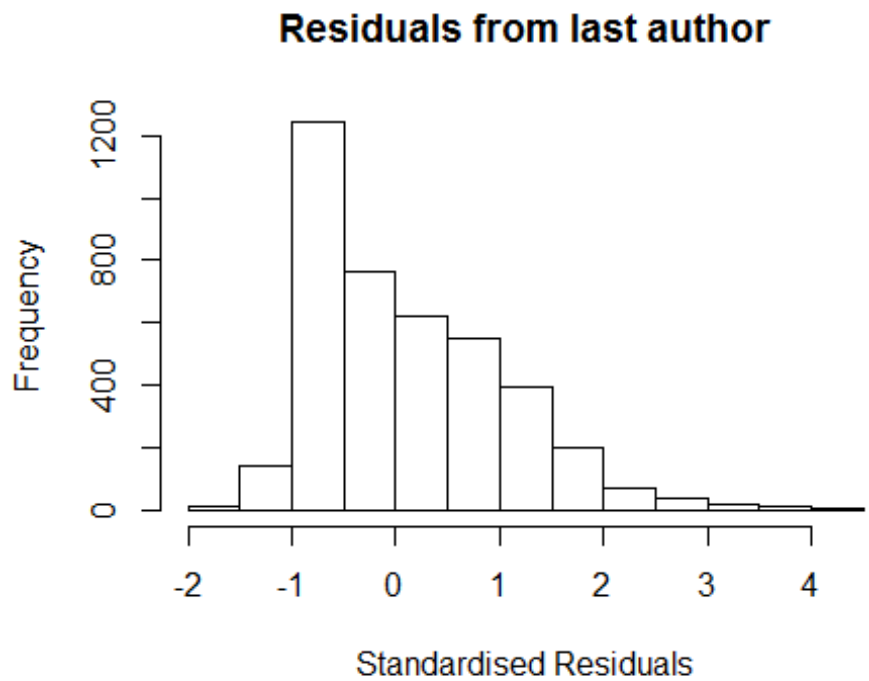



```
## [1] "List of  77 outliers with residuals above  2.5"
##         ScopusId  NLCS Year OneField Fields residuals
## 6     0030276739 3.458 1996     2906      1     2.642
## 16    0030228453 3.680 1996     2906      1     2.864
## 41    0029689219 3.737 1996     2906      1     3.451
## 52    0029692337 3.344 1996     2906      2     2.796
## 204   0242505996 3.218 1996     2902      5     2.670
## 242   0031279020 3.439 1997     2906      1     2.700
## 273   0031229882 4.173 1997     2906      1     3.434
## 274   0031229883 4.722 1997     2906      1     3.721
```

```
## 315   0031132148 4.483 1997     2906      1     3.744
## 317   0031132946 3.952 1997     2906      1     3.213
## 318   0031135082 4.667 1997     2906      1     3.928
## 387   0030638711 3.280 1997     2906      1     2.541
## 591   0032199161 2.868 1998     2906      1     2.688
## 616   0032162185 3.348 1998     2906      1     2.638
## 617   0032163901 3.532 1998     2906      1     2.560
## 638   0032110648 3.871 1998     2906      1     3.429
## 639   0032110728 3.224 1998     2906      1     2.514
## 643   0032113631 4.552 1998     2906      1     3.842
## 674   0032063376 4.172 1998     2906      1     3.462
## 675   0032065176 4.592 1998     2906      1     3.882
## 676   0032065459 4.912 1998     2906      1     3.940
## 679   0032067286 3.376 1998     2906      1     2.666
## 707   0032013132 3.376 1998     2906      1     2.666
## 727   0031605853 3.997 1998     2906      1     3.287
## 728   0031606773 4.095 1998     2906      1     3.653
## 729   0031608014 3.666 1998     2906      1     2.694
## 730   0031613267 4.082 1998     2906      1     3.372
## 943   0032605184 3.621 1999     2906      1     3.140
## 970   0033227121 3.809 1999     2906      1     3.059
## 998   0033192424 4.157 1999     2906      1     3.407
## 1000  0033192728 3.119 1999     2906      1     2.638
## 1025  0033160572 4.842 1999     2906      1     4.092
## 1062  0033088424 4.183 1999     2906      1     3.171
## 1309  0040756608 3.986 2000     2906      1     2.618
## 1673  0035514591 3.553 2001     2906      1     2.793
## 1689  0035409423 4.066 2001     2906      1     2.776
## 1870  0036357147 4.632 2002     2906      1     3.618
## 1871  0036365051 4.514 2002     2906      1     3.500
## 1872  0036559999 4.698 2002     2906      1     3.684
## 1873  0036560012 4.740 2002     2906      1     3.726
## 1876  0036633995 4.964 2002     2906      1     4.481
## 1883  0036834614 3.864 2002     2906      1     2.850
## 1884  0036834950 4.839 2002     2906      1     3.563
## 1885  0036981338 3.820 2002     2906      1     2.806
## 1890  0040675271 3.948 2002     2906      1     2.934
## 1931  0036730953 4.099 2002     2906      1     2.823
## 1932  0040081731 5.249 2002     2906      1     4.235
## 1936  0036675310 3.948 2002     2906      1     2.934
## 1962  0036517418 4.839 2002     2906      1     3.563
## 1964  0036519032 4.781 2002     2906      1     3.505
## 1965  0036519092 3.907 2002     2906      1     2.631
## 1977  0036155135 4.026 2002     2906      1     2.750
## 1980  0036157312 3.322 2002     2906      1     2.577
## 1986  0036593110 3.624 2002     2906      1     2.610
## 1999  0036678690 3.624 2002     2906      1     2.610
## 2011  0036886230 4.463 2002     2906      1     3.449
## 2016  0036887074 4.168 2002     2906      1     2.892
## 2021 85026135285 3.643 2002     2906      2     2.629
```

```
## 2028 85026149648 3.558 2002     2906      2     2.544
## 2389  2142856146 4.241 2003     2906      1     3.322
## 2398  1542649646 3.456 2003     2906      1     2.537
## 2406  0642313730 3.924 2003     2906      1     3.536
## 2410  0038486780 3.621 2003     2906      1     2.702
## 2411  0038486782 4.391 2003     2906      1     3.210
## 2448  0037362753 3.671 2003     2906      1     2.752
## 2449  0038742958 3.621 2003     2906      1     2.702
## 2490  0038175222 3.521 2003     2906      3     2.602
## 2872  7644225814 3.466 2004     2906      1     2.709
## 2874  7644237468 3.648 2004     2906      1     3.159
## 2900  7644230128 3.299 2004     2906      1     2.810
## 2919  4344577382 4.012 2004     2906      1     3.255
## 2940  3042778188 3.339 2004     2906      1     2.582
## 2955  1842539567 3.994 2004     2906      1     3.237
## 3223 27644548535 4.281 2005     2906      1     3.422
## 3291 12744281198 3.159 2005     2906      1     2.830
## 3303 12144271596 3.905 2005     2906      1     3.046
## 5069 77956697132 2.978 2010     2906      1     2.638
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##    Min     1Q Median     3Q    Max
## -1.660 -0.772 -0.114  0.714  4.263
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)          0.5090     0.1193    4.27  2.0e-05 ***
## FirstAuthorFemale1   0.3242     0.0366    8.86  < 2e-16 ***
## Year1997            -0.0565     0.1570   -0.36    0.719
## Year1998            -0.1054     0.1457   -0.72    0.469
## Year1999            -0.0614     0.1426   -0.43    0.667
## Year2000             0.8269     0.2028    4.08  4.6e-05 ***
## Year2001             0.5142     0.1665    3.09    0.002 **
## Year2002             0.1920     0.1405    1.37    0.172
## Year2003             0.1117     0.1318    0.85    0.397
## Year2004            -0.0612     0.1263   -0.48    0.628
## Year2005             0.0454     0.1261    0.36    0.719
## Year2006             0.0969     0.1254    0.77    0.440
## Year2007             0.0125     0.1288    0.10    0.923
## Year2008             0.0362     0.1224    0.30    0.767
## Year2009             0.1459     0.1216    1.20    0.230
## Year2010             0.0584     0.1221    0.48    0.633
## Year2011             0.1267     0.1212    1.05    0.296
## Year2012             0.1004     0.1211    0.83    0.407
## ---
```

```
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.861
## Multiple R-squared:  0.0396, Adjusted R-squared:  0.0356
## Convergence in 15 IRWLS iterations
##
## Robustness weights:
##  4 observations c(249,385,655,704)
##   are outliers with |weight| = 0 ( < 2.5e-05);
##  218 weights are ~= 1. The remaining 3844 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0019  0.8890  0.9280  0.8950  0.9770  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          2.46e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                    psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##                   GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.037  1           1.018
## Year             1.037 16           1.001
```

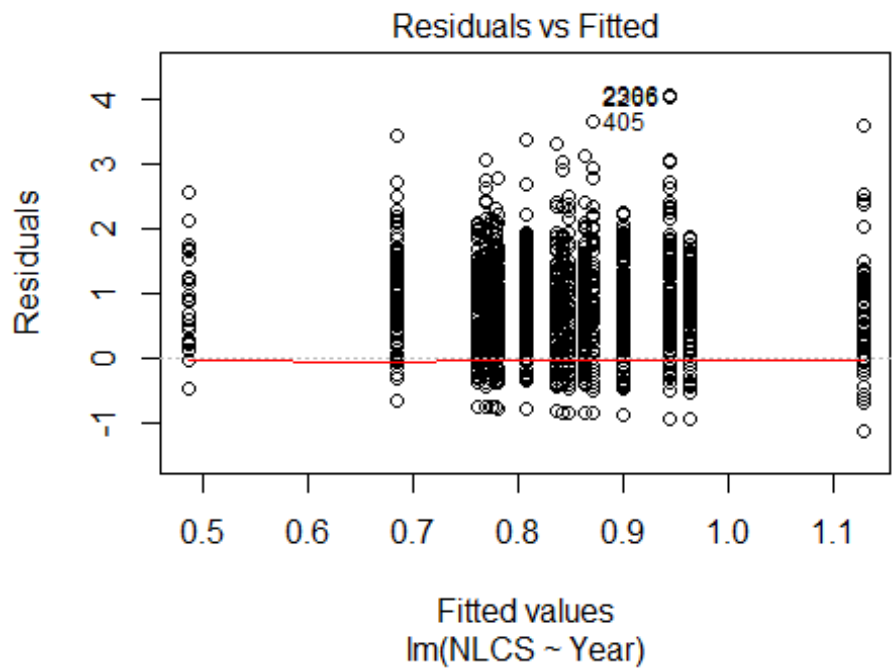



```
## [1] "List of  77 outliers with residuals above  2.5"
##         ScopusId  NLCS Year OneField Fields residuals
## 6     0030276739 3.458 1996     2906      1     2.642
## 16    0030228453 3.680 1996     2906      1     2.864
## 41    0029689219 3.737 1996     2906      1     3.451
## 52    0029692337 3.344 1996     2906      2     2.796
## 204   0242505996 3.218 1996     2902      5     2.670
## 242   0031279020 3.439 1997     2906      1     2.700
## 273   0031229882 4.173 1997     2906      1     3.434
## 274   0031229883 4.722 1997     2906      1     3.721
## 315   0031132148 4.483 1997     2906      1     3.744
## 317   0031132946 3.952 1997     2906      1     3.213
## 318   0031135082 4.667 1997     2906      1     3.928
## 387   0030638711 3.280 1997     2906      1     2.541
## 591   0032199161 2.868 1998     2906      1     2.688
## 616   0032162185 3.348 1998     2906      1     2.638
## 617   0032163901 3.532 1998     2906      1     2.560
## 638   0032110648 3.871 1998     2906      1     3.429
## 639   0032110728 3.224 1998     2906      1     2.514
## 643   0032113631 4.552 1998     2906      1     3.842
## 674   0032063376 4.172 1998     2906      1     3.462
## 675   0032065176 4.592 1998     2906      1     3.882
## 676   0032065459 4.912 1998     2906      1     3.940
## 679   0032067286 3.376 1998     2906      1     2.666
## 707   0032013132 3.376 1998     2906      1     2.666
## 727   0031605853 3.997 1998     2906      1     3.287
## 728   0031606773 4.095 1998     2906      1     3.653
```

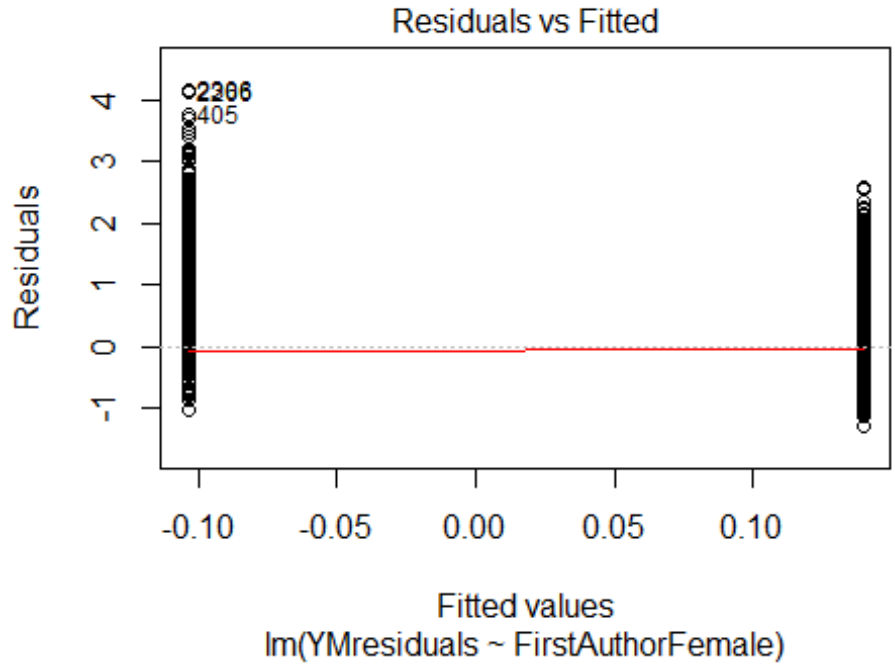
```
## 729   0031608014 3.666 1998     2906      1     2.694
## 730   0031613267 4.082 1998     2906      1     3.372
## 943   0032605184 3.621 1999     2906      1     3.140
## 970   0033227121 3.809 1999     2906      1     3.059
## 998   0033192424 4.157 1999     2906      1     3.407
## 1000  0033192728 3.119 1999     2906      1     2.638
## 1025  0033160572 4.842 1999     2906      1     4.092
## 1062  0033088424 4.183 1999     2906      1     3.171
## 1309  0040756608 3.986 2000     2906      1     2.618
## 1673  0035514591 3.553 2001     2906      1     2.793
## 1689  0035409423 4.066 2001     2906      1     2.776
## 1870  0036357147 4.632 2002     2906      1     3.618
## 1871  0036365051 4.514 2002     2906      1     3.500
## 1872  0036559999 4.698 2002     2906      1     3.684
## 1873  0036560012 4.740 2002     2906      1     3.726
## 1876  0036633995 4.964 2002     2906      1     4.481
## 1883  0036834614 3.864 2002     2906      1     2.850
## 1884  0036834950 4.839 2002     2906      1     3.563
## 1885  0036981338 3.820 2002     2906      1     2.806
## 1890  0040675271 3.948 2002     2906      1     2.934
## 1931  0036730953 4.099 2002     2906      1     2.823
## 1932  0040081731 5.249 2002     2906      1     4.235
## 1936  0036675310 3.948 2002     2906      1     2.934
## 1962  0036517418 4.839 2002     2906      1     3.563
## 1964  0036519032 4.781 2002     2906      1     3.505
## 1965  0036519092 3.907 2002     2906      1     2.631
## 1977  0036155135 4.026 2002     2906      1     2.750
## 1980  0036157312 3.322 2002     2906      1     2.577
## 1986  0036593110 3.624 2002     2906      1     2.610
## 1999  0036678690 3.624 2002     2906      1     2.610
## 2011  0036886230 4.463 2002     2906      1     3.449
## 2016  0036887074 4.168 2002     2906      1     2.892
## 2021 85026135285 3.643 2002     2906      2     2.629
## 2028 85026149648 3.558 2002     2906      2     2.544
## 2389  2142856146 4.241 2003     2906      1     3.322
## 2398  1542649646 3.456 2003     2906      1     2.537
## 2406  0642313730 3.924 2003     2906      1     3.536
## 2410  0038486780 3.621 2003     2906      1     2.702
## 2411  0038486782 4.391 2003     2906      1     3.210
## 2448  0037362753 3.671 2003     2906      1     2.752
## 2449  0038742958 3.621 2003     2906      1     2.702
## 2490  0038175222 3.521 2003     2906      3     2.602
## 2872  7644225814 3.466 2004     2906      1     2.709
## 2874  7644237468 3.648 2004     2906      1     3.159
## 2900  7644230128 3.299 2004     2906      1     2.810
## 2919  4344577382 4.012 2004     2906      1     3.255
## 2940  3042778188 3.339 2004     2906      1     2.582
## 2955  1842539567 3.994 2004     2906      1     3.237
## 3223 27644548535 4.281 2005     2906      1     3.422
## 3291 12744281198 3.159 2005     2906      1     2.830
```

```
## 3303 12144271596 3.905 2005     2906      1     3.046
## 5069 77956697132 2.978 2010     2906      1     2.638
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.6527 -0.7653 -0.0954  0.7220  4.3027
##
## Coefficients:
##                   Estimate Std. Error t value Pr(>|t|)
## (Intercept)        0.72341    0.12325    5.87  4.7e-09 ***
## LastAuthorFemale1  0.08339    0.03914    2.13   0.0332 *
## Year1997          -0.05136    0.15691   -0.33   0.7434
## Year1998          -0.11409    0.14782   -0.77   0.4403
## Year1999          -0.07782    0.14533   -0.54   0.5923
## Year2000           0.84593    0.20202    4.19  2.9e-05 ***
## Year2001           0.53713    0.16880    3.18   0.0015 **
## Year2002           0.19141    0.14151    1.35   0.1762
## Year2003           0.08044    0.13435    0.60   0.5494
## Year2004          -0.07256    0.12779   -0.57   0.5702
## Year2005           0.04192    0.12742    0.33   0.7422
## Year2006           0.07547    0.12654    0.60   0.5509
## Year2007          -0.00297    0.13004   -0.02   0.9818
## Year2008           0.02109    0.12435    0.17   0.8653
## Year2009           0.14130    0.12313    1.15   0.2512
## Year2010           0.05703    0.12349    0.46   0.6442
## Year2011           0.11767    0.12281    0.96   0.3381
## Year2012           0.08667    0.12265    0.71   0.4798
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.864
## Multiple R-squared:  0.0249, Adjusted R-squared:  0.0208
## Convergence in 15 IRWLS iterations
##
## Robustness weights:
##  4 observations c(122,249,385,704)
##   are outliers with |weight| = 0 ( < 2.5e-05);
##  229 weights are ~= 1. The remaining 3833 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0017  0.8930  0.9230  0.8940  0.9770  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          2.46e-05          1.82e-12
```

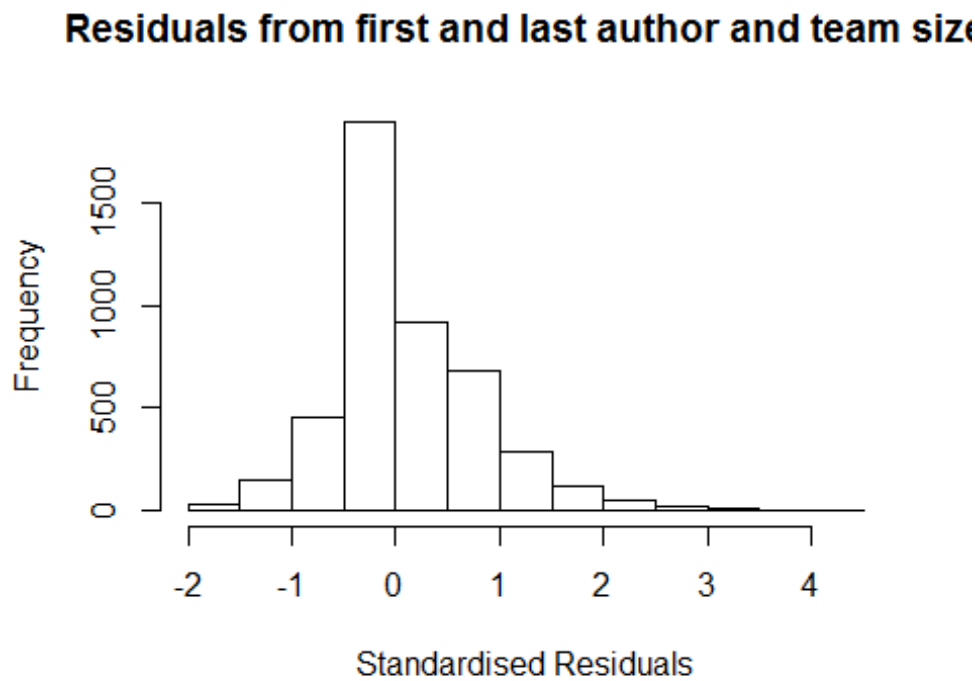
```
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Sample size for the above analysis:  4066"
## [1] ""
## [1] ""
## [1] "####################################"
## [1] "Analysis of AJSC 2907"
## [1] "####################################"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  219  414  387  398  411  313  547  496  487  339  305  308  418  480  557
## 2011 2012
##  446  447
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  140  209  189  108  139  181  299  285  399  267  270  265  376  432  494
## 2011 2012
##  407  392
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  134  202  178  106  125  167  293  271  389  249  252  256  363  406  462
## 2011 2012
##  387  373
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
##  Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 72, df = 16, p-value = 4e-09
```




```
##
##  Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 2.3, df = 1, p-value = 0.1
```



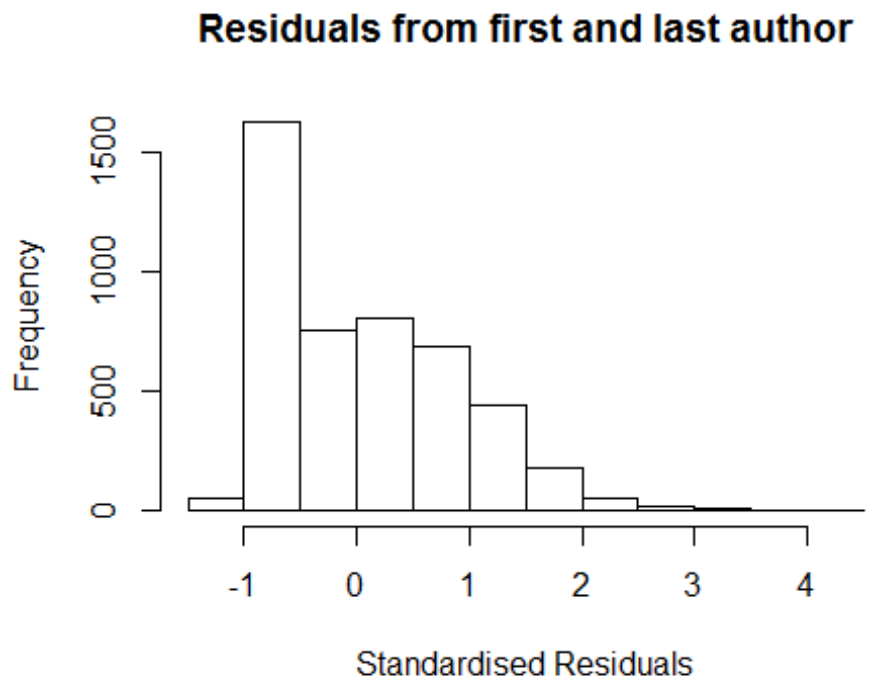

```
## [1] "Female first author team size 2018 geometric mean: 2.5949716491686"
## [1] "Male first author team size 2018 geometric mean: 4.31533112513958"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 8200, p-value = 1e-08
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 2.43938221644925"
## [1] "Male last author team size 2018 geometric mean: 4.44388556496607"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 7500, p-value = 5e-12
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 3.204  1           1.790
## LastAuthorFemale  3.324  1           1.823
## UniqueAuthors     1.325  4           1.036
## Year              1.241 16           1.007
```




```
## [1] "List of  36 outliers with residuals above  2.5"
##         ScopusId  NLCS Year OneField Fields residuals
## 83    0030095790 2.987 1996     2906      2     2.742
## 384   0031064803 3.798 1997     2907      1     3.144
## 405   0031155309 4.527 1997     2907      1     3.873
## 423   0031203501 3.645 1997     2907      1     3.433
## 437   0031254508 3.040 1997     2907      1     2.828
## 451   0031298130 2.848 1997     2907      1     2.636
## 455   0031302585 2.848 1997     2907      1     2.636
## 467   0031309414 3.202 1997     2907      1     2.548
## 507   0031180540 3.068 1997     2906      2     2.856
## 802   0031990054 3.733 1998     2907      1     2.591
## 847   0032133322 3.868 1998     2907      1     3.562
## 1645  0034169785 3.538 2000     2907      1     3.113
## 1658  0034200397 4.716 2000     2907      1     3.456
## 1673  0034239103 3.484 2000     2907      1     2.617
## 1686  0034288015 3.639 2000     2907      1     3.214
## 1976  0035373746 2.850 2001     2907      1     2.579
## 2266  0036008249 4.970 2002     2907      1     4.186
## 2277  0036012829 4.003 2002     2907      1     2.816
## 2278  0036012830 3.956 2002     2907      1     3.172
## 2306  0036776261 4.988 2002     2907      1     3.366
## 2454 85026135285 3.643 2002     2906      2     3.301
## 2461 85026149648 3.558 2002     2906      2     2.774
## 2777  0037323743 2.897 2003     2907      1     2.633
## 2795  0037398605 3.515 2003     2907      1     3.251
## 2796  0037398606 3.810 2003     2907      1     3.546
```

```
## 2830  0347588237 3.366 2003     2907      1     3.102
## 2942 85026140202 3.147 2003     2906      2     2.883
## 3391  1342267540 4.111 2004     2907      1     2.950
## 3402  1942517874 3.396 2004     2907      1     2.824
## 3449  3042563662 2.965 2004     2906      2     2.852
## 3564 85026136040 2.864 2004     2906      2     2.751
## 3572 85026145804 2.709 2004     2906      2     2.596
## 4181 33646805797 4.136 2006     2907      1     2.645
## 4243 31344447300 3.172 2006     2907      1     2.510
## 4246 31344456993 3.220 2006     2907      1     2.982
## 5822 77956649584 4.174 2010     2907      1     2.634
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
UniqueAuthors +
##     Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##     k.max = 1000))
##  \--> method = "MM"
## Residuals:
##    Min     1Q Median     3Q    Max
## -1.723 -0.318 -0.113  0.502  4.186
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         0.26313    0.06267    4.20  2.7e-05 ***
## FirstAuthorFemale1 -0.02719    0.03518   -0.77    0.440
## LastAuthorFemale1   0.00952    0.03568    0.27    0.790
## UniqueAuthors2      0.44204    0.03424   12.91  < 2e-16 ***
## UniqueAuthors3      0.84520    0.05214   16.21  < 2e-16 ***
## UniqueAuthors4      1.05809    0.04757   22.24  < 2e-16 ***
## UniqueAuthors5      1.28028    0.03687   34.72  < 2e-16 ***
## Year1997           -0.03321    0.08374   -0.40    0.692
## Year1998            0.06074    0.08190    0.74    0.458
## Year1999           -0.16293    0.07935   -2.05    0.040 *
## Year2000            0.17913    0.09253    1.94    0.053 .
## Year2001            0.02594    0.07625    0.34    0.734
## Year2002            0.09672    0.07286    1.33    0.184
## Year2003            0.01870    0.07063    0.26    0.791
## Year2004           -0.13272    0.07255   -1.83    0.067 .
## Year2005            0.05174    0.07143    0.72    0.469
## Year2006           -0.02543    0.07202   -0.35    0.724
## Year2007           -0.04803    0.07092   -0.68    0.498
## Year2008            0.02435    0.06796    0.36    0.720
## Year2009            0.05445    0.06852    0.79    0.427
## Year2010            0.01410    0.06648    0.21    0.832
## Year2011            0.11054    0.07084    1.56    0.119
## Year2012            0.10499    0.07026    1.49    0.135
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

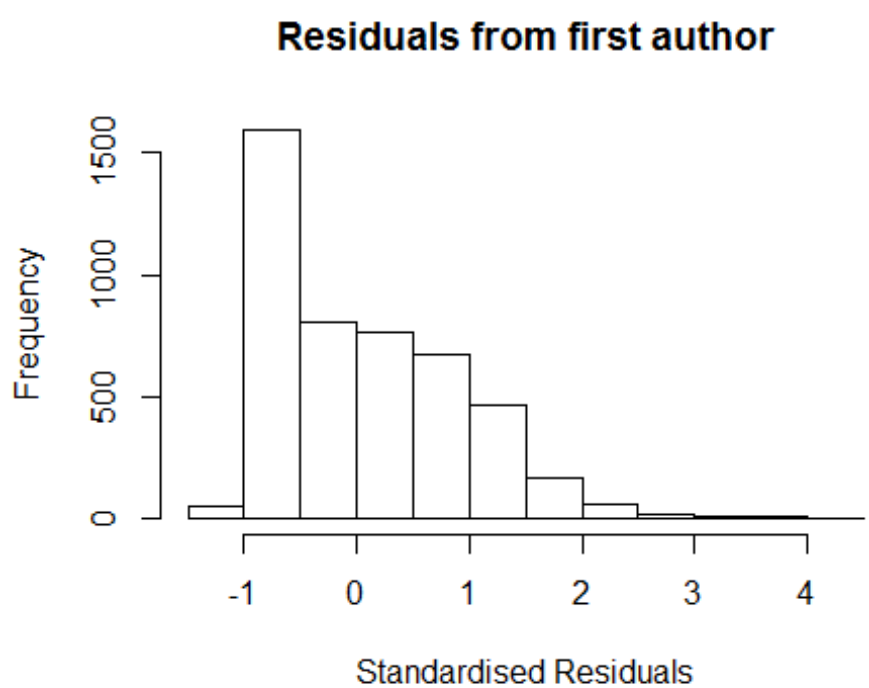
```
##
## Robust residual standard error: 0.512
## Multiple R-squared:  0.425,  Adjusted R-squared:  0.422
## Convergence in 17 IRWLS iterations
##
## Robustness weights:
##  45 observations
c(17,173,185,193,198,204,207,210,216,223,236,379,398,642,650,657,662,780,793,
973,984,985,998,1007,1010,1078,1126,1133,1225,1242,1243,1268,1348,1584,1594,1
631,1734,1742,2167,2217,2220,3410,3469,3549,4534)
##   are outliers with |weight| <= 1.5e-06 ( < 2.2e-05);
##  140 weights are ~= 1. The remaining 4428 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0002  0.8220  0.9560  0.8630  0.9770  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          2.17e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 2.380  1           1.543
## LastAuthorFemale  2.391  1           1.546
## Year              1.037 16           1.001
```




```
## [1] "List of  34 outliers with residuals above  2.5"
##         ScopusId  NLCS Year OneField Fields residuals
## 384   0031064803 3.798 1997     2907      1     3.231
## 405   0031155309 4.527 1997     2907      1     3.960
## 423   0031203501 3.645 1997     2907      1     3.078
## 467   0031309414 3.202 1997     2907      1     2.635
## 507   0031180540 3.068 1997     2906      2     2.501
## 802   0031990054 3.733 1998     2907      1     2.912
## 847   0032133322 3.868 1998     2907      1     3.313
## 1645  0034169785 3.538 2000     2907      1     2.707
## 1658  0034200397 4.716 2000     2907      1     3.620
## 1673  0034239103 3.484 2000     2907      1     2.653
## 1686  0034288015 3.639 2000     2907      1     2.808
## 1954  0035314990 3.542 2001     2907      1     3.028
## 2266  0036008249 4.970 2002     2907      1     4.361
## 2277  0036012829 4.003 2002     2907      1     3.394
## 2278  0036012830 3.956 2002     2907      1     3.347
## 2306  0036776261 4.988 2002     2907      1     4.379
## 2311  0036923166 3.278 2002     2907      1     2.669
## 2454 85026135285 3.643 2002     2906      2     3.034
## 2457 85026139222 3.255 2002     2906      2     2.646
## 2461 85026149648 3.558 2002     2906      2     2.949
## 2760  0042829280 3.196 2003     2705      3     2.744
## 2795  0037398605 3.515 2003     2907      1     3.063
## 2796  0037398606 3.810 2003     2907      1     3.358
## 2830  0347588237 3.366 2003     2907      1     2.914
## 2942 85026140202 3.147 2003     2906      2     2.695
```

```
## 3391  1342267540 4.111 2004     2907      1     3.444
## 3397  1342331383 3.157 2004     2907      1     2.756
## 3402  1942517874 3.396 2004     2907      1     2.641
## 3449  3042563662 2.965 2004     2906      2     2.564
## 4181 33646805797 4.136 2006     2907      1     3.262
## 4243 31344447300 3.172 2006     2907      1     2.563
## 5493 61649109012 3.256 2009     2907      1     2.562
## 5822 77956649584 4.174 2010     2907      1     3.644
## 5904 77954533425 3.488 2010     2907      1     2.869
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.1846 -0.5877 -0.0692  0.6497  4.3794
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         0.93095    0.07137   13.04  < 2e-16 ***
## FirstAuthorFemale1 -0.08883    0.04161   -2.13  0.03283 *
## LastAuthorFemale1  -0.26523    0.04055   -6.54  6.8e-11 ***
## Year1997           -0.00996    0.09449   -0.11  0.91610
## Year1998           -0.02151    0.09329   -0.23  0.81764
## Year1999           -0.30154    0.09018   -3.34  0.00083 ***
## Year2000            0.25362    0.11205    2.26  0.02365 *
## Year2001           -0.06257    0.09306   -0.67  0.50139
## Year2002            0.03171    0.08854    0.36  0.72024
## Year2003           -0.12471    0.08537   -1.46  0.14411
## Year2004           -0.17556    0.07964   -2.20  0.02754 *
## Year2005            0.01079    0.08210    0.13  0.89546
## Year2006            0.03207    0.08442    0.38  0.70404
## Year2007           -0.06091    0.08381   -0.73  0.46741
## Year2008           -0.02351    0.08199   -0.29  0.77429
## Year2009            0.02868    0.07962    0.36  0.71870
## Year2010           -0.04667    0.07864   -0.59  0.55295
## Year2011            0.06744    0.08514    0.79  0.42834
## Year2012            0.19055    0.08117    2.35  0.01894 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.734
## Multiple R-squared:  0.0578, Adjusted R-squared:  0.0541
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
##  6 observations c(185,650,973,1007,1584,3469)
##   are outliers with |weight| = 0 ( < 2.2e-05);
```

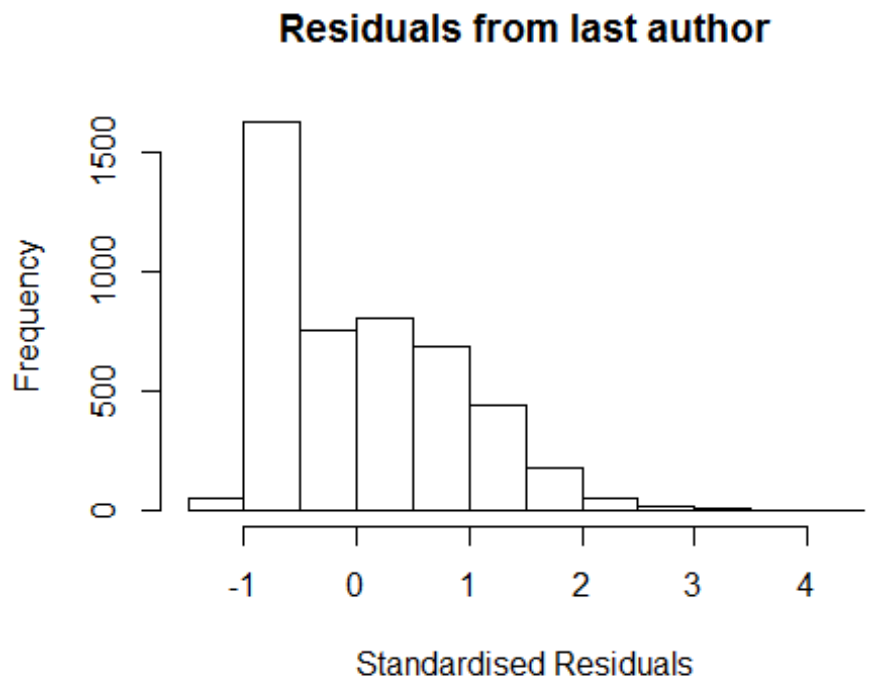
```
##  256 weights are ~= 1. The remaining 4351 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0007  0.8650  0.9380  0.8910  0.9690  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          2.17e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.015  1           1.007
## Year              1.015 16           1.000
```




```
## [1] "List of  34 outliers with residuals above  2.5"
##         ScopusId  NLCS Year OneField Fields residuals
## 384   0031064803 3.798 1997     2907      1     3.231
```

```
## 405   0031155309 4.527 1997     2907      1     3.960
## 423   0031203501 3.645 1997     2907      1     3.078
## 467   0031309414 3.202 1997     2907      1     2.635
## 507   0031180540 3.068 1997     2906      2     2.501
## 802   0031990054 3.733 1998     2907      1     2.912
## 847   0032133322 3.868 1998     2907      1     3.313
## 1645  0034169785 3.538 2000     2907      1     2.707
## 1658  0034200397 4.716 2000     2907      1     3.620
## 1673  0034239103 3.484 2000     2907      1     2.653
## 1686  0034288015 3.639 2000     2907      1     2.808
## 1954  0035314990 3.542 2001     2907      1     3.028
## 2266  0036008249 4.970 2002     2907      1     4.361
## 2277  0036012829 4.003 2002     2907      1     3.394
## 2278  0036012830 3.956 2002     2907      1     3.347
## 2306  0036776261 4.988 2002     2907      1     4.379
## 2311  0036923166 3.278 2002     2907      1     2.669
## 2454 85026135285 3.643 2002     2906      2     3.034
## 2457 85026139222 3.255 2002     2906      2     2.646
## 2461 85026149648 3.558 2002     2906      2     2.949
## 2760  0042829280 3.196 2003     2705      3     2.744
## 2795  0037398605 3.515 2003     2907      1     3.063
## 2796  0037398606 3.810 2003     2907      1     3.358
## 2830  0347588237 3.366 2003     2907      1     2.914
## 2942 85026140202 3.147 2003     2906      2     2.695
## 3391  1342267540 4.111 2004     2907      1     3.444
## 3397  1342331383 3.157 2004     2907      1     2.756
## 3402  1942517874 3.396 2004     2907      1     2.641
## 3449  3042563662 2.965 2004     2906      2     2.564
## 4181 33646805797 4.136 2006     2907      1     3.262
## 4243 31344447300 3.172 2006     2907      1     2.563
## 5493 61649109012 3.256 2009     2907      1     2.562
## 5822 77956649584 4.174 2010     2907      1     3.644
## 5904 77954533425 3.488 2010     2907      1     2.869
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.1590 -0.6116 -0.0826  0.6548  4.3475
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         0.88478    0.07054   12.54   <2e-16 ***
## FirstAuthorFemale1 -0.28438    0.02701  -10.53   <2e-16 ***
## Year1997            0.01122    0.09389    0.12   0.9049
## Year1998           -0.00271    0.09393   -0.03   0.9770
## Year1999           -0.29328    0.09015   -3.25   0.0011 **
```

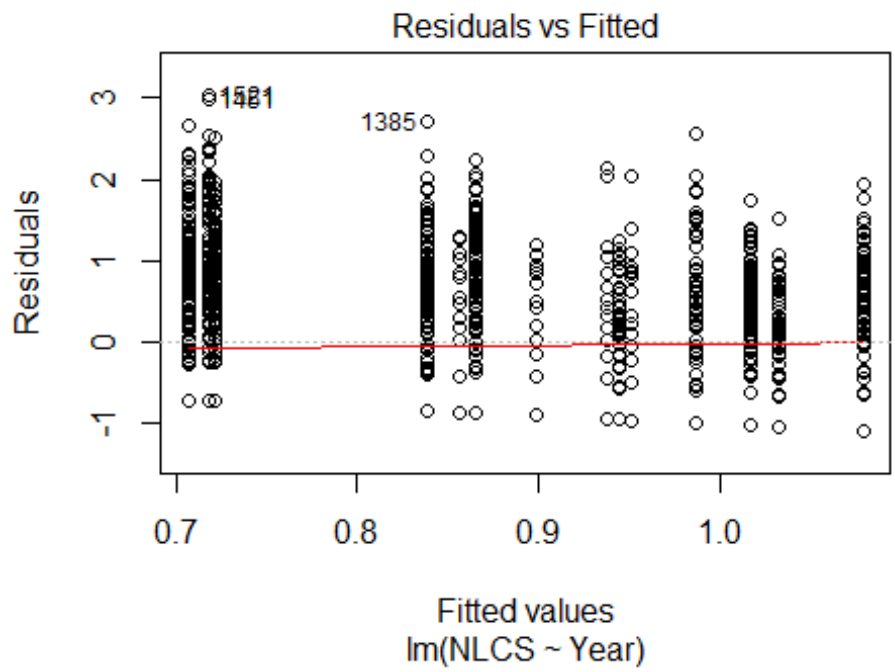
```
## Year2000            0.27418    0.11294    2.43   0.0152 *
## Year2001           -0.05453    0.09347   -0.58   0.5596
## Year2002            0.04010    0.08911    0.45   0.6527
## Year2003           -0.10918    0.08462   -1.29   0.1970
## Year2004           -0.16907    0.07929   -2.13   0.0330 *
## Year2005            0.01373    0.08217    0.17   0.8673
## Year2006            0.04098    0.08548    0.48   0.6317
## Year2007           -0.03891    0.08407   -0.46   0.6435
## Year2008           -0.01684    0.08212   -0.21   0.8375
## Year2009            0.04248    0.07963    0.53   0.5937
## Year2010           -0.03308    0.07856   -0.42   0.6737
## Year2011            0.07357    0.08493    0.87   0.3864
## Year2012            0.20462    0.08107    2.52   0.0116 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.741
## Multiple R-squared:  0.0459, Adjusted R-squared:  0.0424
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
##  7 observations c(185,650,973,1007,1584,2167,3469)
##   are outliers with |weight| = 0 ( < 2.2e-05);
##  194 weights are ~= 1. The remaining 4412 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0038  0.8710  0.9330  0.8930  0.9690  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          2.17e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##                   GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.021  1           1.010
## Year             1.021 16           1.001
```



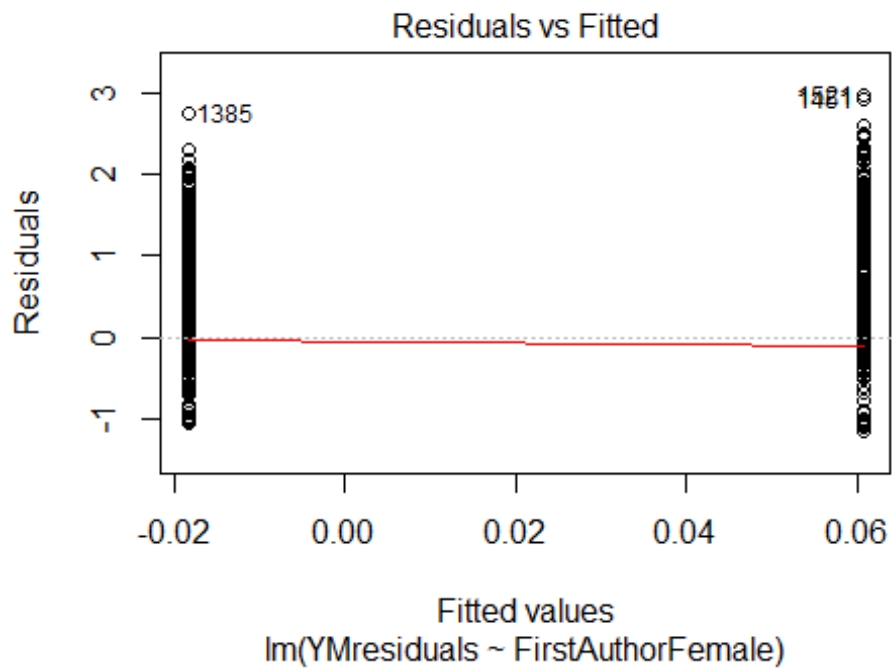

```
## [1] "List of  34 outliers with residuals above  2.5"
##         ScopusId  NLCS Year OneField Fields residuals
## 384   0031064803 3.798 1997     2907      1     3.231
## 405   0031155309 4.527 1997     2907      1     3.960
## 423   0031203501 3.645 1997     2907      1     3.078
## 467   0031309414 3.202 1997     2907      1     2.635
## 507   0031180540 3.068 1997     2906      2     2.501
## 802   0031990054 3.733 1998     2907      1     2.912
## 847   0032133322 3.868 1998     2907      1     3.313
## 1645  0034169785 3.538 2000     2907      1     2.707
## 1658  0034200397 4.716 2000     2907      1     3.620
## 1673  0034239103 3.484 2000     2907      1     2.653
## 1686  0034288015 3.639 2000     2907      1     2.808
## 1954  0035314990 3.542 2001     2907      1     3.028
## 2266  0036008249 4.970 2002     2907      1     4.361
## 2277  0036012829 4.003 2002     2907      1     3.394
## 2278  0036012830 3.956 2002     2907      1     3.347
## 2306  0036776261 4.988 2002     2907      1     4.379
## 2311  0036923166 3.278 2002     2907      1     2.669
## 2454 85026135285 3.643 2002     2906      2     3.034
## 2457 85026139222 3.255 2002     2906      2     2.646
## 2461 85026149648 3.558 2002     2906      2     2.949
## 2760  0042829280 3.196 2003     2705      3     2.744
## 2795  0037398605 3.515 2003     2907      1     3.063
## 2796  0037398606 3.810 2003     2907      1     3.358
## 2830  0347588237 3.366 2003     2907      1     2.914
## 2942 85026140202 3.147 2003     2906      2     2.695
```

```
## 3391  1342267540 4.111 2004     2907      1     3.444
## 3397  1342331383 3.157 2004     2907      1     2.756
## 3402  1942517874 3.396 2004     2907      1     2.641
## 3449  3042563662 2.965 2004     2906      2     2.564
## 4181 33646805797 4.136 2006     2907      1     3.262
## 4243 31344447300 3.172 2006     2907      1     2.563
## 5493 61649109012 3.256 2009     2907      1     2.562
## 5822 77956649584 4.174 2010     2907      1     3.644
## 5904 77954533425 3.488 2010     2907      1     2.869
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.1707 -0.5964 -0.0745  0.6447  4.3702
##
## Coefficients:
##                   Estimate Std. Error t value Pr(>|t|)
## (Intercept)        0.91812    0.07130   12.88  < 2e-16 ***
## LastAuthorFemale1 -0.32866    0.02636  -12.47  < 2e-16 ***
## Year1997          -0.01284    0.09530   -0.13  0.89285
## Year1998          -0.02811    0.09359   -0.30  0.76390
## Year1999          -0.30748    0.09051   -3.40  0.00069 ***
## Year2000           0.25260    0.11209    2.25  0.02428 *
## Year2001          -0.06349    0.09339   -0.68  0.49663
## Year2002           0.02833    0.08921    0.32  0.75083
## Year2003          -0.12710    0.08577   -1.48  0.13843
## Year2004          -0.17431    0.08010   -2.18  0.02960 *
## Year2005           0.00697    0.08255    0.08  0.93273
## Year2006           0.02812    0.08477    0.33  0.74009
## Year2007          -0.06895    0.08420   -0.82  0.41291
## Year2008          -0.02608    0.08233   -0.32  0.75147
## Year2009           0.02610    0.07998    0.33  0.74418
## Year2010          -0.04950    0.07897   -0.63  0.53085
## Year2011           0.06523    0.08540    0.76  0.44501
## Year2012           0.18412    0.08171    2.25  0.02428 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.741
## Multiple R-squared:  0.0562, Adjusted R-squared:  0.0527
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
##  5 observations c(185,650,973,1007,3469)
##   are outliers with |weight| = 0 ( < 2.2e-05);
##  244 weights are ~= 1. The remaining 4364 ones are summarized as
```

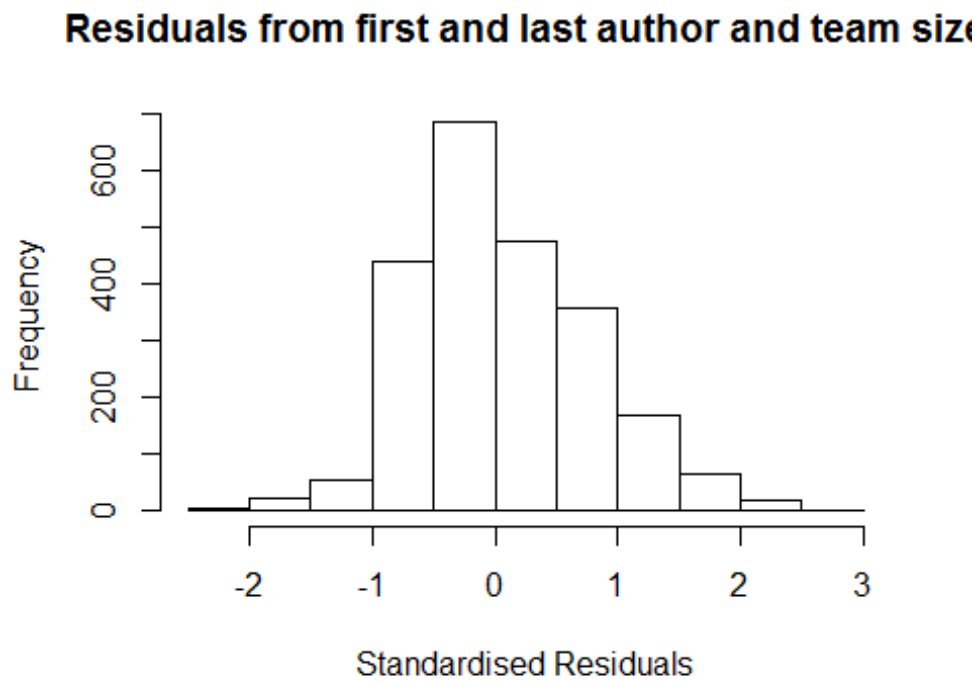
```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0024  0.8690  0.9380  0.8930  0.9700  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          2.17e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Sample size for the above analysis:  4613"
## [1] ""
## [1] ""
## [1] "####################################"
## [1] "Analysis of AJSC 2908"
## [1] "####################################"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##   14   50  102   99   94   90  101   89   95  101  104  144  283  407  446
## 2011 2012
##  441  417
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    0    0   26   31   40   29   92   82   86   91   99  122  253  364  377
## 2011 2012
##  374  363
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    0    0   26   30   40   28   88   77   84   85   97  113  239  328  350
## 2011 2012
##  353  344
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
##  Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 55, df = 14, p-value = 8e-07
```




```
##
##  Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 68, df = 1, p-value <2e-16
```

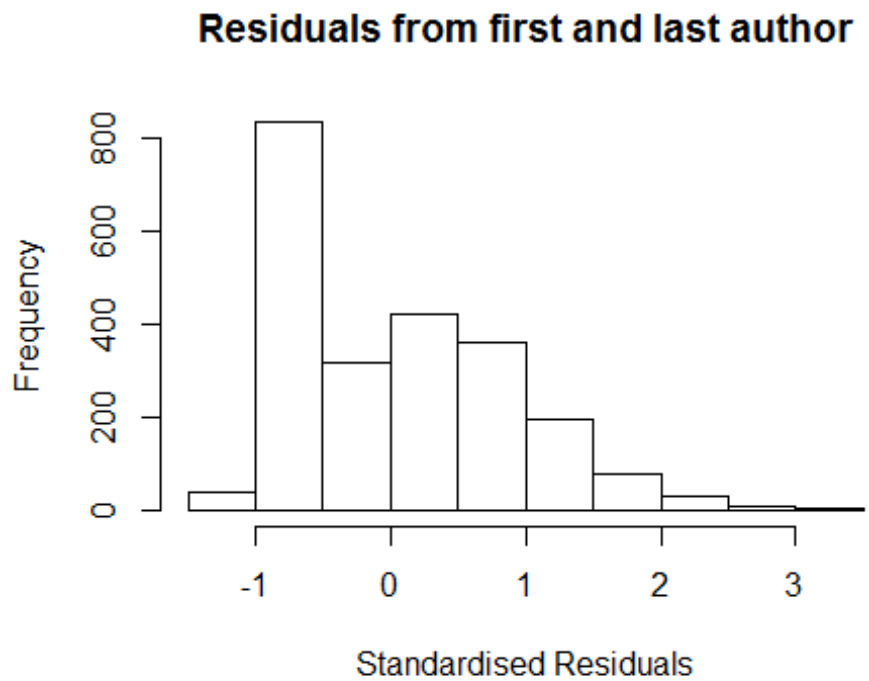



```
## [1] "Female first author team size 2018 geometric mean: 2.5962612267861"
## [1] "Male first author team size 2018 geometric mean: 3.97626502298039"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 3500, p-value = 3e-04
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 2.42770512130897"
## [1] "Male last author team size 2018 geometric mean: 4.14641275261914"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 4000, p-value = 7e-08
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.572  1           1.254
## LastAuthorFemale  1.748  1           1.322
## UniqueAuthors     1.457  4           1.048
## Year              1.470 14           1.014
```



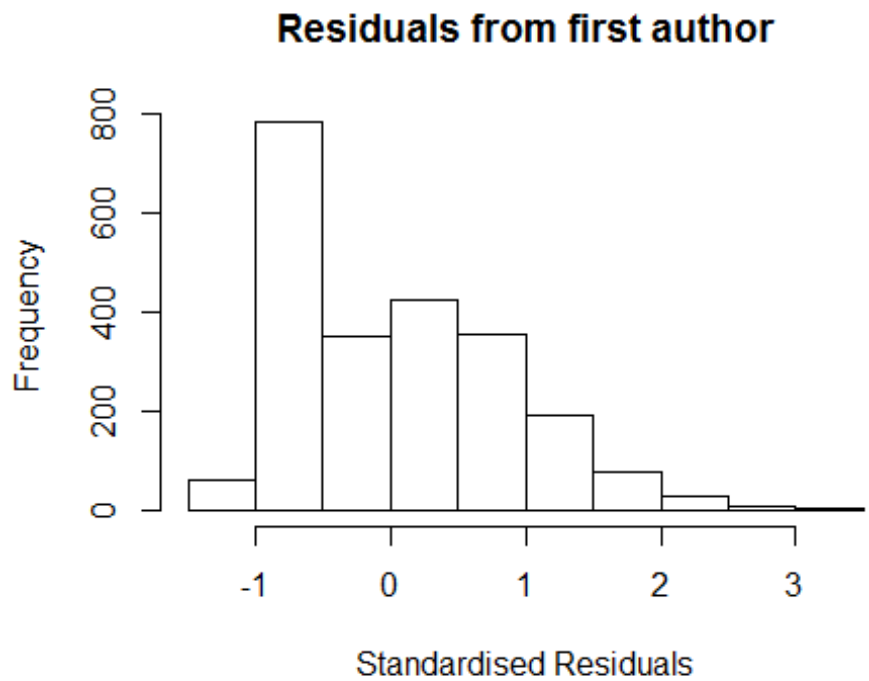

```
## [1] "List of  1 outliers with residuals above  2.5"
##         ScopusId  NLCS Year OneField Fields residuals
## 1947 77956631044 3.365 2010     2719      5     2.748
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
UniqueAuthors +
##     Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##     k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -2.0951 -0.4640 -0.0607  0.5450  2.7481
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)          0.53190    0.17174    3.10    0.002 **
## FirstAuthorFemale1   0.09665    0.04494    2.15    0.032 *
## LastAuthorFemale1    0.05317    0.04424    1.20    0.230
## UniqueAuthors2       0.03717    0.03593    1.03    0.301
## UniqueAuthors3       0.30274    0.05700    5.31  1.2e-07 ***
## UniqueAuthors4       0.59576    0.07761    7.68  2.4e-14 ***
## UniqueAuthors5       1.16381    0.07253   16.05  < 2e-16 ***
## Year1999             0.10455    0.21266    0.49    0.623
## Year2000             0.08263    0.23799    0.35    0.728
## Year2001             0.03501    0.23577    0.15    0.882
```

```
## Year2002            0.17726    0.17938    0.99    0.323
## Year2003            0.15902    0.18910    0.84    0.400
## Year2004            0.12878    0.18994    0.68    0.498
## Year2005            0.24956    0.17977    1.39    0.165
## Year2006           -0.00489    0.19957   -0.02    0.980
## Year2007            0.16167    0.18846    0.86    0.391
## Year2008           -0.08258    0.17529   -0.47    0.638
## Year2009           -0.26628    0.17182   -1.55    0.121
## Year2010           -0.21771    0.17075   -1.28    0.202
## Year2011           -0.22650    0.17051   -1.33    0.184
## Year2012           -0.12755    0.17085   -0.75    0.455
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.646
## Multiple R-squared:  0.239,  Adjusted R-squared:  0.233
## Convergence in 16 IRWLS iterations
##
## Robustness weights:
##  134 weights are ~= 1. The remaining 2148 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0309  0.8680  0.9430  0.8890  0.9720  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          4.38e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.520  1           1.233
## LastAuthorFemale  1.529  1           1.236
## Year              1.094 14           1.003
```



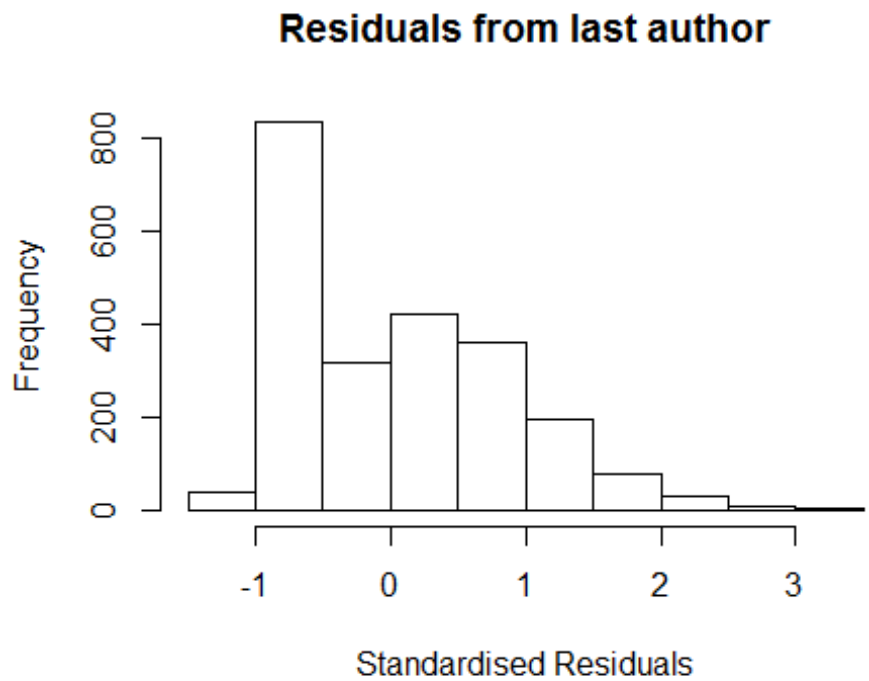

```
## [1] "List of  12 outliers with residuals above  2.5"
##         ScopusId  NLCS Year OneField Fields residuals
## 898  33947510199 3.556 2006     2719      5     2.711
## 1385 40249115349 3.558 2008     2719      5     2.675
## 1419 73849118275 3.246 2009     2719      5     2.852
## 1451 67449130309 3.706 2009     2719      5     3.312
## 1469 70449724751 3.106 2009     2719      5     2.588
## 1518 67649460937 3.078 2009     2719      5     2.560
## 1521 67649482684 3.748 2009     2719      5     3.230
## 1582 66049128443 3.049 2009     2719      5     2.655
## 1584 66049138872 2.930 2009     2719      5     2.536
## 1711 62849111793 3.089 2009     2908      4     2.695
## 1947 77956631044 3.365 2010     2719      5     2.763
## 2561 79952469323 3.228 2011     2719      5     2.596
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.2011 -0.6257 -0.0347  0.6214  3.3122
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)          0.88469    0.15988    5.53  3.5e-08 ***
```

```
## FirstAuthorFemale1  0.12207    0.06132    1.99    0.047 *
## LastAuthorFemale1  -0.12414    0.05629   -2.21    0.028 *
## Year1999           -0.00572    0.21179   -0.03    0.978
## Year2000            0.03095    0.20349    0.15    0.879
## Year2001           -0.10354    0.20828   -0.50    0.619
## Year2002            0.07016    0.16673    0.42    0.674
## Year2003            0.11541    0.17418    0.66    0.508
## Year2004            0.11675    0.17198    0.68    0.497
## Year2005            0.16096    0.16663    0.97    0.334
## Year2006           -0.04005    0.19546   -0.20    0.838
## Year2007            0.19433    0.17936    1.08    0.279
## Year2008           -0.12359    0.16495   -0.75    0.454
## Year2009           -0.36680    0.15916   -2.30    0.021 *
## Year2010           -0.28298    0.15862   -1.78    0.075 .
## Year2011           -0.25310    0.15891   -1.59    0.111
## Year2012           -0.08438    0.16052   -0.53    0.599
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.725
## Multiple R-squared:  0.049,  Adjusted R-squared:  0.0423
## Convergence in 15 IRWLS iterations
##
## Robustness weights:
##  129 weights are ~= 1. The remaining 2153 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0025  0.8810  0.9330  0.8930  0.9650  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          4.38e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.047  1           1.023
## Year              1.047 14           1.002
```




```
## [1] "List of  12 outliers with residuals above  2.5"
##         ScopusId  NLCS Year OneField Fields residuals
## 898  33947510199 3.556 2006     2719      5     2.711
## 1385 40249115349 3.558 2008     2719      5     2.675
## 1419 73849118275 3.246 2009     2719      5     2.852
## 1451 67449130309 3.706 2009     2719      5     3.312
## 1469 70449724751 3.106 2009     2719      5     2.588
## 1518 67649460937 3.078 2009     2719      5     2.560
## 1521 67649482684 3.748 2009     2719      5     3.230
## 1582 66049128443 3.049 2009     2719      5     2.655
## 1584 66049138872 2.930 2009     2719      5     2.536
## 1711 62849111793 3.089 2009     2908      4     2.695
## 1947 77956631044 3.365 2010     2719      5     2.763
## 2561 79952469323 3.228 2011     2719      5     2.596
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.1004 -0.6214 -0.0499  0.6186  3.2518
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)          0.8501     0.1608    5.29  1.4e-07 ***
```

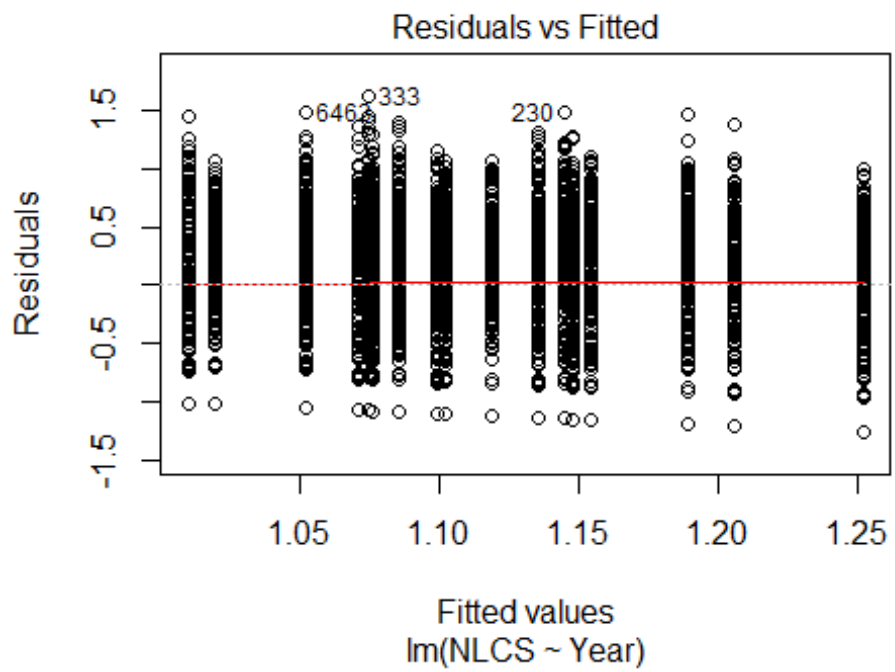
```
## FirstAuthorFemale1   0.0437     0.0512    0.85    0.394
## Year1999            -0.0132     0.2134   -0.06    0.951
## Year2000             0.0250     0.2058    0.12    0.903
## Year2001            -0.1022     0.2097   -0.49    0.626
## Year2002             0.0647     0.1683    0.38    0.701
## Year2003             0.1064     0.1754    0.61    0.544
## Year2004             0.1122     0.1736    0.65    0.518
## Year2005             0.1558     0.1683    0.93    0.355
## Year2006            -0.0323     0.1969   -0.16    0.870
## Year2007             0.2066     0.1805    1.14    0.252
## Year2008            -0.1179     0.1665   -0.71    0.479
## Year2009            -0.3539     0.1608   -2.20    0.028 *
## Year2010            -0.2724     0.1602   -1.70    0.089 .
## Year2011            -0.2425     0.1606   -1.51    0.131
## Year2012            -0.0845     0.1624   -0.52    0.603
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.734
## Multiple R-squared:  0.0451, Adjusted R-squared:  0.0388
## Convergence in 15 IRWLS iterations
##
## Robustness weights:
##  129 weights are ~= 1. The remaining 2153 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.011   0.882   0.936   0.895   0.965   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          4.38e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##                   GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.071  1           1.035
## Year             1.071 14           1.002
```




```
## [1] "List of  12 outliers with residuals above  2.5"
##         ScopusId  NLCS Year OneField Fields residuals
## 898  33947510199 3.556 2006     2719      5     2.711
## 1385 40249115349 3.558 2008     2719      5     2.675
## 1419 73849118275 3.246 2009     2719      5     2.852
## 1451 67449130309 3.706 2009     2719      5     3.312
## 1469 70449724751 3.106 2009     2719      5     2.588
## 1518 67649460937 3.078 2009     2719      5     2.560
## 1521 67649482684 3.748 2009     2719      5     3.230
## 1582 66049128443 3.049 2009     2719      5     2.655
## 1584 66049138872 2.930 2009     2719      5     2.536
## 1711 62849111793 3.089 2009     2908      4     2.695
## 1947 77956631044 3.365 2010     2719      5     2.763
## 2561 79952469323 3.228 2011     2719      5     2.596
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.1215 -0.6217 -0.0476  0.6243  3.1977
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         0.93804    0.15764    5.95  3.1e-09 ***
```

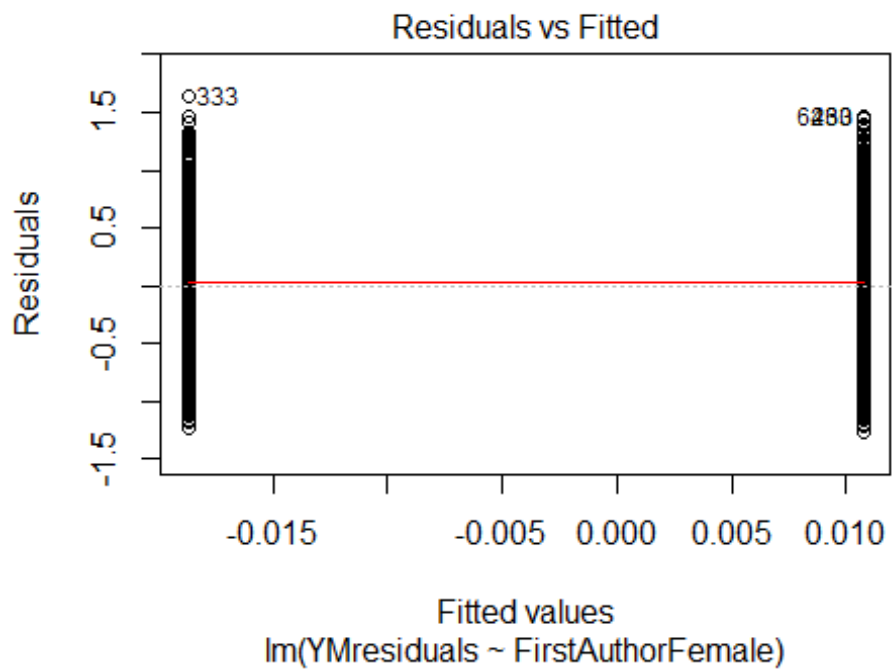
```
## LastAuthorFemale1 -0.06080    0.04755   -1.28    0.201
## Year1999          -0.00221    0.21130   -0.01    0.992
## Year2000           0.02381    0.20310    0.12    0.907
## Year2001          -0.09195    0.20843   -0.44    0.659
## Year2002           0.07413    0.16706    0.44    0.657
## Year2003           0.11902    0.17420    0.68    0.495
## Year2004           0.12099    0.17233    0.70    0.483
## Year2005           0.16565    0.16687    0.99    0.321
## Year2006          -0.03381    0.19539   -0.17    0.863
## Year2007           0.18342    0.17916    1.02    0.306
## Year2008          -0.12895    0.16510   -0.78    0.435
## Year2009          -0.36892    0.15943   -2.31    0.021 *
## Year2010          -0.28906    0.15876   -1.82    0.069 .
## Year2011          -0.25555    0.15909   -1.61    0.108
## Year2012          -0.09211    0.16084   -0.57    0.567
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.726
## Multiple R-squared:  0.0459, Adjusted R-squared:  0.0396
## Convergence in 15 IRWLS iterations
##
## Robustness weights:
##  117 weights are ~= 1. The remaining 2165 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0135  0.8800  0.9340  0.8940  0.9650  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          4.38e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Sample size for the above analysis:  2282"
## [1] ""
## [1] ""
## [1] "####################################"
## [1] "Analysis of AJSC 2909"
## [1] "####################################"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
```

```
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  269  292  256  306  359  440  407  309  309  361  373  386  383  454  512
## 2011 2012
##  420  461
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  212  203  166  170  190  195  354  276  255  305  307  330  323  402  432
## 2011 2012
##  352  386
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  195  191  153  153  178  182  320  249  219  270  283  295  299  376  399
## 2011 2012
##  322  354
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
##  Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 45, df = 16, p-value = 1e-04
```



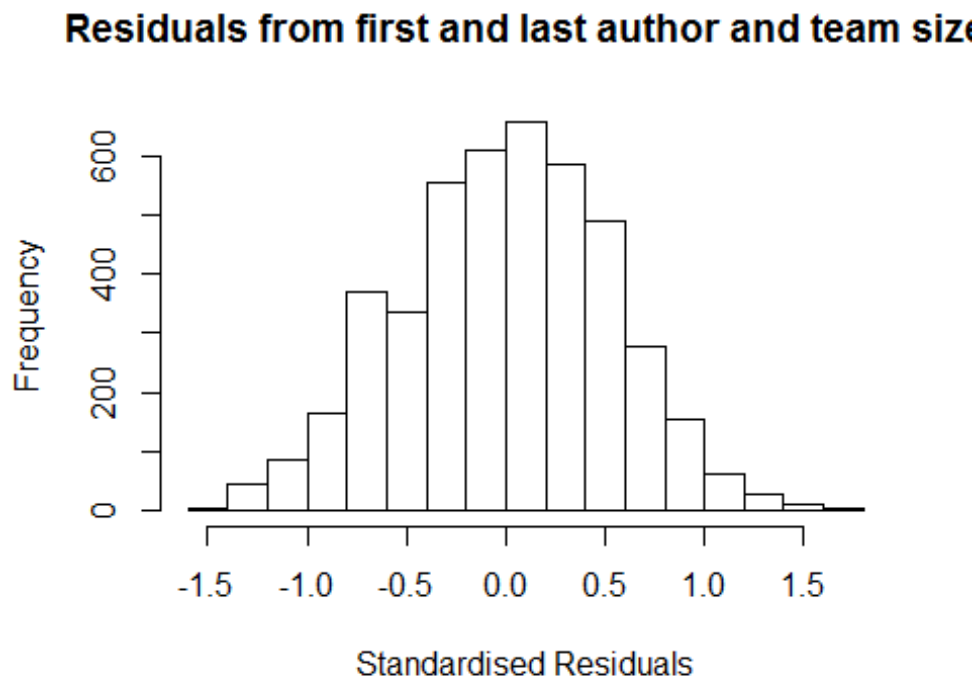

```
##
##  Bartlett test of homogeneity of variances
##
```

```
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 13, df = 1, p-value = 4e-04
```



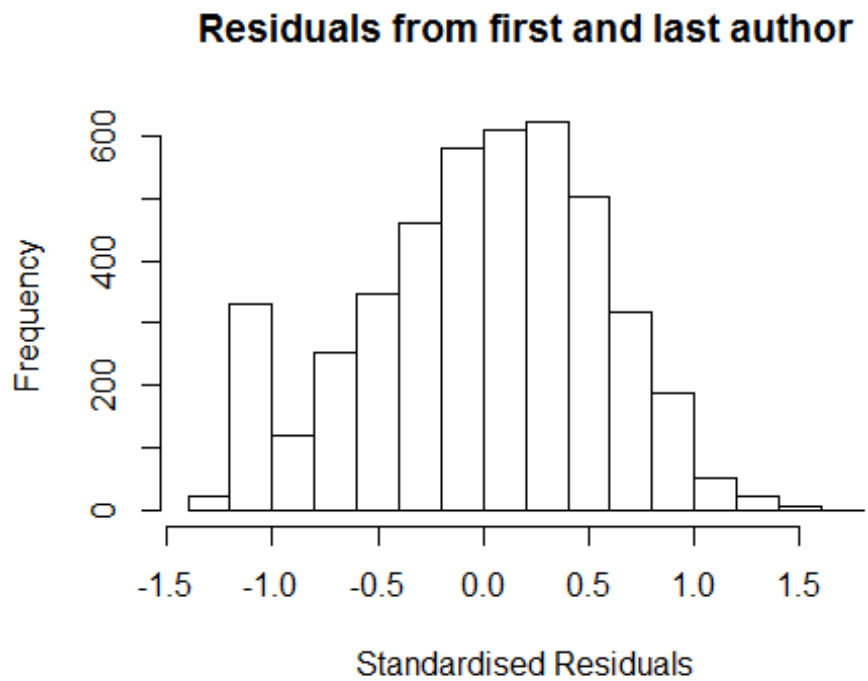

```
## [1] "Female first author team size 2018 geometric mean: 3.46491887012361"
## [1] "Male first author team size 2018 geometric mean: 2.70817103475336"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 22000, p-value = 0.001
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 3.3345554319921"
## [1] "Male last author team size 2018 geometric mean: 3.03561560533904"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 22000, p-value = 0.2
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.071  1           1.035
## LastAuthorFemale  1.091  1           1.045
```

```
## UniqueAuthors     1.157  4           1.018
## Year              1.163 16           1.005
```



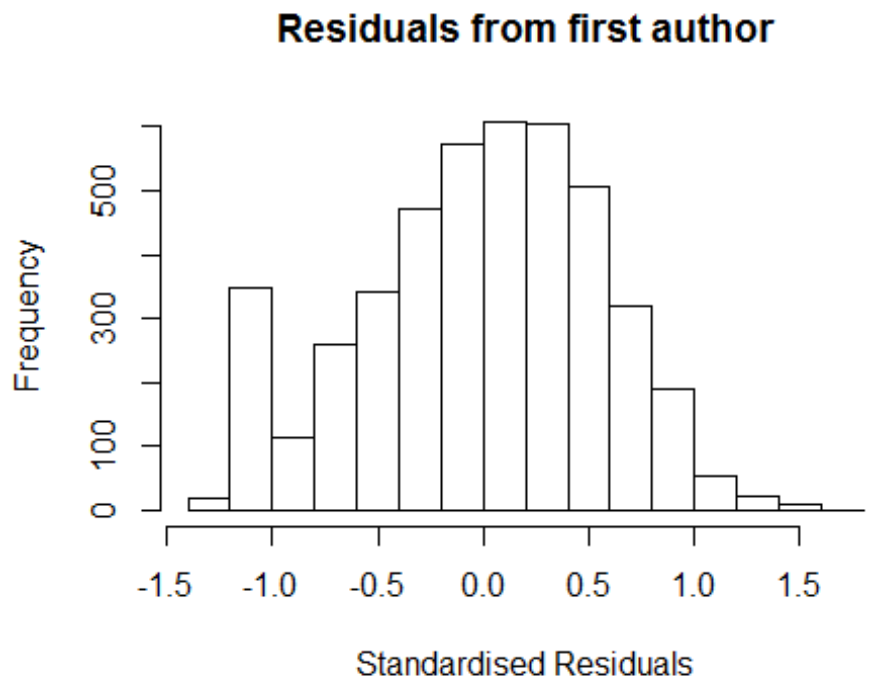

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
UniqueAuthors +
##     Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##     k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.4785 -0.3585  0.0196  0.3694  1.6839
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)          0.85155    0.04819   17.67  < 2e-16 ***
## FirstAuthorFemale1   0.04795    0.01799    2.67  0.00771 **
## LastAuthorFemale1   -0.03029    0.01718   -1.76  0.07786 .
## UniqueAuthors2       0.31098    0.02701   11.51  < 2e-16 ***
## UniqueAuthors3       0.42499    0.02761   15.39  < 2e-16 ***
## UniqueAuthors4       0.50177    0.02952   17.00  < 2e-16 ***
## UniqueAuthors5       0.58488    0.02569   22.77  < 2e-16 ***
```

```
## Year1997           -0.07107    0.06755   -1.05  0.29277
## Year1998           -0.06903    0.06555   -1.05  0.29238
## Year1999            0.03573    0.06407    0.56  0.57713
## Year2000           -0.10273    0.06173   -1.66  0.09615 .
## Year2001           -0.14817    0.06183   -2.40  0.01661 *
## Year2002           -0.09765    0.05439   -1.80  0.07263 .
## Year2003           -0.00585    0.05689   -0.10  0.91816
## Year2004            0.02795    0.05785    0.48  0.62899
## Year2005           -0.05080    0.05589   -0.91  0.36341
## Year2006           -0.08347    0.05644   -1.48  0.13922
## Year2007           -0.14315    0.05347   -2.68  0.00746 **
## Year2008           -0.13635    0.05296   -2.57  0.01006 *
## Year2009           -0.07696    0.05212   -1.48  0.13986
## Year2010           -0.15682    0.05202   -3.01  0.00259 **
## Year2011           -0.21206    0.05363   -3.95  7.8e-05 ***
## Year2012           -0.19697    0.05391   -3.65  0.00026 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.539
## Multiple R-squared:  0.15,   Adjusted R-squared:  0.146
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  344 weights are ~= 1. The remaining 4094 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.309   0.870   0.951   0.912   0.985   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          2.25e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.051  1           1.025
## LastAuthorFemale  1.057  1           1.028
## Year              1.047 16           1.001
```



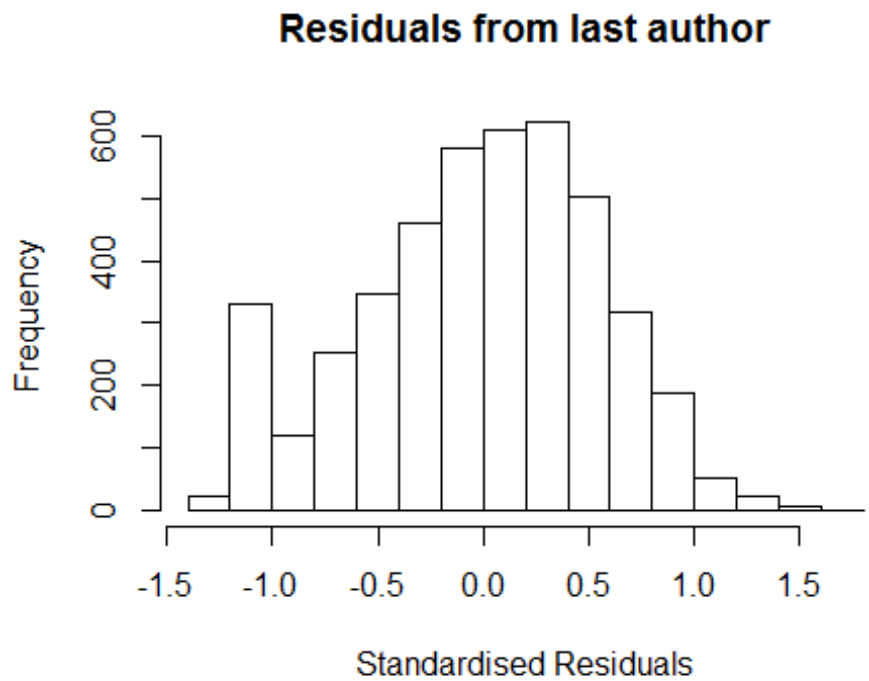

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.3205 -0.3817  0.0306  0.3909  1.6199
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)          1.16223    0.05375   21.62  < 2e-16 ***
## FirstAuthorFemale1  0.05746    0.01958    2.94   0.0034 **
## LastAuthorFemale1  -0.08400    0.01823   -4.61  4.2e-06 ***
## Year1997           -0.08911    0.07529   -1.18   0.2367
## Year1998           -0.03715    0.07429   -0.50   0.6170
## Year1999            0.08133    0.07077    1.15   0.2506
## Year2000           -0.08784    0.07073   -1.24   0.2143
## Year2001           -0.15553    0.07080   -2.20   0.0281 *
## Year2002           -0.06575    0.06264   -1.05   0.2940
## Year2003            0.03623    0.06487    0.56   0.5765
## Year2004            0.10083    0.06534    1.54   0.1229
## Year2005           -0.00584    0.06526   -0.09   0.9287
```

```
## Year2006           -0.00751    0.06362   -0.12   0.9060
## Year2007           -0.03855    0.06183   -0.62   0.5330
## Year2008           -0.03146    0.06049   -0.52   0.6030
## Year2009            0.02223    0.05980    0.37   0.7101
## Year2010           -0.08083    0.05960   -1.36   0.1751
## Year2011           -0.11181    0.06069   -1.84   0.0655 .
## Year2012           -0.10114    0.06244   -1.62   0.1054
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.566
## Multiple R-squared:  0.0169, Adjusted R-squared:  0.0129
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
##  346 weights are ~= 1. The remaining 4092 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.393   0.869   0.949   0.909   0.985   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          2.25e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.021  1           1.010
## Year              1.021 16           1.001
```




```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.2619 -0.3821  0.0333  0.3944  1.6468
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)          1.13453    0.05293   21.44   <2e-16 ***
## FirstAuthorFemale1  0.03080    0.01936    1.59    0.112
## Year1997           -0.08834    0.07529   -1.17    0.241
## Year1998           -0.03891    0.07455   -0.52    0.602
## Year1999            0.08198    0.07066    1.16    0.246
## Year2000           -0.09527    0.07028   -1.36    0.175
## Year2001           -0.15916    0.07074   -2.25    0.025 *
## Year2002           -0.07413    0.06229   -1.19    0.234
## Year2003            0.02915    0.06436    0.45    0.651
## Year2004            0.09658    0.06505    1.48    0.138
## Year2005           -0.00746    0.06506   -0.11    0.909
## Year2006           -0.00786    0.06339   -0.12    0.901
```

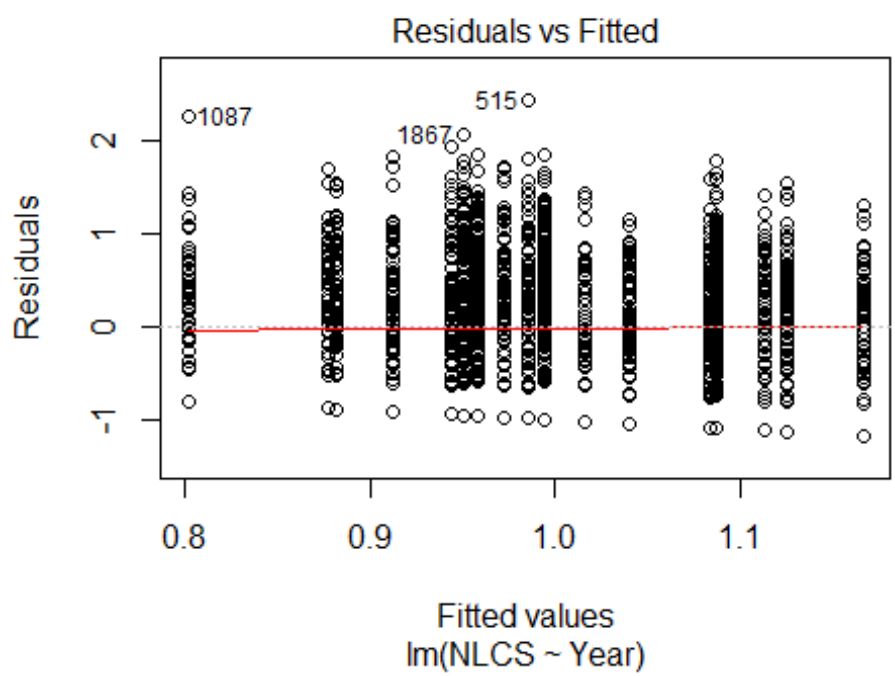
```
## Year2007           -0.03741    0.06147   -0.61    0.543
## Year2008           -0.03321    0.06027   -0.55    0.582
## Year2009            0.01835    0.05939    0.31    0.757
## Year2010           -0.08306    0.05934   -1.40    0.162
## Year2011           -0.11720    0.06045   -1.94    0.053 .
## Year2012           -0.10219    0.06221   -1.64    0.101
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.567
## Multiple R-squared:  0.0124, Adjusted R-squared:  0.00861
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
##  361 weights are ~= 1. The remaining 4077 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.379   0.870   0.950   0.908   0.984   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          2.25e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##                   GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.026  1           1.013
## Year             1.026 16           1.001
```



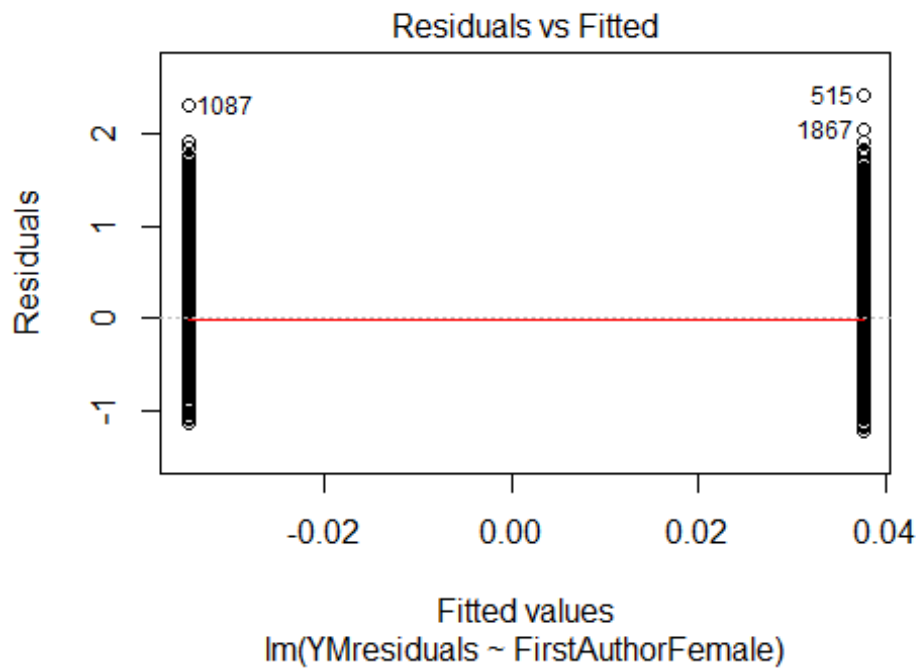

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.2924 -0.3851  0.0302  0.3926  1.5926
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.183200   0.052937   22.35  < 2e-16 ***
## LastAuthorFemale1 -0.067243   0.018197   -3.70  0.00022 ***
## Year1997          -0.082822   0.074903   -1.11  0.26890
## Year1998          -0.032229   0.074024   -0.44  0.66330
## Year1999           0.086440   0.070549    1.23  0.22054
## Year2000          -0.084968   0.070436   -1.21  0.22776
## Year2001          -0.146151   0.070547   -2.07  0.03835 *
## Year2002          -0.060145   0.062376   -0.96  0.33498
## Year2003           0.043249   0.064611    0.67  0.50329
## Year2004           0.109172   0.064847    1.68  0.09234 .
## Year2005           0.000521   0.065055    0.01  0.99361
## Year2006          -0.001100   0.063373   -0.02  0.98615
```

```
## Year2007          -0.033585   0.061640   -0.54  0.58588
## Year2008          -0.024782   0.060206   -0.41  0.68063
## Year2009           0.028354   0.059393    0.48  0.63311
## Year2010          -0.069166   0.059078   -1.17  0.24176
## Year2011          -0.102923   0.060283   -1.71  0.08783 .
## Year2012          -0.091179   0.062041   -1.47  0.14173
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.566
## Multiple R-squared:  0.0151, Adjusted R-squared:  0.0113
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  351 weights are ~= 1. The remaining 4087 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.408   0.870   0.950   0.908   0.984   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          2.25e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Sample size for the above analysis:  4438"
## [1] ""
## [1] ""
## [1] "####################################"
## [1] "Analysis of AJSC 2910"
## [1] "####################################"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  112   87   97  124  143  194  138  135  121  135  108  126  177  381  487
## 2011 2012
##  491  484
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##   75   70   79   97  109  151  123  117  104  112   92  111  148  329  436
## 2011 2012
```

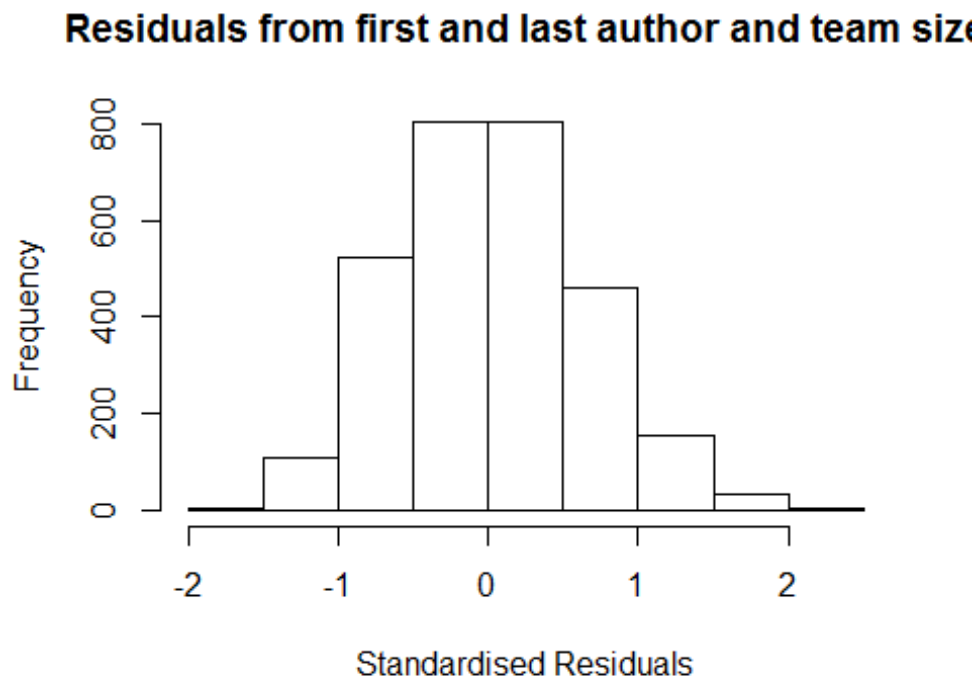
```
##  435  433
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##   75   68   76   95  106  144  120  113   99  105   84  103  138  314  420
## 2011 2012
##  421  416
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
##  Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 24, df = 16, p-value = 0.09
```




```
##
##  Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 6.1, df = 1, p-value = 0.01
```

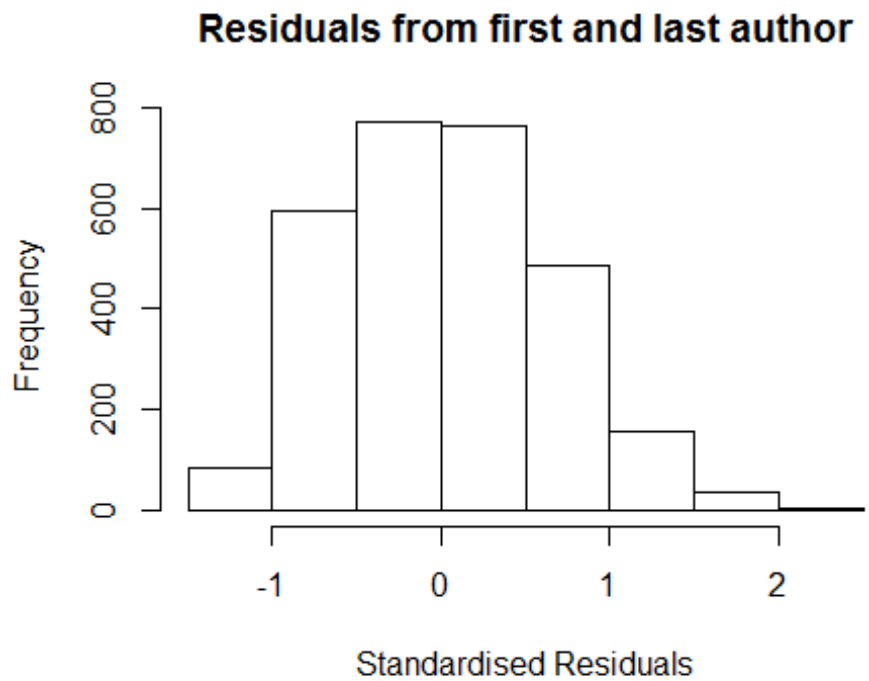



```
## [1] "Female first author team size 2018 geometric mean: 1.85791926001733"
## [1] "Male first author team size 2018 geometric mean: 1.49957970865817"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 11000, p-value = 0.005
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 1.87953589511945"
## [1] "Male last author team size 2018 geometric mean: 1.49256693121686"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 11000, p-value = 0.007
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 2.648  1           1.627
## LastAuthorFemale  2.545  1           1.595
## UniqueAuthors     1.192  4           1.022
## Year              1.188 16           1.005
```



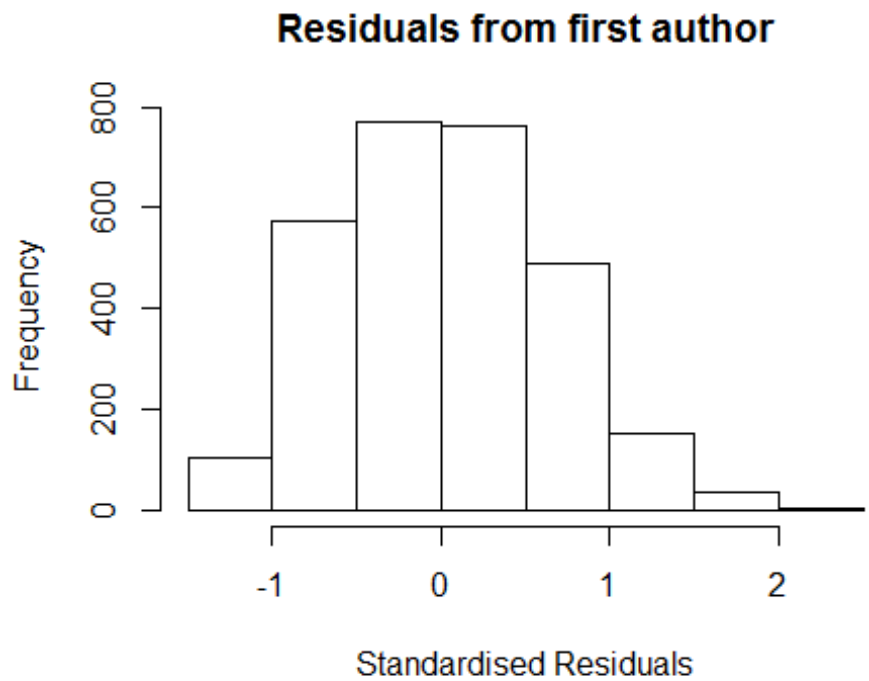

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
UniqueAuthors +
##     Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##     k.max = 1000))
##  \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.56803 -0.43930  0.00617  0.44695  2.44854
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         0.95211    0.07753   12.28  < 2e-16 ***
## FirstAuthorFemale1  0.03969    0.03994    0.99  0.32038
## LastAuthorFemale1  -0.00969    0.03950   -0.25  0.80613
## UniqueAuthors2      0.14910    0.03365    4.43  9.7e-06 ***
## UniqueAuthors3      0.27145    0.04783    5.68  1.5e-08 ***
## UniqueAuthors4      0.41014    0.07298    5.62  2.1e-08 ***
## UniqueAuthors5      0.56610    0.06559    8.63  < 2e-16 ***
## Year1997            0.16261    0.11495    1.41  0.15729
## Year1998            0.14894    0.10914    1.36  0.17243
## Year1999           -0.12890    0.10800   -1.19  0.23275
```

```
## Year2000           -0.09177    0.10937   -0.84  0.40149
## Year2001           -0.10256    0.09957   -1.03  0.30311
## Year2002           -0.13651    0.10108   -1.35  0.17695
## Year2003           -0.17180    0.09748   -1.76  0.07811 .
## Year2004           -0.33465    0.09983   -3.35  0.00081 ***
## Year2005           -0.17965    0.10448   -1.72  0.08564 .
## Year2006           -0.06162    0.10119   -0.61  0.54259
## Year2007            0.01899    0.10029    0.19  0.84986
## Year2008            0.01565    0.09302    0.17  0.86641
## Year2009           -0.09021    0.08476   -1.06  0.28733
## Year2010            0.01982    0.08358    0.24  0.81257
## Year2011           -0.11842    0.08219   -1.44  0.14976
## Year2012           -0.11603    0.08432   -1.38  0.16890
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.661
## Multiple R-squared:  0.0642, Adjusted R-squared:  0.057
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  247 weights are ~= 1. The remaining 2650 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.141   0.861   0.951   0.914   0.985   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          3.45e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 2.548  1           1.596
## LastAuthorFemale  2.550  1           1.597
## Year              1.072 16           1.002
```



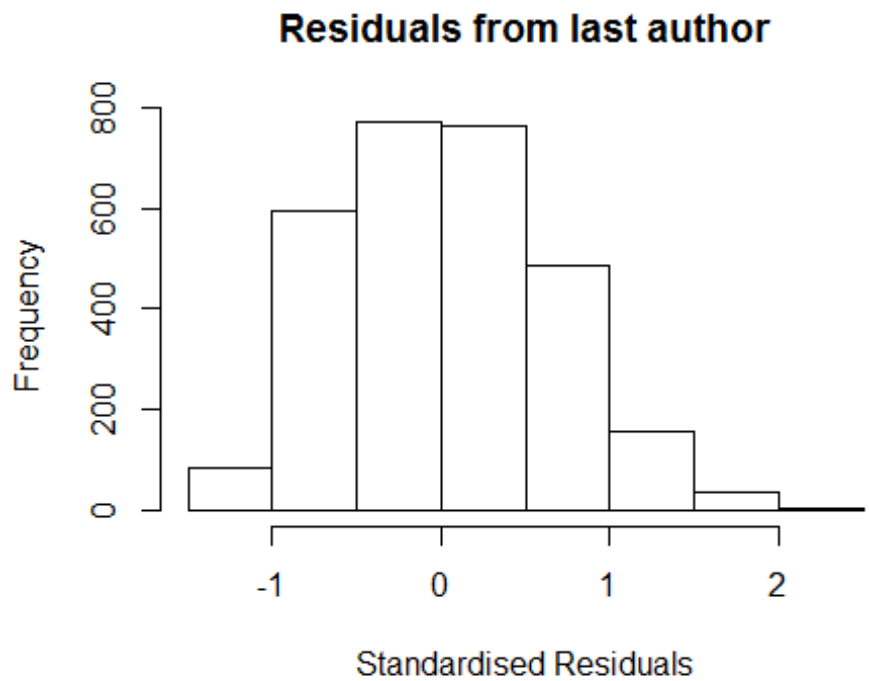

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.25245 -0.48845 -0.00308  0.46342  2.44765
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)           0.9820     0.0782   12.56   <2e-16 ***
## FirstAuthorFemale1    0.0929     0.0403    2.30   0.0213 *
## LastAuthorFemale1    -0.0208     0.0406   -0.51   0.6091
## Year1997              0.1984     0.1149    1.73   0.0844 .
## Year1998              0.1410     0.1088    1.30   0.1950
## Year1999             -0.1500     0.1084   -1.38   0.1667
## Year2000             -0.0727     0.1109   -0.66   0.5118
## Year2001             -0.0925     0.1012   -0.91   0.3609
## Year2002             -0.1089     0.1014   -1.07   0.2830
## Year2003             -0.1667     0.0993   -1.68   0.0933 .
## Year2004             -0.2851     0.1028   -2.77   0.0056 **
## Year2005             -0.1436     0.1056   -1.36   0.1741
```

```
## Year2006            -0.0110     0.1011   -0.11   0.9137
## Year2007             0.0856     0.1021    0.84   0.4017
## Year2008             0.0598     0.0946    0.63   0.5276
## Year2009            -0.0771     0.0860   -0.90   0.3699
## Year2010             0.0526     0.0843    0.62   0.5322
## Year2011            -0.0765     0.0831   -0.92   0.3575
## Year2012            -0.0548     0.0852   -0.64   0.5204
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.691
## Multiple R-squared:  0.0215, Adjusted R-squared:  0.0154
## Convergence in 10 IRWLS iterations
##
## Robustness weights:
##  246 weights are ~= 1. The remaining 2651 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.184   0.866   0.950   0.918   0.986   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          3.45e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.045  1           1.022
## Year              1.045 16           1.001
```




```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.25487 -0.48663 -0.00185  0.46155  2.44405
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)           0.9803     0.0779   12.58   <2e-16 ***
## FirstAuthorFemale1    0.0770     0.0258    2.98   0.0029 **
## Year1997              0.1976     0.1148    1.72   0.0855 .
## Year1998              0.1411     0.1087    1.30   0.1945
## Year1999             -0.1494     0.1083   -1.38   0.1677
## Year2000             -0.0724     0.1108   -0.65   0.5137
## Year2001             -0.0934     0.1012   -0.92   0.3559
## Year2002             -0.1099     0.1013   -1.08   0.2781
## Year2003             -0.1663     0.0992   -1.68   0.0938 .
## Year2004             -0.2861     0.1028   -2.78   0.0054 **
## Year2005             -0.1441     0.1056   -1.36   0.1725
## Year2006             -0.0118     0.1009   -0.12   0.9066
```

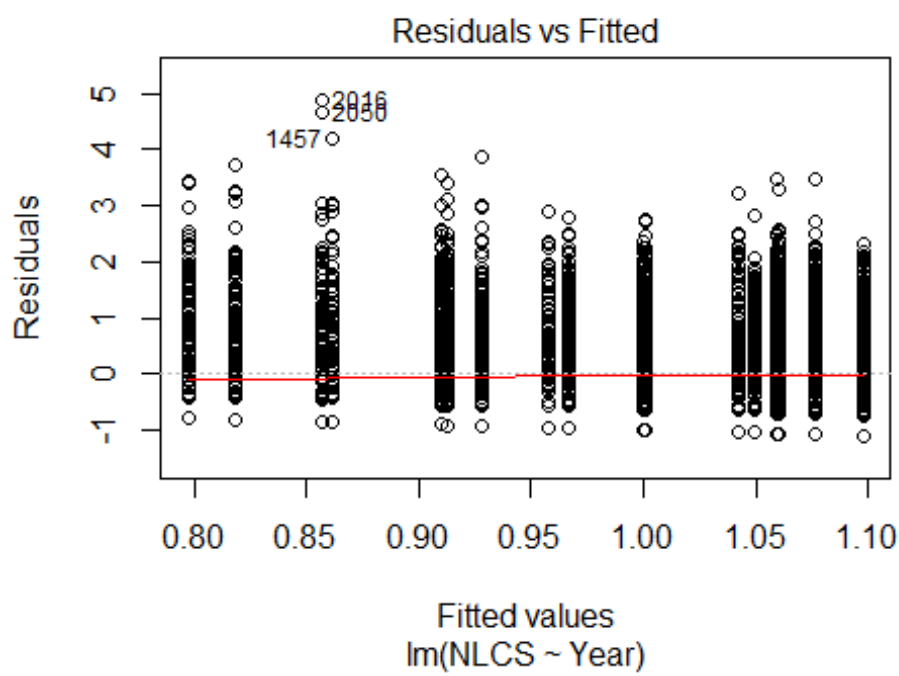
```
## Year2007             0.0866     0.1021    0.85   0.3966
## Year2008             0.0594     0.0945    0.63   0.5298
## Year2009            -0.0770     0.0859   -0.90   0.3702
## Year2010             0.0528     0.0841    0.63   0.5302
## Year2011            -0.0767     0.0830   -0.92   0.3560
## Year2012            -0.0549     0.0851   -0.64   0.5193
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.692
## Multiple R-squared:  0.0214, Adjusted R-squared:  0.0156
## Convergence in 10 IRWLS iterations
##
## Robustness weights:
##  260 weights are ~= 1. The remaining 2637 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.186   0.864   0.950   0.918   0.985   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          3.45e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##                   GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.047  1           1.023
## Year             1.047 16           1.001
```



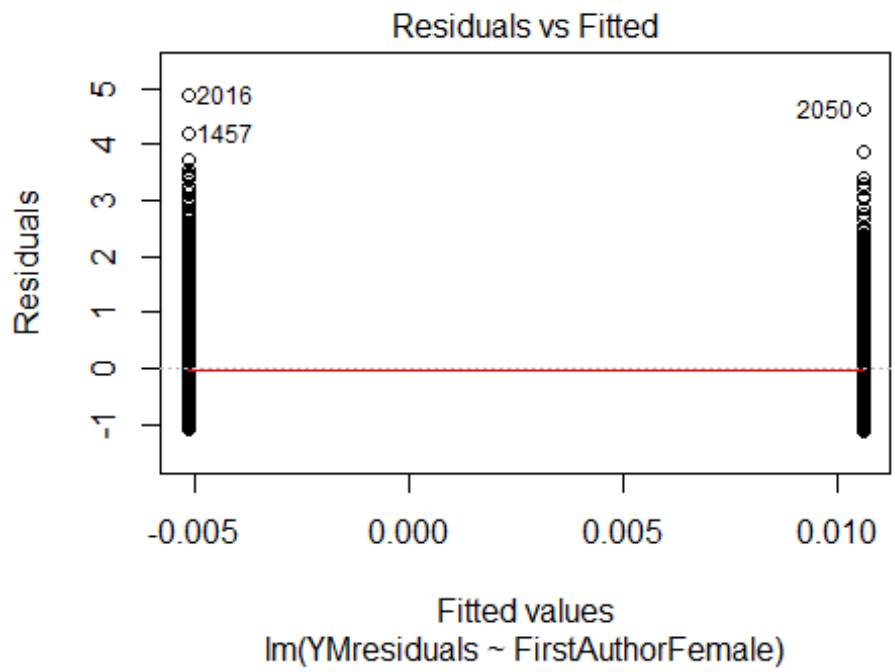

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.22866 -0.49892 -0.00464  0.46957  2.45128
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         0.98967    0.07888   12.55   <2e-16 ***
## LastAuthorFemale1   0.05021    0.02603    1.93   0.0538 .
## Year1997            0.18877    0.11549    1.63   0.1023
## Year1998            0.14219    0.10964    1.30   0.1948
## Year1999           -0.15031    0.10868   -1.38   0.1667
## Year2000           -0.06216    0.11114   -0.56   0.5760
## Year2001           -0.08849    0.10157   -0.87   0.3837
## Year2002           -0.10292    0.10185   -1.01   0.3123
## Year2003           -0.15760    0.09972   -1.58   0.1141
## Year2004           -0.27722    0.10333   -2.68   0.0073 **
## Year2005           -0.13409    0.10623   -1.26   0.2070
## Year2006           -0.00673    0.10125   -0.07   0.9470
```

```
## Year2007           0.09681    0.10344    0.94   0.3494
## Year2008           0.06357    0.09544    0.67   0.5054
## Year2009          -0.07394    0.08661   -0.85   0.3934
## Year2010           0.05781    0.08485    0.68   0.4957
## Year2011          -0.07375    0.08376   -0.88   0.3787
## Year2012          -0.05124    0.08585   -0.60   0.5507
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.694
## Multiple R-squared:  0.0195, Adjusted R-squared:  0.0137
## Convergence in 10 IRWLS iterations
##
## Robustness weights:
##  236 weights are ~= 1. The remaining 2661 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.186   0.866   0.951   0.919   0.986   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          3.45e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Sample size for the above analysis:  2897"
## [1] ""
## [1] ""
## [1] "####################################"
## [1] "Analysis of AJSC 2911"
## [1] "####################################"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  539  585  585  564  692  728  727  601  518  608  507  627  754  831  851
## 2011 2012
## 1046  726
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  256  231  249  299  392  427  666  529  465  541  446  548  673  739  737
## 2011 2012
```

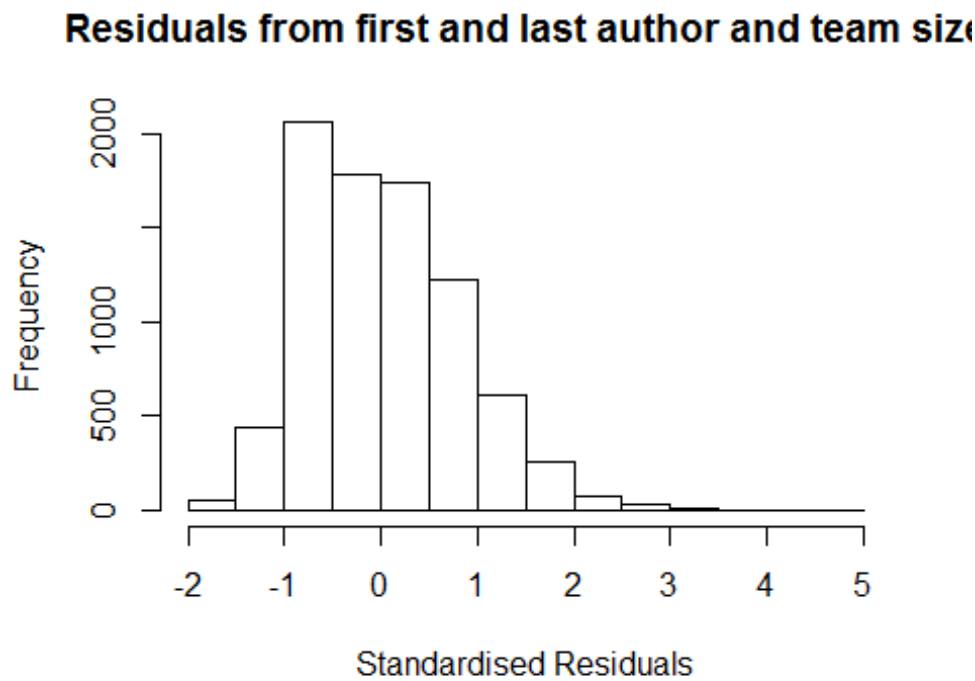
```
##  935  640
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  246  218  237  285  362  414  632  496  447  518  422  515  636  689  692
## 2011 2012
##  889  596
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
##  Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 62, df = 16, p-value = 3e-07
```




```
##
##  Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 0.71, df = 1, p-value = 0.4
```



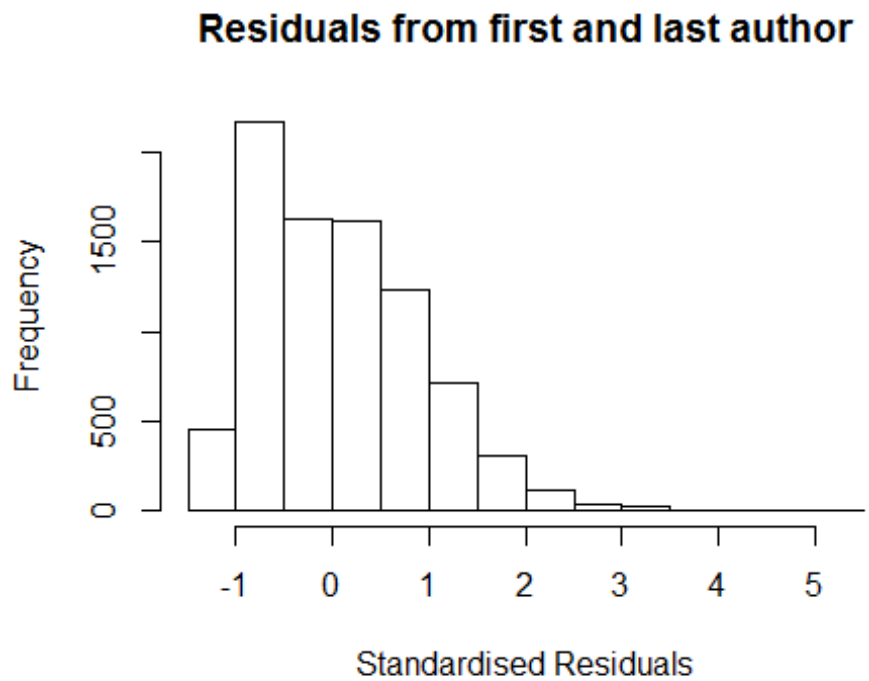

```
## [1] "Female first author team size 2018 geometric mean: 2.71610194296933"
## [1] "Male first author team size 2018 geometric mean: 3.27465578171249"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 26000, p-value = 0.007
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 2.57539720822308"
## [1] "Male last author team size 2018 geometric mean: 3.57219206481451"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 24000, p-value = 3e-07
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.930  1           1.389
## LastAuthorFemale  1.973  1           1.405
## UniqueAuthors     1.124  4           1.015
## Year              1.123 16           1.004
```




```
## [1] "List of  40 outliers with residuals above  2.5"
##          ScopusId  NLCS Year OneField Fields residuals
## 217    0030002054 3.857 1996     2911      1     2.857
## 648    0030767805 4.270 1997     2911      1     3.594
## 1008  17644439119 3.543 1997     2911      1     2.867
## 1348   0031968557 3.847 1998     2911      1     2.741
## 1350   0031969182 3.760 1998     2911      1     2.654
## 1351   0031978565 3.901 1998     2911      1     2.795
## 1457   0031852597 5.054 1998     2911      1     3.540
## 2016   0344572791 5.730 1999     2911      1     4.798
## 2160   0033233216 3.895 1999     2911      1     3.381
## 2186   0033497846 3.590 1999     2911      1     2.658
## 2189   0033503671 3.731 1999     2911      1     2.774
## 2728   0034154033 4.072 2000     2911      1     3.226
## 2760   0034294822 4.543 2000     2911      1     3.495
## 2762   0034302388 3.904 2000     2911      1     2.535
## 2767   0034330243 4.048 2000     2911      1     2.952
## 2980  84993710017 3.429 2000     2910      2     2.583
## 3190   0038417818 3.189 2001     2911      1     2.696
## 3334   0002087395 3.266 2001     2911      1     2.773
## 3372   0035151691 4.191 2001     2911      1     3.102
## 3427   0035744421 3.759 2001     2911      1     2.823
## 3428   0035744427 4.234 2001     2911      1     3.120
## 3523  33745269242 3.080 2001     2911      1     2.587
## 3964   0036853306 3.424 2002     2911      1     2.831
## 4132   0036310834 3.219 2002     2911      1     2.626
## 4240   0036718125 4.315 2002     2911      1     2.780
```

```
## 4859   0037613484 4.445 2003     2911      1     2.861
## 4860   0038288856 3.497 2003     2911      1     2.523
## 4947   0037607052 3.155 2003     2911      1     2.599
## 5394   8544254752 3.927 2004     2911      1     2.623
## 5611   4043063731 4.798 2004     2911      1     3.234
## 5615   4043160018 3.891 2004     2911      1     2.875
## 6105  20144381547 3.357 2005     2911      1     2.720
## 6150  23844542129 4.350 2005     2911      1     2.709
## 6880  33748693969 4.535 2006     2911      1     3.300
## 7274  34548240535 3.431 2007     2911      1     2.769
## 7814  57049157007 3.212 2008     2911      1     2.584
## 8098  44249112154 4.538 2008     2911      1     2.882
## 8316  42149106861 3.786 2008     2911      1     2.715
## 10922 79960941189 3.752 2011     2911      1     2.559
## 11696 84865733130 3.859 2012     2911      1     2.564
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
UniqueAuthors +
##     Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##     k.max = 1000))
##  \--> method = "MM"
## Residuals:
##    Min     1Q Median     3Q    Max
## -1.703 -0.563 -0.046  0.545  4.798
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         0.54436    0.05982    9.10   <2e-16 ***
## FirstAuthorFemale1  0.06196    0.02526    2.45    0.014 *
## LastAuthorFemale1  -0.02455    0.02506   -0.98    0.327
## UniqueAuthors2      0.41861    0.02414   17.34   <2e-16 ***
## UniqueAuthors3      0.59623    0.02902   20.55   <2e-16 ***
## UniqueAuthors4      0.70623    0.03450   20.47   <2e-16 ***
## UniqueAuthors5      1.00405    0.03141   31.96   <2e-16 ***
## Year1997            0.09445    0.08143    1.16    0.246
## Year1998           -0.07185    0.07841   -0.92    0.359
## Year1999           -0.06793    0.07920   -0.86    0.391
## Year2000           -0.15437    0.07313   -2.11    0.035 *
## Year2001           -0.08885    0.06885   -1.29    0.197
## Year2002            0.01117    0.06537    0.17    0.864
## Year2003           -0.02625    0.06801   -0.39    0.700
## Year2004            0.01585    0.06825    0.23    0.816
## Year2005            0.09262    0.06752    1.37    0.170
## Year2006            0.03211    0.06949    0.46    0.644
## Year2007            0.08042    0.06769    1.19    0.235
## Year2008            0.04577    0.06564    0.70    0.486
## Year2009           -0.00697    0.06534   -0.11    0.915
## Year2010           -0.00881    0.06392   -0.14    0.890
```

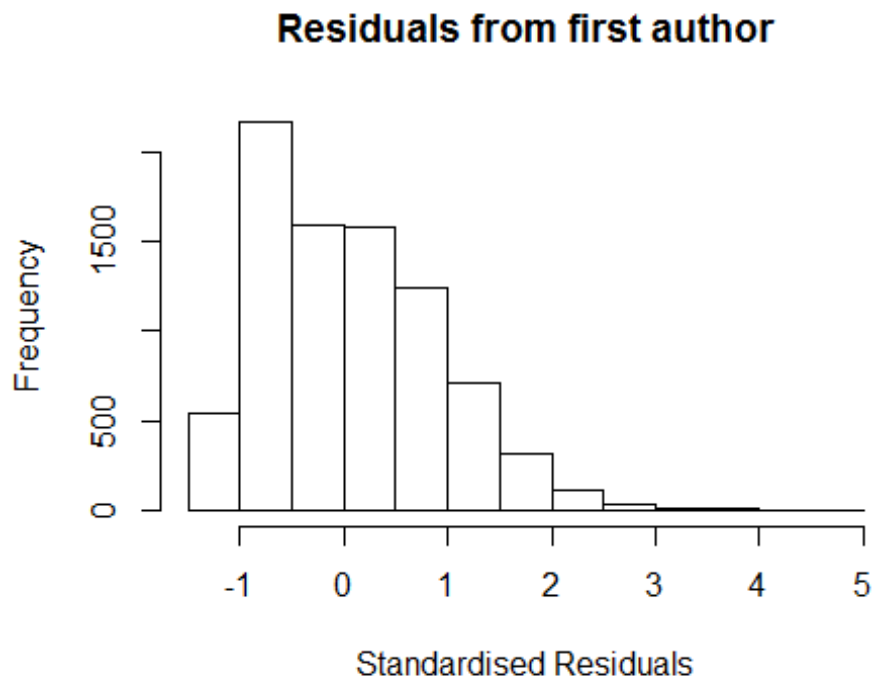
```
## Year2011           -0.00914    0.06302   -0.15    0.885
## Year2012            0.00655    0.06496    0.10    0.920
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.731
## Multiple R-squared:  0.192,  Adjusted R-squared:  0.19
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
##  4 observations c(273,529,752,1136)
##   are outliers with |weight| = 0 ( < 1.2e-05);
##  575 weights are ~= 1. The remaining 7715 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0006  0.8700  0.9450  0.9010  0.9780  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.21e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.801  1           1.342
## LastAuthorFemale  1.797  1           1.340
## Year              1.036 16           1.001
```




```
## [1] "List of  51 outliers with residuals above  2.5"
##          ScopusId  NLCS Year OneField Fields residuals
## 217    0030002054 3.857 1996     2911      1     2.989
## 607    0030776626 3.521 1997     2911      1     2.592
## 648    0030767805 4.270 1997     2911      1     3.341
## 1008  17644439119 3.543 1997     2911      1     2.614
## 1348   0031968557 3.847 1998     2911      1     3.144
## 1349   0031968843 3.316 1998     2911      1     2.613
## 1350   0031969182 3.760 1998     2911      1     3.057
## 1351   0031978565 3.901 1998     2911      1     3.198
## 1457   0031852597 5.054 1998     2911      1     4.351
## 1460   0031875098 3.269 1998     2911      1     2.566
## 2016   0344572791 5.730 1999     2911      1     5.019
## 2160   0033233216 3.895 1999     2911      1     3.184
## 2186   0033497846 3.590 1999     2911      1     2.879
## 2189   0033503671 3.731 1999     2911      1     2.868
## 2663   0033627592 3.876 2000     2911      1     3.106
## 2728   0034154033 4.072 2000     2911      1     3.382
## 2760   0034294822 4.543 2000     2911      1     3.702
## 2762   0034302388 3.904 2000     2911      1     3.286
## 2767   0034330243 4.048 2000     2911      1     3.278
## 2980  84993710017 3.429 2000     2910      2     2.739
## 3190   0038417818 3.189 2001     2911      1     2.516
## 3273   0035014575 3.189 2001     2911      1     2.516
## 3334   0002087395 3.266 2001     2911      1     2.593
## 3372   0035151691 4.191 2001     2911      1     3.518
## 3375   0035191266 3.189 2001     2911      1     2.516
```

```
## 3403   0035406399 3.337 2001     2911      1     2.584
## 3427   0035744421 3.759 2001     2911      1     2.934
## 3428   0035744427 4.234 2001     2911      1     3.409
## 3964   0036853306 3.424 2002     2911      1     2.596
## 4240   0036718125 4.315 2002     2911      1     3.559
## 4274   0036884123 3.314 2002     2911      1     2.558
## 4859   0037613484 4.445 2003     2911      1     3.486
## 4860   0038288856 3.497 2003     2911      1     2.690
## 4895   0037335173 3.386 2003     2911      1     2.579
## 4923   0037323010 3.919 2003     2911      1     3.112
## 5394   8544254752 3.927 2004     2911      1     3.091
## 5611   4043063731 4.798 2004     2911      1     3.882
## 5615   4043160018 3.891 2004     2911      1     3.055
## 6132  22144451148 3.602 2005     2911      1     2.544
## 6150  23844542129 4.350 2005     2911      1     3.292
## 6734  33947510199 3.556 2006     2719      5     2.526
## 6857  33745493311 3.519 2006     2911      1     2.569
## 6880  33748693969 4.535 2006     2911      1     3.434
## 8098  44249112154 4.538 2008     2911      1     3.389
## 8316  42149106861 3.786 2008     2911      1     2.637
## 8753  67449130309 3.706 2009     2719      5     2.881
## 8824  67649482684 3.748 2009     2719      5     2.772
## 9246  68749112732 3.466 2009     2911      1     2.570
## 10922 79960941189 3.752 2011     2911      1     2.712
## 11202 78651374145 3.444 2011     2911      1     2.555
## 11696 84865733130 3.859 2012     2911      1     2.857
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.1630 -0.7112 -0.0327  0.6029  5.0188
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)          0.9481     0.0602   15.75  < 2e-16 ***
## FirstAuthorFemale1   0.0715     0.0276    2.59   0.0095 **
## LastAuthorFemale1   -0.1517     0.0268   -5.65  1.6e-08 ***
## Year1997             0.0615     0.0911    0.68   0.4997
## Year1998            -0.1648     0.0822   -2.00   0.0451 *
## Year1999            -0.1567     0.0806   -1.95   0.0518 .
## Year2000            -0.1784     0.0750   -2.38   0.0174 *
## Year2001            -0.1949     0.0724   -2.69   0.0071 **
## Year2002            -0.0401     0.0677   -0.59   0.5536
## Year2003            -0.0610     0.0715   -0.85   0.3939
## Year2004            -0.0317     0.0715   -0.44   0.6581
## Year2005             0.1096     0.0716    1.53   0.1257
```

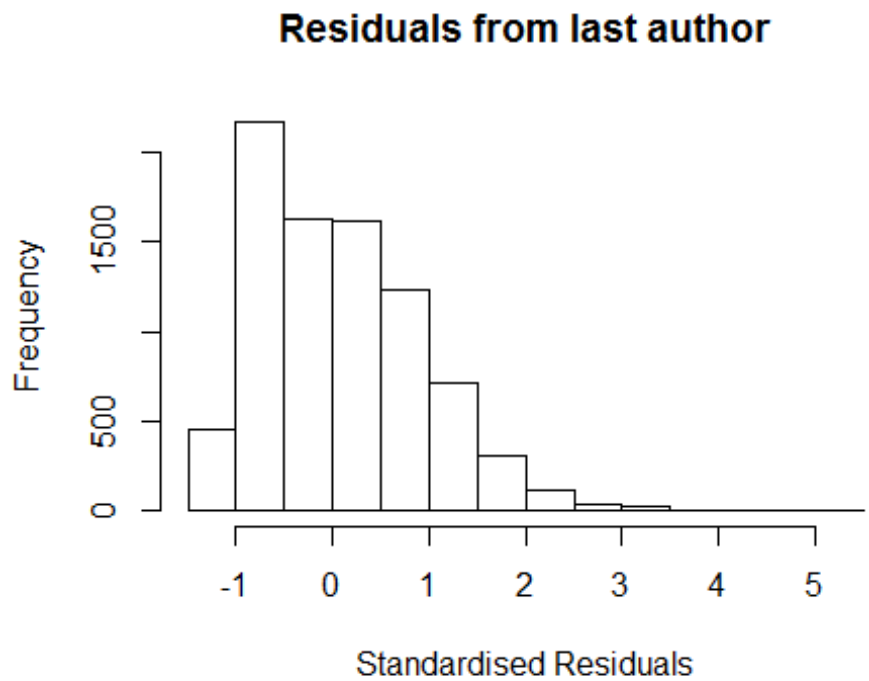
```
## Year2006             0.0818     0.0738    1.11   0.2681
## Year2007             0.1434     0.0722    1.99   0.0469 *
## Year2008             0.1290     0.0683    1.89   0.0590 .
## Year2009             0.0281     0.0684    0.41   0.6807
## Year2010             0.0628     0.0672    0.93   0.3499
## Year2011             0.0209     0.0659    0.32   0.7512
## Year2012             0.1341     0.0681    1.97   0.0488 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.864
## Multiple R-squared:  0.0194, Adjusted R-squared:  0.0173
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  2 observations c(529,752) are outliers with |weight| = 0 ( < 1.2e-05);
##  656 weights are ~= 1. The remaining 7636 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0066  0.8870  0.9390  0.9130  0.9810  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.21e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.023  1           1.012
## Year              1.023 16           1.001
```




```
## [1] "List of  51 outliers with residuals above  2.5"
##          ScopusId  NLCS Year OneField Fields residuals
## 217    0030002054 3.857 1996     2911      1     2.989
## 607    0030776626 3.521 1997     2911      1     2.592
## 648    0030767805 4.270 1997     2911      1     3.341
## 1008  17644439119 3.543 1997     2911      1     2.614
## 1348   0031968557 3.847 1998     2911      1     3.144
## 1349   0031968843 3.316 1998     2911      1     2.613
## 1350   0031969182 3.760 1998     2911      1     3.057
## 1351   0031978565 3.901 1998     2911      1     3.198
## 1457   0031852597 5.054 1998     2911      1     4.351
## 1460   0031875098 3.269 1998     2911      1     2.566
## 2016   0344572791 5.730 1999     2911      1     5.019
## 2160   0033233216 3.895 1999     2911      1     3.184
## 2186   0033497846 3.590 1999     2911      1     2.879
## 2189   0033503671 3.731 1999     2911      1     2.868
## 2663   0033627592 3.876 2000     2911      1     3.106
## 2728   0034154033 4.072 2000     2911      1     3.382
## 2760   0034294822 4.543 2000     2911      1     3.702
## 2762   0034302388 3.904 2000     2911      1     3.286
## 2767   0034330243 4.048 2000     2911      1     3.278
## 2980  84993710017 3.429 2000     2910      2     2.739
## 3190   0038417818 3.189 2001     2911      1     2.516
## 3273   0035014575 3.189 2001     2911      1     2.516
## 3334   0002087395 3.266 2001     2911      1     2.593
## 3372   0035151691 4.191 2001     2911      1     3.518
## 3375   0035191266 3.189 2001     2911      1     2.516
```

```
## 3403   0035406399 3.337 2001     2911      1     2.584
## 3427   0035744421 3.759 2001     2911      1     2.934
## 3428   0035744427 4.234 2001     2911      1     3.409
## 3964   0036853306 3.424 2002     2911      1     2.596
## 4240   0036718125 4.315 2002     2911      1     3.559
## 4274   0036884123 3.314 2002     2911      1     2.558
## 4859   0037613484 4.445 2003     2911      1     3.486
## 4860   0038288856 3.497 2003     2911      1     2.690
## 4895   0037335173 3.386 2003     2911      1     2.579
## 4923   0037323010 3.919 2003     2911      1     3.112
## 5394   8544254752 3.927 2004     2911      1     3.091
## 5611   4043063731 4.798 2004     2911      1     3.882
## 5615   4043160018 3.891 2004     2911      1     3.055
## 6132  22144451148 3.602 2005     2911      1     2.544
## 6150  23844542129 4.350 2005     2911      1     3.292
## 6734  33947510199 3.556 2006     2719      5     2.526
## 6857  33745493311 3.519 2006     2911      1     2.569
## 6880  33748693969 4.535 2006     2911      1     3.434
## 8098  44249112154 4.538 2008     2911      1     3.389
## 8316  42149106861 3.786 2008     2911      1     2.637
## 8753  67449130309 3.706 2009     2719      5     2.881
## 8824  67649482684 3.748 2009     2719      5     2.772
## 9246  68749112732 3.466 2009     2911      1     2.570
## 10922 79960941189 3.752 2011     2911      1     2.712
## 11202 78651374145 3.444 2011     2911      1     2.555
## 11696 84865733130 3.859 2012     2911      1     2.857
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.0664 -0.7157 -0.0335  0.6092  4.9987
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)          0.9212     0.0595   15.48   <2e-16 ***
## FirstAuthorFemale1  -0.0271     0.0206   -1.31   0.1894
## Year1997             0.0560     0.0906    0.62   0.5366
## Year1998            -0.1618     0.0824   -1.96   0.0497 *
## Year1999            -0.1628     0.0803   -2.03   0.0427 *
## Year2000            -0.1819     0.0747   -2.44   0.0149 *
## Year2001            -0.2054     0.0719   -2.86   0.0043 **
## Year2002            -0.0486     0.0673   -0.72   0.4702
## Year2003            -0.0659     0.0711   -0.93   0.3539
## Year2004            -0.0388     0.0712   -0.54   0.5861
## Year2005             0.1063     0.0713    1.49   0.1359
## Year2006             0.0843     0.0734    1.15   0.2507
```

```
## Year2007             0.1453     0.0719    2.02   0.0435 *
## Year2008             0.1210     0.0679    1.78   0.0749 .
## Year2009             0.0280     0.0680    0.41   0.6803
## Year2010             0.0595     0.0668    0.89   0.3737
## Year2011             0.0166     0.0655    0.25   0.7995
## Year2012             0.1264     0.0679    1.86   0.0627 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.869
## Multiple R-squared:  0.0149, Adjusted R-squared:  0.0129
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  2 observations c(529,752) are outliers with |weight| = 0 ( < 1.2e-05);
##  656 weights are ~= 1. The remaining 7636 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0056  0.8880  0.9370  0.9130  0.9810  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.21e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##                  GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.02  1           1.010
## Year             1.02 16           1.001
```

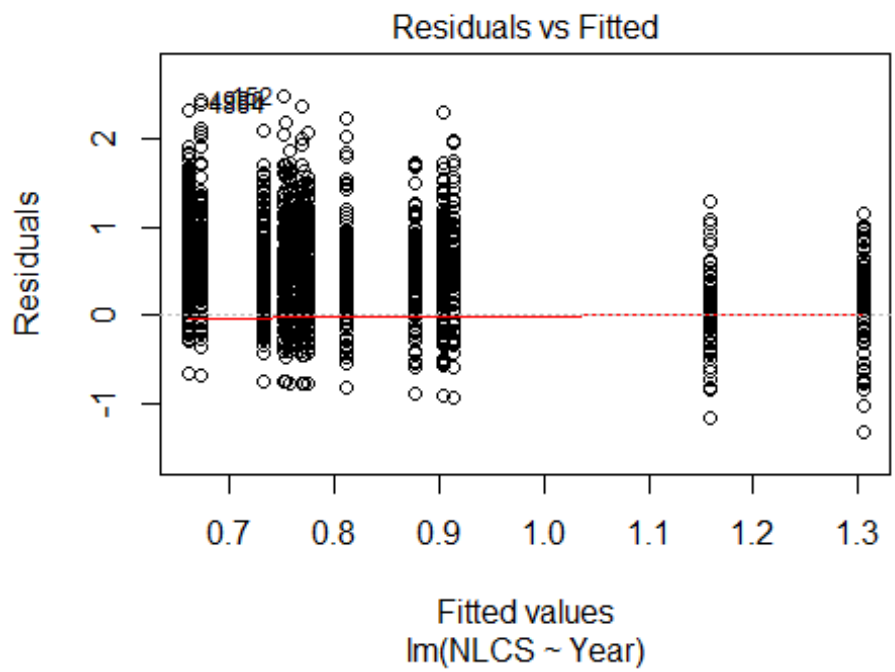



```
## [1] "List of  51 outliers with residuals above  2.5"
##          ScopusId  NLCS Year OneField Fields residuals
## 217    0030002054 3.857 1996     2911      1     2.989
## 607    0030776626 3.521 1997     2911      1     2.592
## 648    0030767805 4.270 1997     2911      1     3.341
## 1008  17644439119 3.543 1997     2911      1     2.614
## 1348   0031968557 3.847 1998     2911      1     3.144
## 1349   0031968843 3.316 1998     2911      1     2.613
## 1350   0031969182 3.760 1998     2911      1     3.057
## 1351   0031978565 3.901 1998     2911      1     3.198
## 1457   0031852597 5.054 1998     2911      1     4.351
## 1460   0031875098 3.269 1998     2911      1     2.566
## 2016   0344572791 5.730 1999     2911      1     5.019
## 2160   0033233216 3.895 1999     2911      1     3.184
## 2186   0033497846 3.590 1999     2911      1     2.879
## 2189   0033503671 3.731 1999     2911      1     2.868
## 2663   0033627592 3.876 2000     2911      1     3.106
## 2728   0034154033 4.072 2000     2911      1     3.382
## 2760   0034294822 4.543 2000     2911      1     3.702
## 2762   0034302388 3.904 2000     2911      1     3.286
## 2767   0034330243 4.048 2000     2911      1     3.278
## 2980  84993710017 3.429 2000     2910      2     2.739
## 3190   0038417818 3.189 2001     2911      1     2.516
## 3273   0035014575 3.189 2001     2911      1     2.516
## 3334   0002087395 3.266 2001     2911      1     2.593
## 3372   0035151691 4.191 2001     2911      1     3.518
## 3375   0035191266 3.189 2001     2911      1     2.516
```

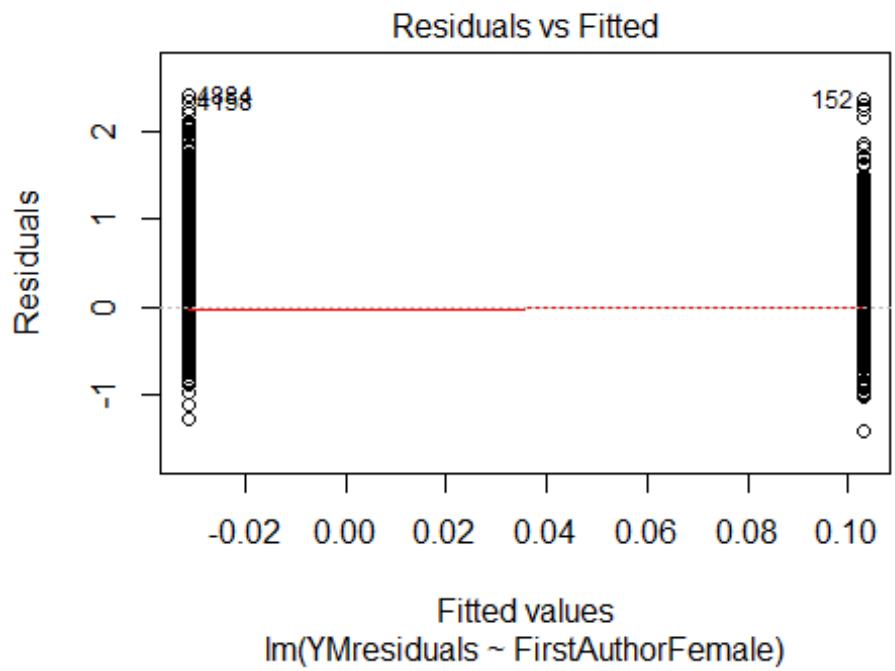
```
## 3403   0035406399 3.337 2001     2911      1     2.584
## 3427   0035744421 3.759 2001     2911      1     2.934
## 3428   0035744427 4.234 2001     2911      1     3.409
## 3964   0036853306 3.424 2002     2911      1     2.596
## 4240   0036718125 4.315 2002     2911      1     3.559
## 4274   0036884123 3.314 2002     2911      1     2.558
## 4859   0037613484 4.445 2003     2911      1     3.486
## 4860   0038288856 3.497 2003     2911      1     2.690
## 4895   0037335173 3.386 2003     2911      1     2.579
## 4923   0037323010 3.919 2003     2911      1     3.112
## 5394   8544254752 3.927 2004     2911      1     3.091
## 5611   4043063731 4.798 2004     2911      1     3.882
## 5615   4043160018 3.891 2004     2911      1     3.055
## 6132  22144451148 3.602 2005     2911      1     2.544
## 6150  23844542129 4.350 2005     2911      1     3.292
## 6734  33947510199 3.556 2006     2719      5     2.526
## 6857  33745493311 3.519 2006     2911      1     2.569
## 6880  33748693969 4.535 2006     2911      1     3.434
## 8098  44249112154 4.538 2008     2911      1     3.389
## 8316  42149106861 3.786 2008     2911      1     2.637
## 8753  67449130309 3.706 2009     2719      5     2.881
## 8824  67649482684 3.748 2009     2719      5     2.772
## 9246  68749112732 3.466 2009     2911      1     2.570
## 10922 79960941189 3.752 2011     2911      1     2.712
## 11202 78651374145 3.444 2011     2911      1     2.555
## 11696 84865733130 3.859 2012     2911      1     2.857
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.1118 -0.7017 -0.0328  0.6069  5.0283
##
## Coefficients:
##                   Estimate Std. Error t value Pr(>|t|)
## (Intercept)         0.9626     0.0600   16.06  < 2e-16 ***
## LastAuthorFemale1  -0.1071     0.0203   -5.28  1.4e-07 ***
## Year1997            0.0642     0.0909    0.71   0.4802
## Year1998           -0.1604     0.0822   -1.95   0.0512 .
## Year1999           -0.1538     0.0805   -1.91   0.0561 .
## Year2000           -0.1764     0.0748   -2.36   0.0184 *
## Year2001           -0.1896     0.0722   -2.63   0.0086 **
## Year2002           -0.0338     0.0674   -0.50   0.6156
## Year2003           -0.0578     0.0713   -0.81   0.4175
## Year2004           -0.0253     0.0713   -0.35   0.7232
## Year2005            0.1148     0.0713    1.61   0.1075
## Year2006            0.0878     0.0735    1.19   0.2326
```

```
## Year2007            0.1491     0.0719    2.07   0.0382 *
## Year2008            0.1323     0.0681    1.94   0.0520 .
## Year2009            0.0352     0.0681    0.52   0.6056
## Year2010            0.0694     0.0669    1.04   0.2995
## Year2011            0.0277     0.0655    0.42   0.6730
## Year2012            0.1402     0.0678    2.07   0.0389 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.865
## Multiple R-squared:  0.0185, Adjusted R-squared:  0.0164
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  2 observations c(529,752) are outliers with |weight| = 0 ( < 1.2e-05);
##  643 weights are ~= 1. The remaining 7649 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0087  0.8890  0.9400  0.9130  0.9820  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.21e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Sample size for the above analysis:  8294"
## [1] ""
## [1] ""
## [1] "####################################"
## [1] "Analysis of AJSC 2912"
## [1] "####################################"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  182  374  389  323  358  226  536  475  397  341  288  254  394  454  433
## 2011 2012
##  504  493
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  118  149  172  123   95   88  301  275  319  274  254  223  366  415  384
```

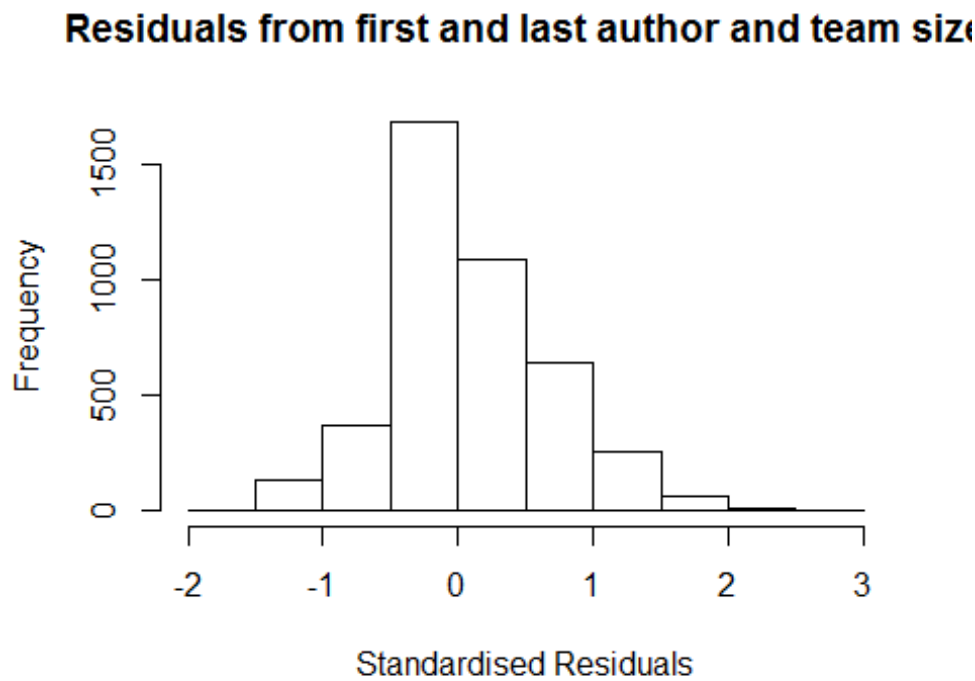
```
## 2011 2012
##  454  435
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  115  139  161  119   86   83  282  266  313  255  246  219  354  394  369
## 2011 2012
##  423  420
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
##  Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 43, df = 16, p-value = 3e-04
```




```
##
##  Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 0.46, df = 1, p-value = 0.5
```

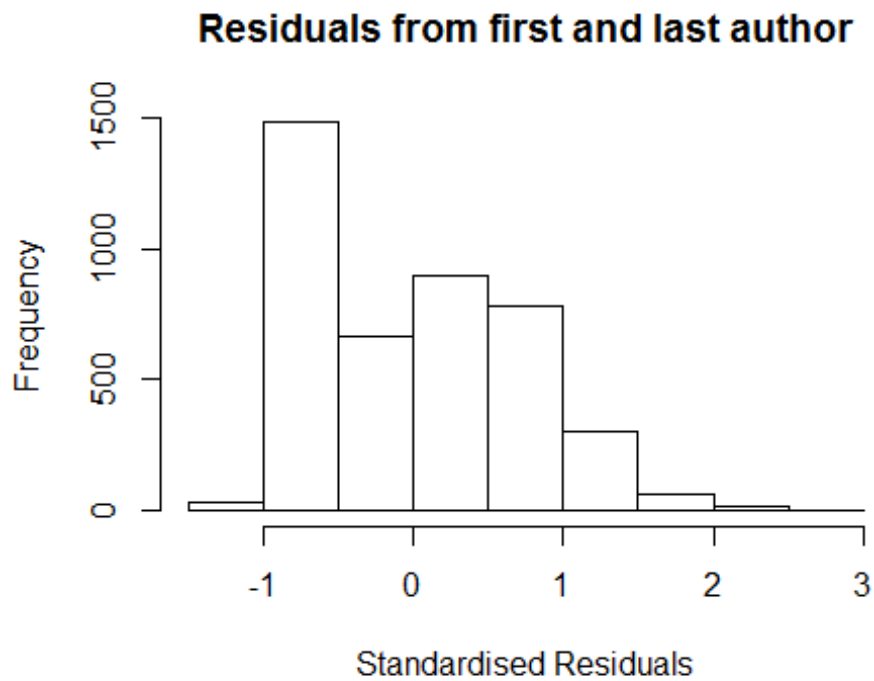



```
## [1] "Female first author team size 2018 geometric mean: 2.02288752594169"
## [1] "Male first author team size 2018 geometric mean: 2.08046083936429"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 6000, p-value = 0.7
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 1.9652062453304"
## [1] "Male last author team size 2018 geometric mean: 2.34067843718076"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 5700, p-value = 0.04
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.786  1           1.337
## LastAuthorFemale  1.842  1           1.357
## UniqueAuthors     1.292  4           1.033
## Year              1.285 16           1.008
```



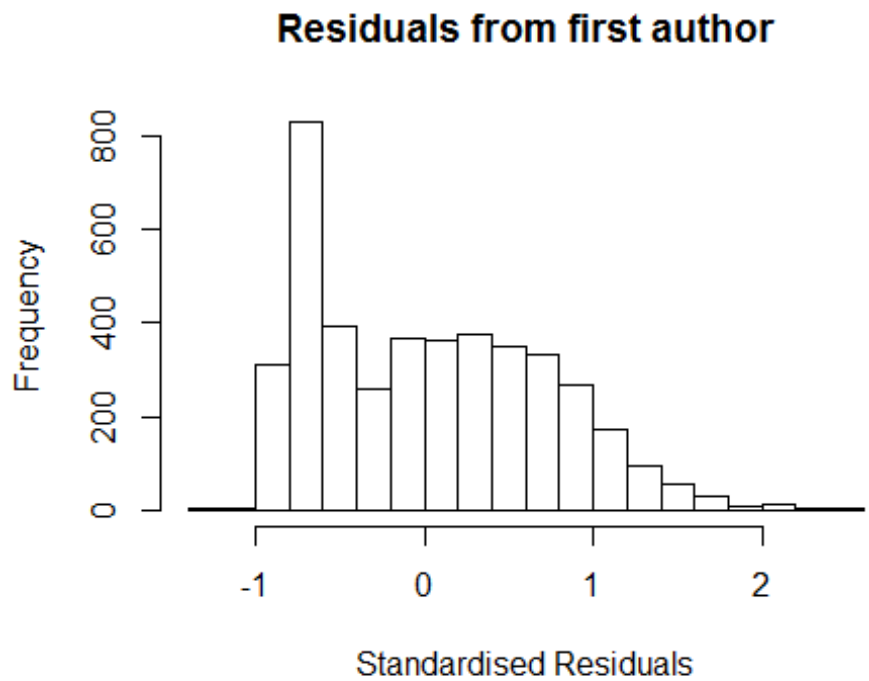

```
## [1] "List of  1 outliers with residuals above  2.5"
##         ScopusId  NLCS Year OneField Fields residuals
## 4884 67650084651 3.058 2009     2902      4     2.787
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
UniqueAuthors +
##     Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##     k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.6767 -0.3834 -0.0646  0.4557  2.7866
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         0.49116    0.05969    8.23  2.5e-16 ***
## FirstAuthorFemale1 -0.00904    0.03057   -0.30    0.767
## LastAuthorFemale1  -0.12063    0.02940   -4.10  4.1e-05 ***
## UniqueAuthors2      0.59156    0.03051   19.39  < 2e-16 ***
## UniqueAuthors3      0.77877    0.03720   20.93  < 2e-16 ***
## UniqueAuthors4      0.84055    0.03998   21.02  < 2e-16 ***
## UniqueAuthors5      0.91493    0.04731   19.34  < 2e-16 ***
## Year1997            0.05859    0.08519    0.69    0.492
## Year1998            0.02102    0.07908    0.27    0.790
## Year1999           -0.06077    0.07771   -0.78    0.434
```

```
## Year2000            0.40024    0.08908    4.49  7.2e-06 ***
## Year2001            0.19508    0.07989    2.44    0.015 *
## Year2002            0.12183    0.07191    1.69    0.090 .
## Year2003           -0.03863    0.06562   -0.59    0.556
## Year2004           -0.03290    0.06727   -0.49    0.625
## Year2005            0.10503    0.06970    1.51    0.132
## Year2006            0.05336    0.07222    0.74    0.460
## Year2007           -0.10778    0.06595   -1.63    0.102
## Year2008           -0.07084    0.06500   -1.09    0.276
## Year2009           -0.09007    0.06351   -1.42    0.156
## Year2010            0.03887    0.06489    0.60    0.549
## Year2011           -0.04300    0.06399   -0.67    0.502
## Year2012           -0.05640    0.06495   -0.87    0.385
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.5
## Multiple R-squared:  0.324,  Adjusted R-squared:  0.321
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
##  2 observations c(2809,2888) are outliers with |weight| = 0 ( < 2.4e-05);
##  148 weights are ~= 1. The remaining 4094 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0347  0.8440  0.9380  0.8760  0.9730  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          2.36e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                    psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.771  1           1.331
## LastAuthorFemale  1.815  1           1.347
## Year              1.076 16           1.002
```



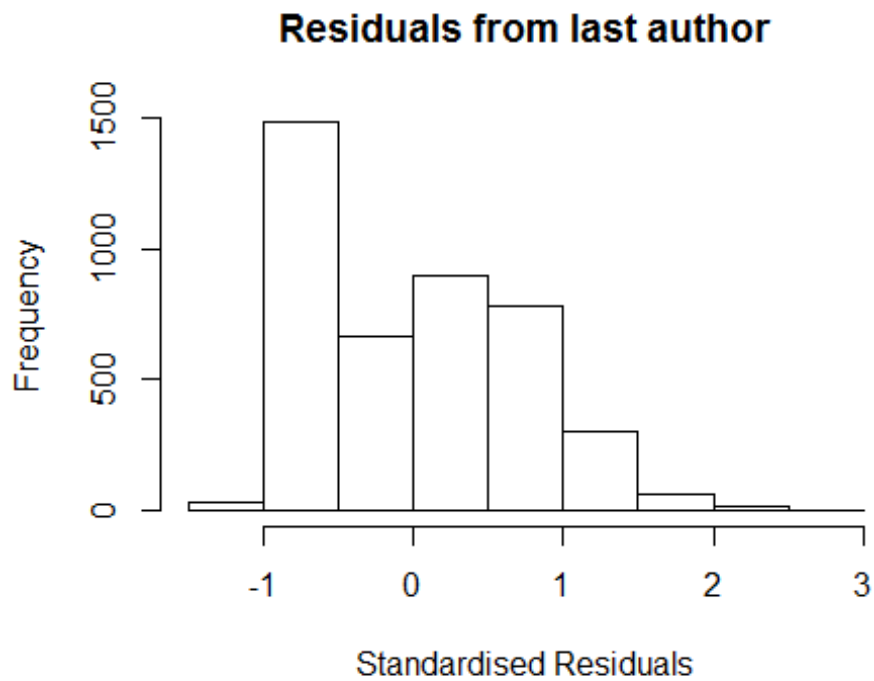

```
## [1] "List of  2 outliers with residuals above  2.5"
##         ScopusId  NLCS Year OneField Fields residuals
## 4884 67650084651 3.058 2009     2902      4     2.511
## 4980 62849111793 3.089 2009     2908      4     2.565
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.4121 -0.6075 -0.0364  0.5500  2.5651
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)           0.8204     0.0690   11.89  < 2e-16 ***
## FirstAuthorFemale1    0.0229     0.0375    0.61   0.5412
## LastAuthorFemale1    -0.2612     0.0348   -7.50  7.6e-14 ***
## Year1997              0.1585     0.0975    1.63   0.1042
## Year1998              0.0241     0.0892    0.27   0.7869
## Year1999              0.0254     0.0928    0.27   0.7844
## Year2000              0.5918     0.0923    6.41  1.6e-10 ***
## Year2001              0.4409     0.0841    5.24  1.6e-07 ***
## Year2002              0.2452     0.0826    2.97   0.0030 **
## Year2003              0.0366     0.0812    0.45   0.6526
## Year2004              0.0543     0.0773    0.70   0.4825
```

```
## Year2005             0.2261     0.0797    2.84   0.0046 **
## Year2006             0.1254     0.0834    1.50   0.1329
## Year2007            -0.0398     0.0822   -0.48   0.6289
## Year2008            -0.0361     0.0774   -0.47   0.6412
## Year2009            -0.0352     0.0750   -0.47   0.6383
## Year2010             0.1114     0.0757    1.47   0.1411
## Year2011             0.0966     0.0742    1.30   0.1929
## Year2012             0.0823     0.0755    1.09   0.2759
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.729
## Multiple R-squared:  0.0542, Adjusted R-squared:  0.0501
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  323 weights are ~= 1. The remaining 3921 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.190   0.891   0.935   0.918   0.975   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          2.36e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                  psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.037  1           1.018
## Year              1.037 16           1.001
```




```
## [1] "List of  2 outliers with residuals above  2.5"
##         ScopusId  NLCS Year OneField Fields residuals
## 4884 67650084651 3.058 2009     2902      4     2.511
## 4980 62849111793 3.089 2009     2908      4     2.565
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.3775 -0.6564 -0.0387  0.5535  2.4826
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)           0.7667     0.0689   11.13  < 2e-16 ***
## FirstAuthorFemale1   -0.1484     0.0276   -5.38  7.8e-08 ***
## Year1997              0.1717     0.0992    1.73   0.0835 .
## Year1998              0.0236     0.0903    0.26   0.7935
## Year1999              0.0388     0.0938    0.41   0.6789
## Year2000              0.6108     0.0934    6.54  6.8e-11 ***
## Year2001              0.4821     0.0848    5.69  1.4e-08 ***
## Year2002              0.2322     0.0835    2.78   0.0054 **
## Year2003              0.0382     0.0817    0.47   0.6406
## Year2004              0.0440     0.0782    0.56   0.5737
## Year2005              0.2172     0.0807    2.69   0.0071 **
```

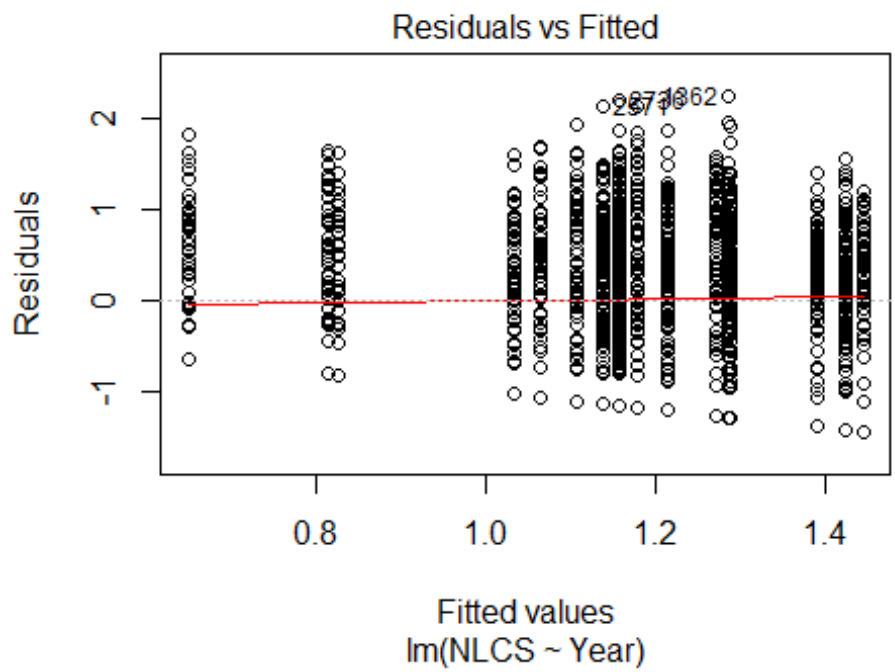
```
## Year2006             0.1188     0.0839    1.42   0.1566
## Year2007            -0.0555     0.0834   -0.67   0.5054
## Year2008            -0.0450     0.0783   -0.58   0.5651
## Year2009            -0.0429     0.0759   -0.57   0.5718
## Year2010             0.0969     0.0765    1.27   0.2053
## Year2011             0.0887     0.0753    1.18   0.2385
## Year2012             0.0732     0.0760    0.96   0.3361
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.746
## Multiple R-squared:  0.0392, Adjusted R-squared:  0.0353
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  277 weights are ~= 1. The remaining 3967 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.245   0.900   0.933   0.921   0.978   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          2.36e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##                   GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.065  1           1.032
## Year             1.065 16           1.002
```



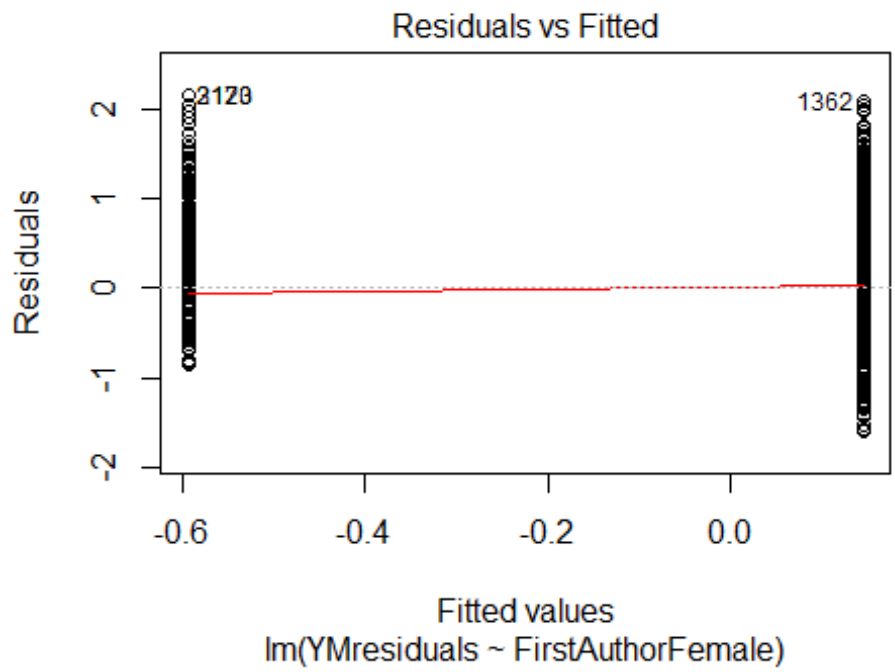

```
## [1] "List of  2 outliers with residuals above  2.5"
##         ScopusId  NLCS Year OneField Fields residuals
## 4884 67650084651 3.058 2009     2902      4     2.511
## 4980 62849111793 3.089 2009     2908      4     2.565
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.4186 -0.6048 -0.0393  0.5517  2.5441
##
## Coefficients:
##                   Estimate Std. Error t value Pr(>|t|)
## (Intercept)         0.8255     0.0679   12.15  < 2e-16 ***
## LastAuthorFemale1  -0.2478     0.0273   -9.09  < 2e-16 ***
## Year1997            0.1607     0.0975    1.65   0.0992 .
## Year1998            0.0257     0.0892    0.29   0.7733
## Year1999            0.0271     0.0927    0.29   0.7702
## Year2000            0.5931     0.0921    6.44  1.3e-10 ***
## Year2001            0.4436     0.0837    5.30  1.2e-07 ***
## Year2002            0.2485     0.0822    3.02   0.0025 **
## Year2003            0.0387     0.0812    0.48   0.6335
## Year2004            0.0565     0.0773    0.73   0.4651
## Year2005            0.2296     0.0796    2.88   0.0039 **
```

```
## Year2006            0.1278     0.0833    1.53   0.1252
## Year2007           -0.0382     0.0822   -0.46   0.6425
## Year2008           -0.0341     0.0774   -0.44   0.6591
## Year2009           -0.0328     0.0749   -0.44   0.6614
## Year2010            0.1144     0.0755    1.52   0.1298
## Year2011            0.0996     0.0741    1.34   0.1788
## Year2012            0.0850     0.0754    1.13   0.2594
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.724
## Multiple R-squared:  0.0543, Adjusted R-squared:  0.0505
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  324 weights are ~= 1. The remaining 3920 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.192   0.890   0.935   0.917   0.974   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          2.36e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                    psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Sample size for the above analysis:  4244"
## [1] ""
## [1] ""
## [1] "####################################"
## [1] "Analysis of AJSC 2913"
## [1] "####################################"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  195  192  178  150  175  143  149  136  123   94  126  144  180  229  273
## 2011 2012
##  280  272
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##   89   99   87   82  112   88  134  123  108   84  113  129  157  202  235
```

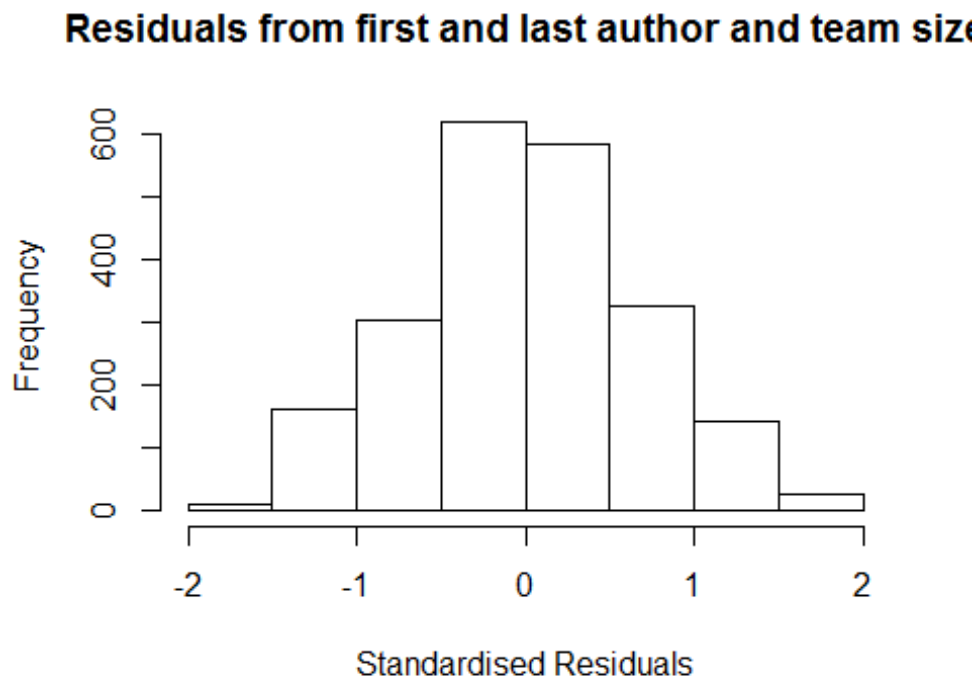
```
## 2011 2012
##  255  236
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##   83   92   82   80  108   84  129  114  100   77  103  119  144  191  219
## 2011 2012
##  237  211
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
##  Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 40, df = 16, p-value = 7e-04
```




```
##
##  Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 6.5, df = 1, p-value = 0.01
```

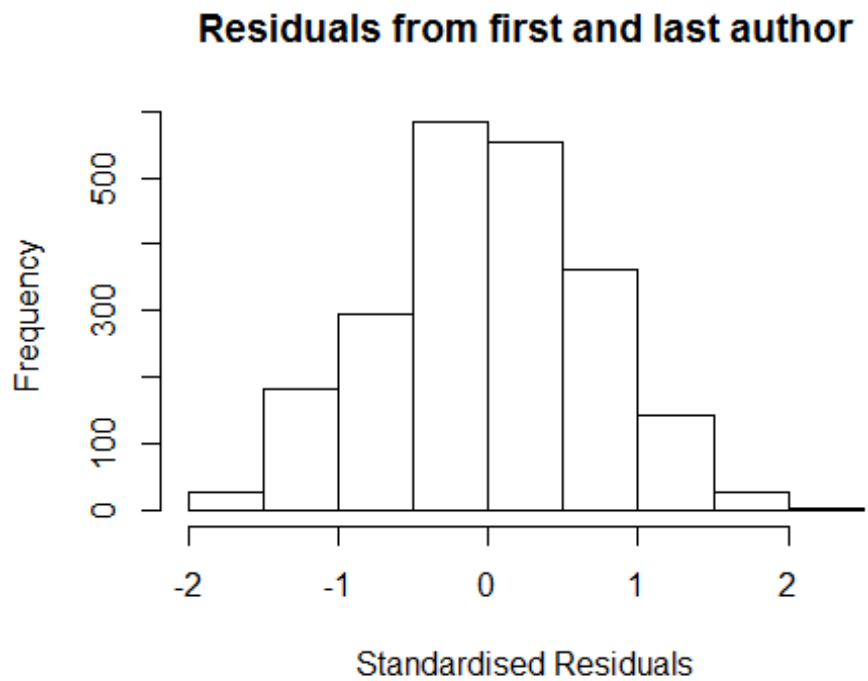



```
## [1] "Female first author team size 2018 geometric mean: 3.01821335758453"
## [1] "Male first author team size 2018 geometric mean: 2.54873734426022"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 3000, p-value = 0.4
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 2.92717574989429"
## [1] "Male last author team size 2018 geometric mean: 3.00150601805117"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 4000, p-value = 0.8
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 2.693  1           1.641
## LastAuthorFemale  2.392  1           1.547
## UniqueAuthors     1.455  4           1.048
## Year              1.410 16           1.011
```



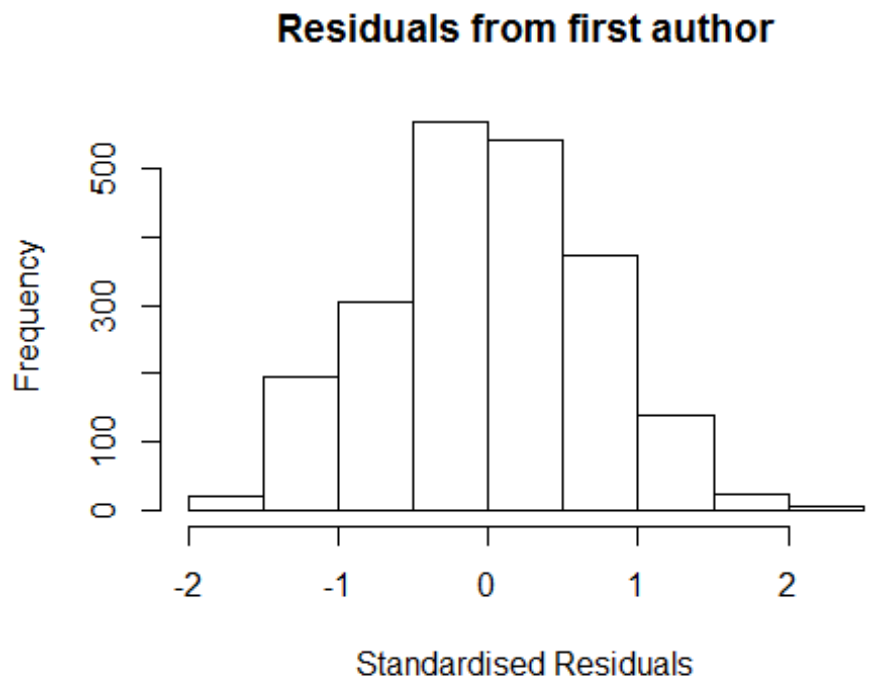

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
UniqueAuthors +
##     Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##     k.max = 1000))
##  \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.83445 -0.44715 -0.00511  0.46635  1.99867
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         -0.0412     0.0733   -0.56  0.57453
## FirstAuthorFemale1   0.5236     0.0501   10.44  < 2e-16 ***
## LastAuthorFemale1    0.3463     0.0461    7.51  8.8e-14 ***
## UniqueAuthors2       0.1414     0.0424    3.33  0.00088 ***
## UniqueAuthors3       0.4149     0.0482    8.60  < 2e-16 ***
## UniqueAuthors4       0.4988     0.0514    9.71  < 2e-16 ***
## UniqueAuthors5       0.5171     0.0484   10.69  < 2e-16 ***
## Year1997             0.1134     0.1045    1.09  0.27760
## Year1998             0.1355     0.1006    1.35  0.17817
## Year1999             0.1615     0.1080    1.50  0.13474
```

```
## Year2000             0.3432     0.1024    3.35  0.00081 ***
## Year2001             0.3525     0.1124    3.14  0.00174 **
## Year2002             0.3724     0.0991    3.76  0.00018 ***
## Year2003             0.4704     0.1004    4.69  3.0e-06 ***
## Year2004             0.4887     0.1160    4.21  2.6e-05 ***
## Year2005             0.5257     0.1017    5.17  2.5e-07 ***
## Year2006             0.4847     0.1044    4.64  3.7e-06 ***
## Year2007             0.4466     0.1026    4.35  1.4e-05 ***
## Year2008             0.5014     0.0938    5.35  9.9e-08 ***
## Year2009             0.3363     0.0903    3.72  0.00020 ***
## Year2010             0.2505     0.0875    2.86  0.00426 **
## Year2011             0.2470     0.0851    2.90  0.00374 **
## Year2012             0.1567     0.0901    1.74  0.08203 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.674
## Multiple R-squared:  0.285,  Adjusted R-squared:  0.277
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  164 weights are ~= 1. The remaining 2009 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.359   0.862   0.953   0.906   0.985   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          4.60e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 2.582  1           1.607
## LastAuthorFemale  2.428  1           1.558
## Year              1.197 16           1.006
```



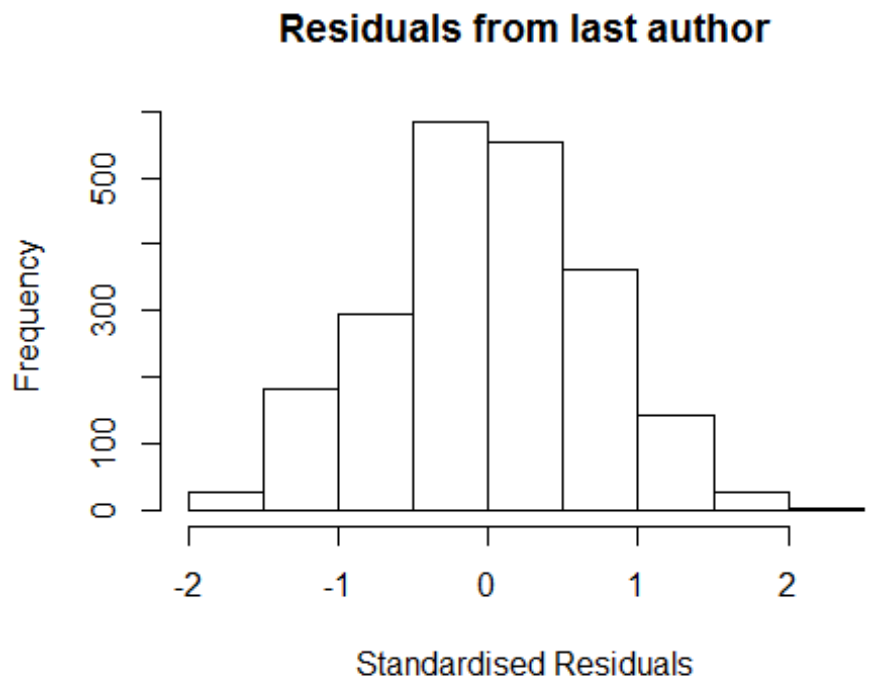

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##       Min        1Q    Median        3Q       Max
## -1.595562 -0.469437  0.000682  0.491438  2.300134
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)           0.0437     0.0739    0.59  0.55428
## FirstAuthorFemale1    0.6327     0.0559   11.32  < 2e-16 ***
## LastAuthorFemale1     0.2809     0.0509    5.52  3.8e-08 ***
## Year1997              0.1153     0.1067    1.08  0.28014
## Year1998              0.1731     0.0997    1.74  0.08258 .
## Year1999              0.2238     0.1105    2.03  0.04297 *
## Year2000              0.3641     0.1042    3.49  0.00049 ***
## Year2001              0.3875     0.1101    3.52  0.00044 ***
## Year2002              0.4234     0.1036    4.09  4.5e-05 ***
## Year2003              0.5011     0.1018    4.92  9.3e-07 ***
## Year2004              0.4808     0.1191    4.04  5.6e-05 ***
## Year2005              0.6204     0.1028    6.04  1.9e-09 ***
```

```
## Year2006             0.6382     0.1048    6.09  1.3e-09 ***
## Year2007             0.5579     0.1031    5.41  6.8e-08 ***
## Year2008             0.6322     0.0946    6.68  2.9e-11 ***
## Year2009             0.4305     0.0924    4.66  3.3e-06 ***
## Year2010             0.3335     0.0909    3.67  0.00025 ***
## Year2011             0.3519     0.0877    4.01  6.2e-05 ***
## Year2012             0.3247     0.0934    3.48  0.00051 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.686
## Multiple R-squared:  0.232,  Adjusted R-squared:  0.225
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
##  164 weights are ~= 1. The remaining 2009 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.238   0.861   0.948   0.902   0.982   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          4.60e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.191  1           1.091
## Year              1.191 16           1.005
```




```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.55990 -0.47861 -0.00705  0.49889  2.26604
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)           0.0758     0.0764    0.99  0.32121
## FirstAuthorFemale1    0.8331     0.0404   20.63  < 2e-16 ***
## Year1997              0.1319     0.1085    1.22  0.22431
## Year1998              0.1937     0.1013    1.91  0.05616 .
## Year1999              0.2224     0.1142    1.95  0.05158 .
## Year2000              0.3977     0.1060    3.75  0.00018 ***
## Year2001              0.4000     0.1124    3.56  0.00038 ***
## Year2002              0.4292     0.1080    3.97  7.3e-05 ***
## Year2003              0.5221     0.1037    5.03  5.2e-07 ***
## Year2004              0.4833     0.1191    4.06  5.1e-05 ***
## Year2005              0.6309     0.1057    5.97  2.8e-09 ***
## Year2006              0.6510     0.1058    6.15  9.1e-10 ***
```

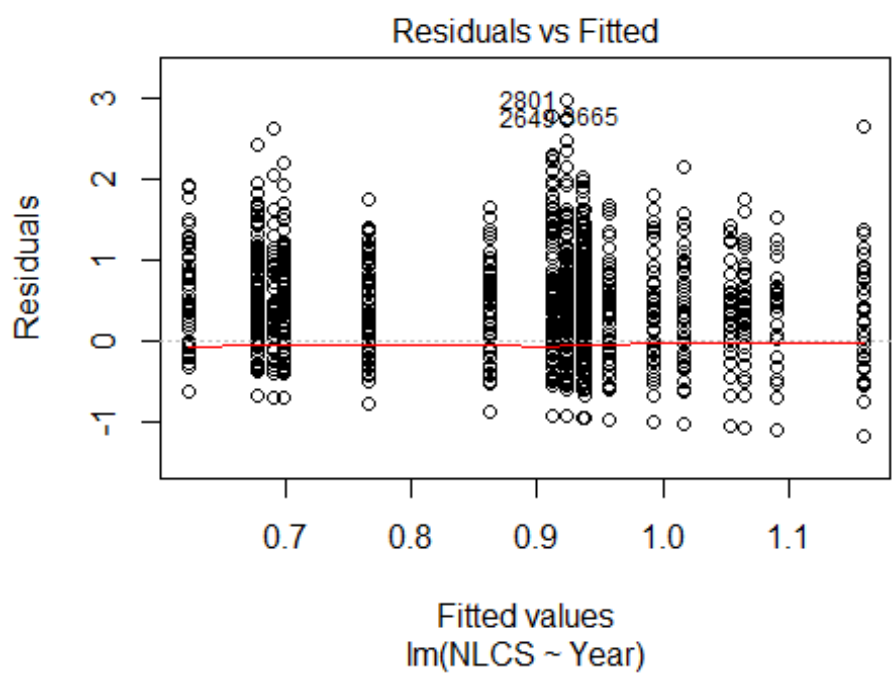
```
## Year2007             0.5615     0.1062    5.29  1.4e-07 ***
## Year2008             0.6342     0.0954    6.65  3.8e-11 ***
## Year2009             0.4320     0.0944    4.58  5.0e-06 ***
## Year2010             0.3465     0.0921    3.76  0.00017 ***
## Year2011             0.3732     0.0892    4.18  3.0e-05 ***
## Year2012             0.3243     0.0949    3.42  0.00065 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.686
## Multiple R-squared:  0.22,   Adjusted R-squared:  0.214
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
##  147 weights are ~= 1. The remaining 2026 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.253   0.862   0.951   0.902   0.983   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          4.60e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##                   GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.131  1           1.063
## Year             1.131 16           1.004
```



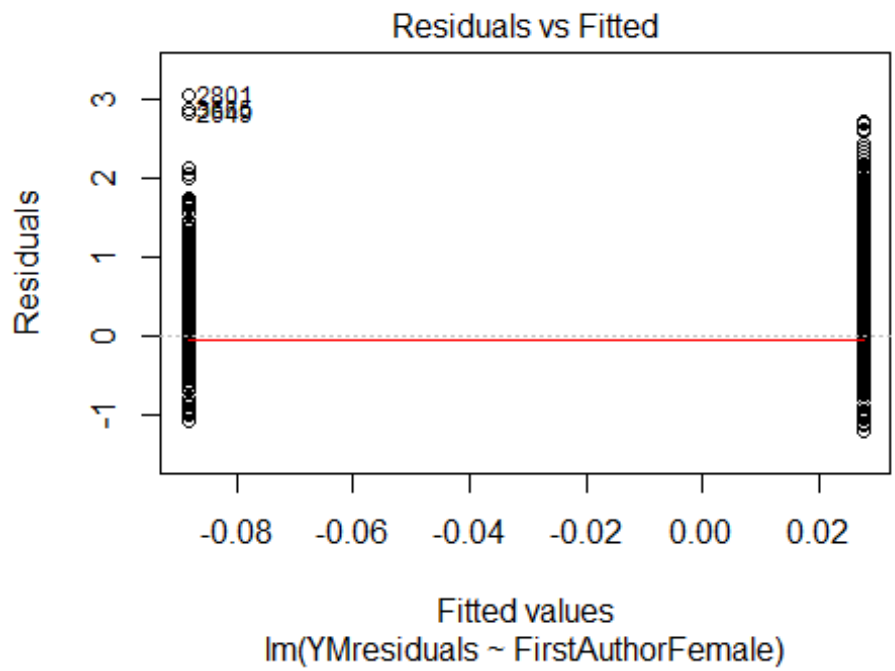

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.58503 -0.55598 -0.00111  0.51906  2.32724
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)          0.2234     0.0693    3.22  0.00128 **
## LastAuthorFemale1    0.6362     0.0415   15.32  < 2e-16 ***
## Year1997             0.1357     0.1018    1.33  0.18256
## Year1998             0.1623     0.0976    1.66  0.09637 .
## Year1999             0.3326     0.1060    3.14  0.00172 **
## Year2000             0.3832     0.1055    3.63  0.00029 ***
## Year2001             0.4230     0.1119    3.78  0.00016 ***
## Year2002             0.4651     0.1012    4.60  4.6e-06 ***
## Year2003             0.5441     0.1022    5.32  1.1e-07 ***
## Year2004             0.5483     0.1228    4.47  8.4e-06 ***
## Year2005             0.7070     0.1010    7.00  3.5e-12 ***
## Year2006             0.7188     0.1039    6.92  5.9e-12 ***
```

```
## Year2007            0.6304     0.1015    6.21  6.3e-10 ***
## Year2008            0.7254     0.0939    7.73  1.7e-14 ***
## Year2009            0.5145     0.0889    5.79  8.1e-09 ***
## Year2010            0.4051     0.0875    4.63  3.8e-06 ***
## Year2011            0.4090     0.0837    4.89  1.1e-06 ***
## Year2012            0.4085     0.0895    4.56  5.4e-06 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.723
## Multiple R-squared:  0.174,  Adjusted R-squared:  0.167
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
##  166 weights are ~= 1. The remaining 2007 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.279   0.872   0.938   0.905   0.983   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          4.60e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Sample size for the above analysis:  2173"
## [1] ""
## [1] ""
## [1] "####################################"
## [1] "Analysis of AJSC 2914"
## [1] "####################################"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  179  234  193  174  188  191  139  148  148  152  125  145  211  289  316
## 2011 2012
##  342  351
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##   74   54   50   46   76   95  127  137  135  143  116  127  193  261  285
## 2011 2012
```

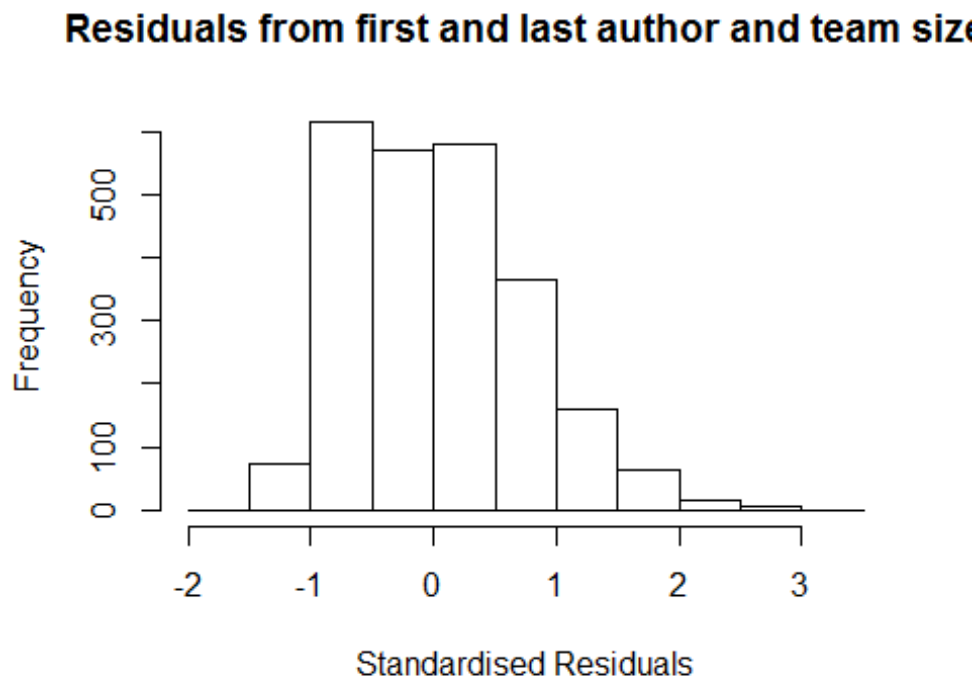
```
##  312  323
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##   72   50   47   44   73   91  124  131  130  139  110  121  187  252  272
## 2011 2012
##  294  311
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
##  Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 59, df = 16, p-value = 7e-07
```




```
##
##  Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 3, df = 1, p-value = 0.08
```



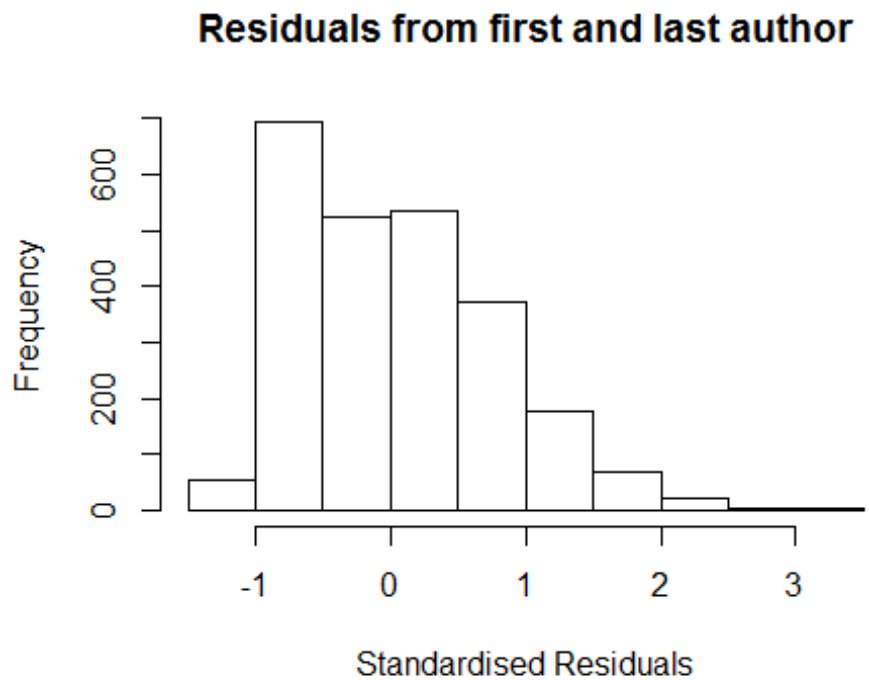

```
## [1] "Female first author team size 2018 geometric mean: 2.19743015937251"
## [1] "Male first author team size 2018 geometric mean: 3.07827392776149"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 3200, p-value = 0.003
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 2.22695095027763"
## [1] "Male last author team size 2018 geometric mean: 2.90393998539292"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 3600, p-value = 0.02
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 2.037  1           1.427
## LastAuthorFemale  2.081  1           1.443
## UniqueAuthors     1.176  4           1.021
## Year              1.258 16           1.007
```




```
## [1] "List of  6 outliers with residuals above  2.5"
##         ScopusId  NLCS Year OneField Fields residuals
## 2459 66049104209 3.107 2009     2914      1     2.584
## 2683 77957286112 3.677 2010     2914      1     2.989
## 2774 77954153066 3.658 2010     2914      1     2.970
## 2801 77954836425 3.894 2010     2914      1     2.710
## 3583 84859035757 3.186 2012     2914      1     2.537
## 3665 84855180516 3.698 2012     2914      1     3.002
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
UniqueAuthors +
##     Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##     k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.5701 -0.5238 -0.0238  0.4949  3.0023
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)          0.7013     0.1060    6.61  4.6e-11 ***
## FirstAuthorFemale1   0.1790     0.0467    3.84  0.00013 ***
## LastAuthorFemale1   -0.0243     0.0461   -0.53  0.59779
## UniqueAuthors2       0.2257     0.0399    5.66  1.7e-08 ***
## UniqueAuthors3       0.4750     0.0519    9.16  < 2e-16 ***
```

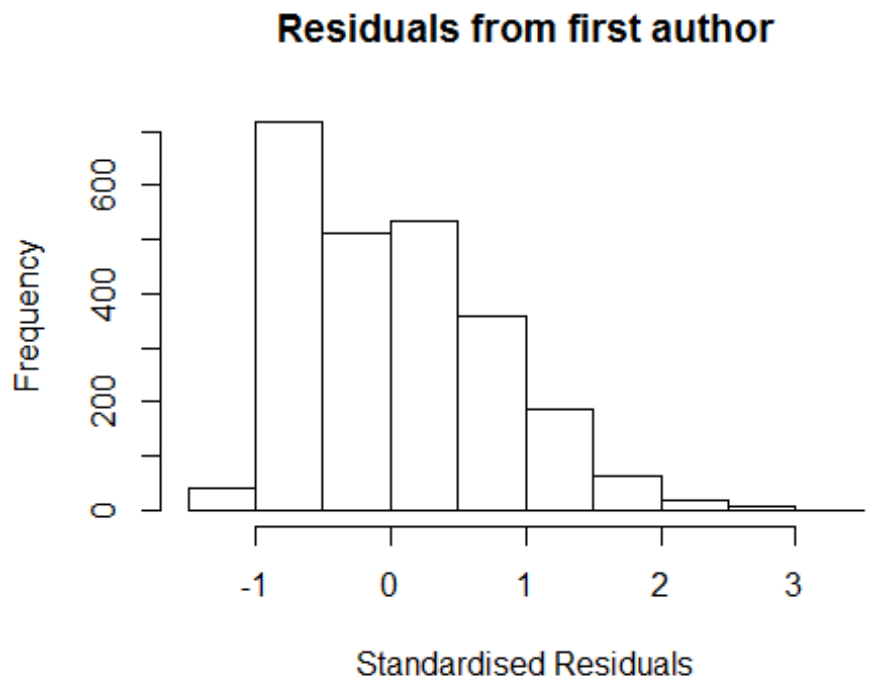
```
## UniqueAuthors4       0.5909     0.0710    8.33  < 2e-16 ***
## UniqueAuthors5       0.6507     0.0625   10.42  < 2e-16 ***
## Year1997             0.0517     0.1459    0.35  0.72335
## Year1998             0.0097     0.1532    0.06  0.94951
## Year1999             0.1233     0.1611    0.77  0.44429
## Year2000             0.0549     0.1328    0.41  0.67941
## Year2001            -0.0365     0.1237   -0.29  0.76832
## Year2002            -0.0397     0.1221   -0.33  0.74495
## Year2003            -0.1237     0.1182   -1.05  0.29508
## Year2004            -0.1329     0.1197   -1.11  0.26669
## Year2005            -0.2748     0.1147   -2.40  0.01664 *
## Year2006            -0.3120     0.1197   -2.61  0.00920 **
## Year2007            -0.3784     0.1164   -3.25  0.00117 **
## Year2008            -0.3071     0.1116   -2.75  0.00597 **
## Year2009            -0.3333     0.1078   -3.09  0.00202 **
## Year2010            -0.1676     0.1122   -1.49  0.13545
## Year2011            -0.1374     0.1111   -1.24  0.21601
## Year2012            -0.2070     0.1119   -1.85  0.06450 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.687
## Multiple R-squared:  0.138,  Adjusted R-squared:  0.131
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
##  184 weights are ~= 1. The remaining 2264 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0168  0.8850  0.9440  0.9030  0.9800  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          4.08e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 2.176  1           1.475
```

```
## LastAuthorFemale  2.177  1           1.476
## Year              1.077 16           1.002
```



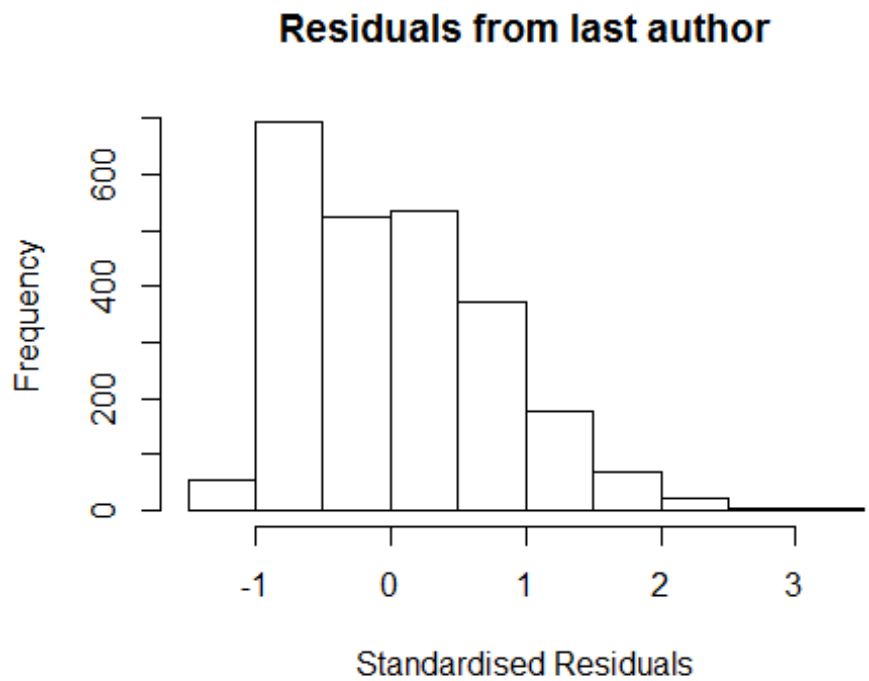

```
## [1] "List of  7 outliers with residuals above  2.5"
##         ScopusId  NLCS Year OneField Fields residuals
## 287   0031156020 3.820 1997     2914      1     2.689
## 1825 33748117943 3.310 2006     2914      1     2.641
## 2649 78049241602 3.677 2010     2914      1     2.901
## 2683 77957286112 3.677 2010     2914      1     2.819
## 2774 77954153066 3.658 2010     2914      1     2.800
## 2801 77954836425 3.894 2010     2914      1     3.118
## 3665 84855180516 3.698 2012     2914      1     3.089
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.1786 -0.6035 -0.0262  0.5422  3.1180
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)          0.9297     0.0983    9.46  < 2e-16 ***
## FirstAuthorFemale1   0.2183     0.0516    4.23  2.4e-05 ***
## LastAuthorFemale1   -0.1361     0.0501   -2.71  0.00671 **
```

```
## Year1997             0.1190     0.1386    0.86  0.39066
## Year1998             0.0306     0.1460    0.21  0.83425
## Year1999             0.1217     0.1540    0.79  0.42952
## Year2000             0.0811     0.1272    0.64  0.52364
## Year2001            -0.0629     0.1192   -0.53  0.59782
## Year2002            -0.0460     0.1154   -0.40  0.69003
## Year2003            -0.0610     0.1140   -0.54  0.59236
## Year2004            -0.1480     0.1129   -1.31  0.18980
## Year2005            -0.2598     0.1100   -2.36  0.01830 *
## Year2006            -0.3428     0.1152   -2.98  0.00296 **
## Year2007            -0.4128     0.1139   -3.62  0.00030 ***
## Year2008            -0.3262     0.1071   -3.05  0.00234 **
## Year2009            -0.3798     0.1043   -3.64  0.00028 ***
## Year2010            -0.1538     0.1097   -1.40  0.16113
## Year2011            -0.1026     0.1075   -0.95  0.33994
## Year2012            -0.1848     0.1088   -1.70  0.08971 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.755
## Multiple R-squared:  0.0447, Adjusted R-squared:  0.0376
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  173 weights are ~= 1. The remaining 2275 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0501  0.8860  0.9420  0.9110  0.9830  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          4.08e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.053  1           1.026
## Year              1.053 16           1.002
```




```
## [1] "List of  7 outliers with residuals above  2.5"
##         ScopusId  NLCS Year OneField Fields residuals
## 287   0031156020 3.820 1997     2914      1     2.689
## 1825 33748117943 3.310 2006     2914      1     2.641
## 2649 78049241602 3.677 2010     2914      1     2.901
## 2683 77957286112 3.677 2010     2914      1     2.819
## 2774 77954153066 3.658 2010     2914      1     2.800
## 2801 77954836425 3.894 2010     2914      1     3.118
## 3665 84855180516 3.698 2012     2914      1     3.089
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##    Min     1Q Median     3Q    Max
## -1.149 -0.589 -0.034  0.537  3.144
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)           0.9007     0.0983    9.16  < 2e-16 ***
## FirstAuthorFemale1    0.1220     0.0355    3.44  0.00059 ***
## Year1997              0.1266     0.1390    0.91  0.36252
## Year1998              0.0352     0.1463    0.24  0.80971
## Year1999              0.1226     0.1536    0.80  0.42481
## Year2000              0.0859     0.1282    0.67  0.50291
```

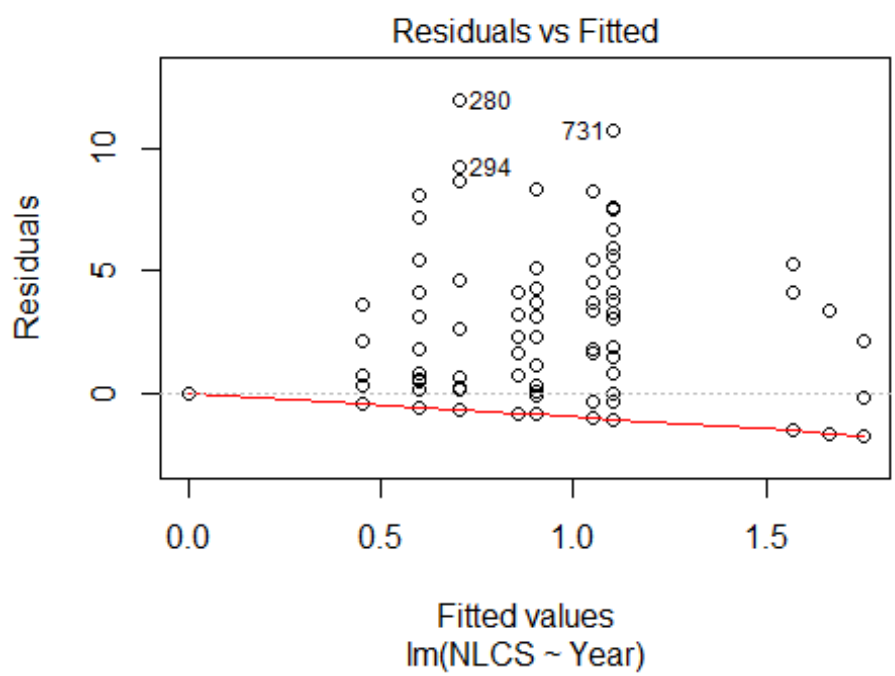
```
## Year2001            -0.0620     0.1195   -0.52  0.60388
## Year2002            -0.0441     0.1155   -0.38  0.70264
## Year2003            -0.0588     0.1142   -0.52  0.60641
## Year2004            -0.1436     0.1136   -1.26  0.20615
## Year2005            -0.2509     0.1108   -2.27  0.02360 *
## Year2006            -0.3318     0.1160   -2.86  0.00429 **
## Year2007            -0.3998     0.1142   -3.50  0.00047 ***
## Year2008            -0.3220     0.1076   -2.99  0.00281 **
## Year2009            -0.3827     0.1048   -3.65  0.00027 ***
## Year2010            -0.1506     0.1103   -1.37  0.17204
## Year2011            -0.0982     0.1080   -0.91  0.36314
## Year2012            -0.1861     0.1094   -1.70  0.08891 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.755
## Multiple R-squared:  0.0412, Adjusted R-squared:  0.0345
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  178 weights are ~= 1. The remaining 2270 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0443  0.8820  0.9390  0.9110  0.9820  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          4.08e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##                   GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.049  1           1.024
## Year             1.049 16           1.001
```




```
## [1] "List of  7 outliers with residuals above  2.5"
##         ScopusId  NLCS Year OneField Fields residuals
## 287   0031156020 3.820 1997     2914      1     2.689
## 1825 33748117943 3.310 2006     2914      1     2.641
## 2649 78049241602 3.677 2010     2914      1     2.901
## 2683 77957286112 3.677 2010     2914      1     2.819
## 2774 77954153066 3.658 2010     2914      1     2.800
## 2801 77954836425 3.894 2010     2914      1     3.118
## 3665 84855180516 3.698 2012     2914      1     3.089
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.1290 -0.6133 -0.0285  0.5332  3.0560
##
## Coefficients:
##                   Estimate Std. Error t value Pr(>|t|)
## (Intercept)        0.98719    0.09769   10.11  < 2e-16 ***
## LastAuthorFemale1  0.00637    0.03588    0.18  0.85919
## Year1997           0.13545    0.13969    0.97  0.33231
## Year1998           0.03069    0.14584    0.21  0.83335
## Year1999           0.13060    0.15351    0.85  0.39497
## Year2000           0.08245    0.12725    0.65  0.51709
```

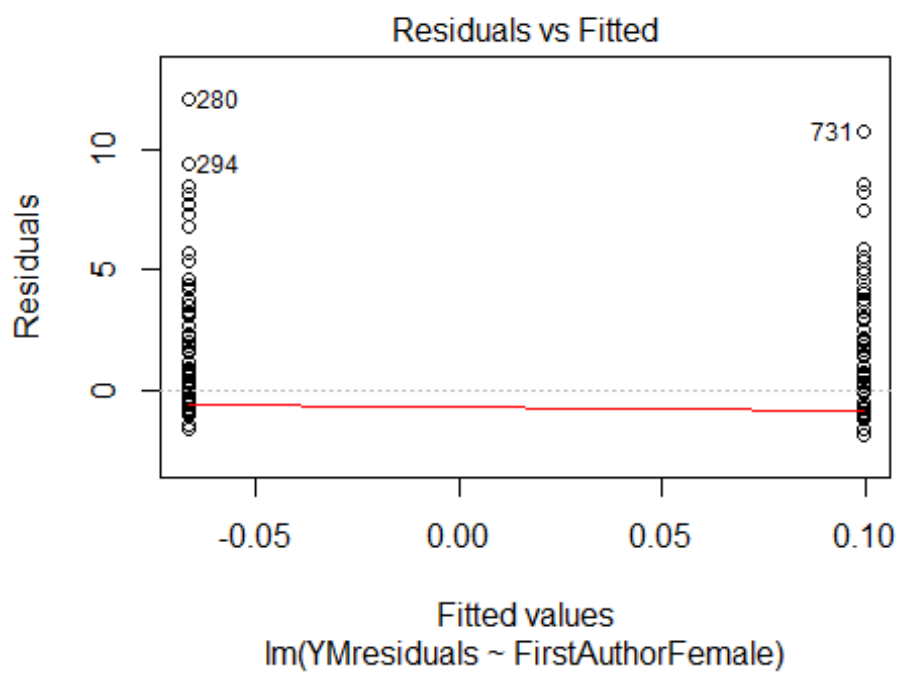
```
## Year2001          -0.05168    0.11853   -0.44  0.66286
## Year2002          -0.04122    0.11450   -0.36  0.71891
## Year2003          -0.05307    0.11334   -0.47  0.63967
## Year2004          -0.14012    0.11266   -1.24  0.21368
## Year2005          -0.24771    0.10975   -2.26  0.02409 *
## Year2006          -0.33835    0.11658   -2.90  0.00374 **
## Year2007          -0.41718    0.11479   -3.63  0.00028 ***
## Year2008          -0.32943    0.10714   -3.07  0.00213 **
## Year2009          -0.38026    0.10404   -3.65  0.00026 ***
## Year2010          -0.14919    0.10959   -1.36  0.17351
## Year2011          -0.09295    0.10706   -0.87  0.38538
## Year2012          -0.17850    0.10855   -1.64  0.10020
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.758
## Multiple R-squared:  0.0367, Adjusted R-squared:  0.03
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  161 weights are ~= 1. The remaining 2287 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0674  0.8880  0.9410  0.9120  0.9830  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          4.08e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Sample size for the above analysis:  2448"
## [1] ""
## [1] ""
## [1] "####################################"
## [1] "Analysis of AJSC 2915"
## [1] "####################################"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1997 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011 2012
##    9   16    5   16    8    5    5   34   49   69   92  119  125   98  123
```

```
##
## 1997 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011 2012
##    8    6    0    8    7    4    3   29   45   62   84  112  116   97  109
##
## 1997 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011 2012
##    8    6    0    7    7    4    3   28   45   61   80  107  115   94  106
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
##  Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = Inf, df = 13, p-value <2e-16
```




```
##
##  Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 5.9, df = 1, p-value = 0.02

## [1] "Female first author team size 2018 geometric mean: 1.92505921911844"
## [1] "Male first author team size 2018 geometric mean: 1.96372648541743"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 880, p-value = 0.9
```

```
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 1.76479392053783"
## [1] "Male last author team size 2018 geometric mean: 2.27763908870955"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 780, p-value = 0.02
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"

## Warning in lmrob.S(x, y, control = control, mf = mf): S-estimated scale ==
## 0: Probably exact fit; check your data

## Warning in cov2cor(v): diag(.) had 0 or NA entries; non-finite result is
## doubtful
```



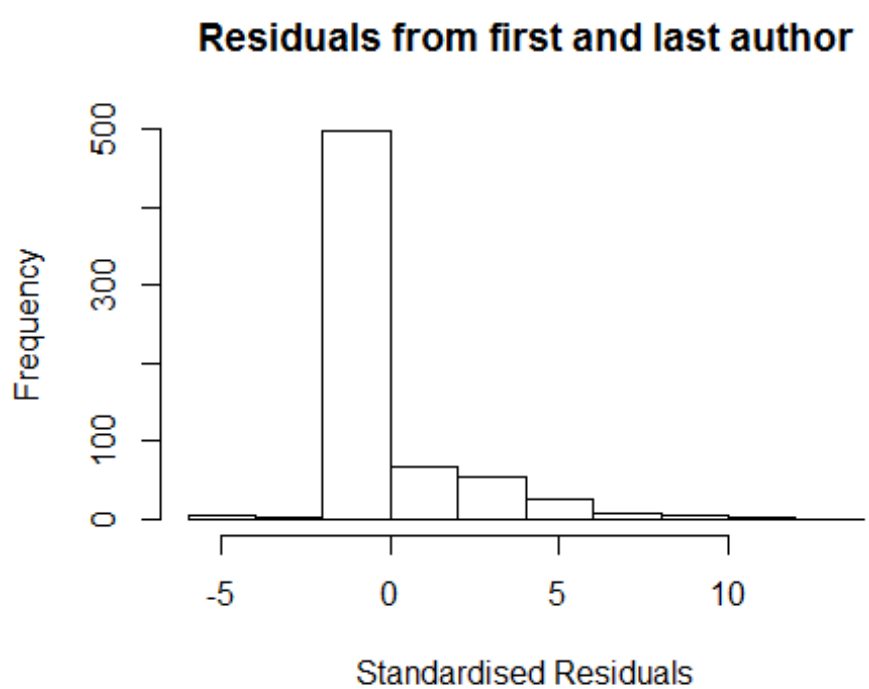

```
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale   NaN  1             NaN
## LastAuthorFemale    NaN  1             NaN
## Year                NaN 13             NaN
```

```
## [1] "List of  97 outliers with residuals above  2.5"
##        ScopusId   NLCS Year OneField Fields residuals
## 1   53249095953  0.000 1997     2915      1    -5.702
## 2   53249095954  0.000 1997     2915      1    -5.702
## 3   53249098867  0.000 1997     2915      1    -5.702
## 5   53249117643  0.000 1997     2915      1    -5.702
## 6   53249121981  0.000 1997     2915      1    -5.702
## 8   53249125913  0.000 1997     2915      1    -5.702
## 47   0036358592  4.726 2002     2915      1     4.726
## 51   0036691206  2.637 2002     2915      1     2.637
## 59   1442359636  0.000 2003     2915      1    -3.870
## 62  16644377794  5.000 2004     2915      1     5.000
## 71  19644380535  4.957 2005     2915      1     4.957
## 72  19644386919  4.042 2005     2915      1     4.042
## 110 33645635094  8.722 2006     2915      1     6.820
## 117 33745267080  4.917 2006     2915      1     3.015
## 125 33746537334  4.417 2006     2915      1     2.515
## 141 37349063360  7.039 2006     2915      1     5.137
## 150 38449116798  2.809 2007     2915      1     2.809
## 157 33847639651  2.809 2007     2915      1     2.809
## 162 33947516595  2.809 2007     2915      1     2.809
## 164 34047195453  5.619 2007     2915      1     5.619
## 167 34047223899  4.453 2007     2915      1     4.453
## 169 34047223902  6.523 2007     2915      1     6.523
## 170 34047229521  2.809 2007     2915      1     2.809
## 172 34247854872  2.809 2007     2915      1     2.809
## 175 34447506593  5.619 2007     2915      1     5.619
## 185 35248818866  2.809 2007     2915      1     2.809
## 188 35248836604  2.809 2007     2915      1     2.809
## 192 36048931651  4.453 2007     2915      1     4.453
## 211 38449105733  4.453 2007     2915      1     4.453
## 213 38449113755  4.453 2007     2915      1     4.453
## 255 39749091154  9.372 2008     2915      1     9.372
## 266 43049134885  5.291 2008     2915      1     5.291
## 270 43849113261  3.338 2008     2915      1     3.338
## 271 44449085442  3.338 2008     2915      1     3.338
## 279 47249152188  3.338 2008     2915      1     3.338
## 280 51249085065 12.710 2008     2915      1    12.710
## 291 54449098418  3.338 2008     2915      1     3.338
## 294 55449135631 10.015 2008     2915      1    10.015
## 310 77954378540  3.161 2009     2915      1     3.161
## 311 80051915932  3.161 2009     2915      1     3.161
## 333 67649282818  5.155 2009     2915      1     5.155
## 357 63349105104  3.988 2009     2915      1     3.988
## 365 65349134659  3.988 2009     2915      1     3.988
## 369 65349168323  3.161 2009     2915      1     3.161
## 373 68049097133  5.155 2009     2915      1     5.155
## 384 70349305314  3.161 2009     2915      1     3.161
## 386 70349322617  9.260 2009     2915      1     9.260
## 388 70349332693  3.988 2009     2915      1     3.988
```

```
## 390 70350455625  3.161 2009     2915      1     3.161
## 394 70350490568  4.630 2009     2915      1     4.630
## 437 78649361325  3.725 2010     2915      1     3.725
## 471 77955497686  3.725 2010     2915      1     3.725
## 488 77951919793  3.725 2010     2915      1     3.725
## 493 77951215761  8.698 2010     2915      1     8.698
## 501 77951883752  3.725 2010     2915      1     3.725
## 528 77956284940  6.076 2010     2915      1     6.076
## 556 79952110823  3.725 2010     2915      1     3.725
## 563 79952111770  7.808 2010     2915      1     7.808
## 575 79952999731  4.701 2010     2915      1     4.701
## 610 80155186701  2.596 2011     2915      1     2.596
## 617 79960070053  2.596 2011     2915      1     2.596
## 629 79954434291  2.596 2011     2915      1     2.596
## 635 79952271257  2.596 2011     2915      1     2.596
## 641 79954450264  2.596 2011     2915      1     2.596
## 658 79960063998  2.596 2011     2915      1     2.596
## 666 79960956243  2.596 2011     2915      1     2.596
## 668 80053308320  2.596 2011     2915      1     2.596
## 682 80053531269  2.596 2011     2915      1     2.596
## 698 84855167539  2.596 2011     2915      1     2.596
## 703 84855179061  4.115 2011     2915      1     4.115
## 707 84856144361  4.115 2011     2915      1     4.115
## 709 84856161379  2.596 2011     2915      1     2.596
## 731 84872593614 11.903 2012     2915      1    11.903
## 758 84864116706  2.596 2012     2915      1     2.596
## 759 84863206357  2.596 2012     2915      1     2.596
## 761 84863219078  4.115 2012     2915      1     4.115
## 772 84860786544  2.596 2012     2915      1     2.596
## 773 84861143519  6.711 2012     2915      1     6.711
## 783 84859378234  6.028 2012     2915      1     6.028
## 786 84859387819  2.596 2012     2915      1     2.596
## 788 84859408406  2.596 2012     2915      1     2.596
## 789 84859409650  4.115 2012     2915      1     4.115
## 790 84859410305  4.115 2012     2915      1     4.115
## 799 84860813506  2.596 2012     2915      1     2.596
## 800 84860819544  8.624 2012     2915      1     8.624
## 807 84861137684  4.115 2012     2915      1     4.115
## 816 84864118197  2.596 2012     2915      1     2.596
## 817 84864124334  6.711 2012     2915      1     6.711
## 819 84864146277  4.115 2012     2915      1     4.115
## 824 84866454565  4.115 2012     2915      1     4.115
## 825 84866458027  5.192 2012     2915      1     5.192
## 826 84866458789  2.596 2012     2915      1     2.596
## 828 84867034981  2.596 2012     2915      1     2.596
## 831 84867077526  2.596 2012     2915      1     2.596
## 839 84871890823  7.788 2012     2915      1     7.788
## 840 84871900701  2.596 2012     2915      1     2.596
## 842 84871915967  4.115 2012     2915      1     4.115
##
```

```
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "S"
## Residuals:
##    Min     1Q Median     3Q    Max
##   -5.7    0.0    0.0    0.0   12.7
##
## Exact fit detected
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)            5.70       0.00      NA       NA
## FirstAuthorFemale1     0.00       0.00      NA       NA
## LastAuthorFemale1      0.00       0.00      NA       NA
## Year1999              -5.70       0.00      NA       NA
## Year2001              -5.70       0.00      NA       NA
## Year2002              -5.70       0.00      NA       NA
## Year2003              -1.83       0.00      NA       NA
## Year2004              -5.70       0.00      NA       NA
## Year2005              -5.70       0.00      NA       NA
## Year2006              -3.80       0.00      NA       NA
## Year2007              -5.70       0.00      NA       NA
## Year2008              -5.70       0.00      NA       NA
## Year2009              -5.70       0.00      NA       NA
## Year2010              -5.70       0.00      NA       NA
## Year2011              -5.70       0.00      NA       NA
## Year2012              -5.70       0.00      NA       NA
##
## Robustness weights:
##  204 observations
c(1,2,3,4,5,6,7,22,25,29,30,32,34,38,40,41,47,52,55,58,62,63,64,65,66,67,68,6
9,70,71,72,73,74,75,76,77,78,79,80,81,82,83,84,85,86,87,88,89,91,92,94,95,96,
98,100,101,102,103,104,105,106,107,110,117,122,124,127,129,130,132,134,144,14
6,148,164,166,169,181,183,192,193,199,209,213,214,221,222,231,234,249,250,251
,253,257,261,266,269,270,275,277,278,282,291,293,294,295,296,300,302,303,310,
312,314,315,317,321,329,334,336,338,346,347,349,360,372,373,375,379,384,387,3
90,393,396,397,400,401,402,403,404,407,409,414,422,434,448,449,450,454,461,46
6,468,471,475,479,480,483,486,491,496,501,506,509,513,526,530,532,539,547,550
,552,554,563,565,568,581,585,586,588,589,591,596,600,601,602,611,613,614,615,
616,622,623,628,635,636,638,643,644,645,646,649,654,655,656,667)
##   are outliers with |weight| = 0 ( < 0.00015);
##  467 weights are ~= 1.
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.49e-04          1.82e-12
## warn.limit.reject warn.limit.meanrw
```

```
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"

## Warning in lmrob.S(x, y, control = control, mf = mf): S-estimated scale ==
## 0: Probably exact fit; check your data

## Warning in lmrob.S(x, y, control = control, mf = mf): diag(.) had 0 or NA
## entries; non-finite result is doubtful
```



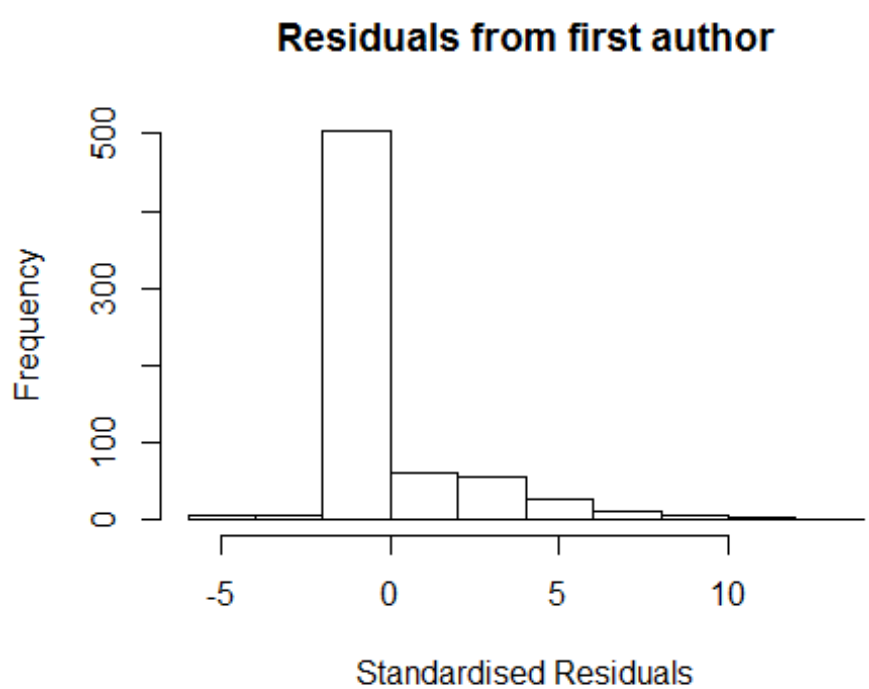

```
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale  NaN  1             NaN
## Year               NaN 13             NaN

## [1] "List of  97 outliers with residuals above  2.5"
##        ScopusId   NLCS Year OneField Fields residuals
## 1   53249095953  0.000 1997     2915      1    -5.702
## 2   53249095954  0.000 1997     2915      1    -5.702
## 3   53249098867  0.000 1997     2915      1    -5.702
```

```
## 5   53249117643  0.000 1997     2915      1    -5.702
## 6   53249121981  0.000 1997     2915      1    -5.702
## 8   53249125913  0.000 1997     2915      1    -5.702
## 47   0036358592  4.726 2002     2915      1     4.726
## 51   0036691206  2.637 2002     2915      1     2.637
## 59   1442359636  0.000 2003     2915      1    -3.870
## 62  16644377794  5.000 2004     2915      1     5.000
## 71  19644380535  4.957 2005     2915      1     4.957
## 72  19644386919  4.042 2005     2915      1     4.042
## 110 33645635094  8.722 2006     2915      1     6.820
## 117 33745267080  4.917 2006     2915      1     3.015
## 125 33746537334  4.417 2006     2915      1     2.515
## 141 37349063360  7.039 2006     2915      1     5.137
## 150 38449116798  2.809 2007     2915      1     2.809
## 157 33847639651  2.809 2007     2915      1     2.809
## 162 33947516595  2.809 2007     2915      1     2.809
## 164 34047195453  5.619 2007     2915      1     5.619
## 167 34047223899  4.453 2007     2915      1     4.453
## 169 34047223902  6.523 2007     2915      1     6.523
## 170 34047229521  2.809 2007     2915      1     2.809
## 172 34247854872  2.809 2007     2915      1     2.809
## 175 34447506593  5.619 2007     2915      1     5.619
## 185 35248818866  2.809 2007     2915      1     2.809
## 188 35248836604  2.809 2007     2915      1     2.809
## 192 36048931651  4.453 2007     2915      1     4.453
## 211 38449105733  4.453 2007     2915      1     4.453
## 213 38449113755  4.453 2007     2915      1     4.453
## 255 39749091154  9.372 2008     2915      1     9.372
## 266 43049134885  5.291 2008     2915      1     5.291
## 270 43849113261  3.338 2008     2915      1     3.338
## 271 44449085442  3.338 2008     2915      1     3.338
## 279 47249152188  3.338 2008     2915      1     3.338
## 280 51249085065 12.710 2008     2915      1    12.710
## 291 54449098418  3.338 2008     2915      1     3.338
## 294 55449135631 10.015 2008     2915      1    10.015
## 310 77954378540  3.161 2009     2915      1     3.161
## 311 80051915932  3.161 2009     2915      1     3.161
## 333 67649282818  5.155 2009     2915      1     5.155
## 357 63349105104  3.988 2009     2915      1     3.988
## 365 65349134659  3.988 2009     2915      1     3.988
## 369 65349168323  3.161 2009     2915      1     3.161
## 373 68049097133  5.155 2009     2915      1     5.155
## 384 70349305314  3.161 2009     2915      1     3.161
## 386 70349322617  9.260 2009     2915      1     9.260
## 388 70349332693  3.988 2009     2915      1     3.988
## 390 70350455625  3.161 2009     2915      1     3.161
## 394 70350490568  4.630 2009     2915      1     4.630
## 437 78649361325  3.725 2010     2915      1     3.725
## 471 77955497686  3.725 2010     2915      1     3.725
## 488 77951919793  3.725 2010     2915      1     3.725
```

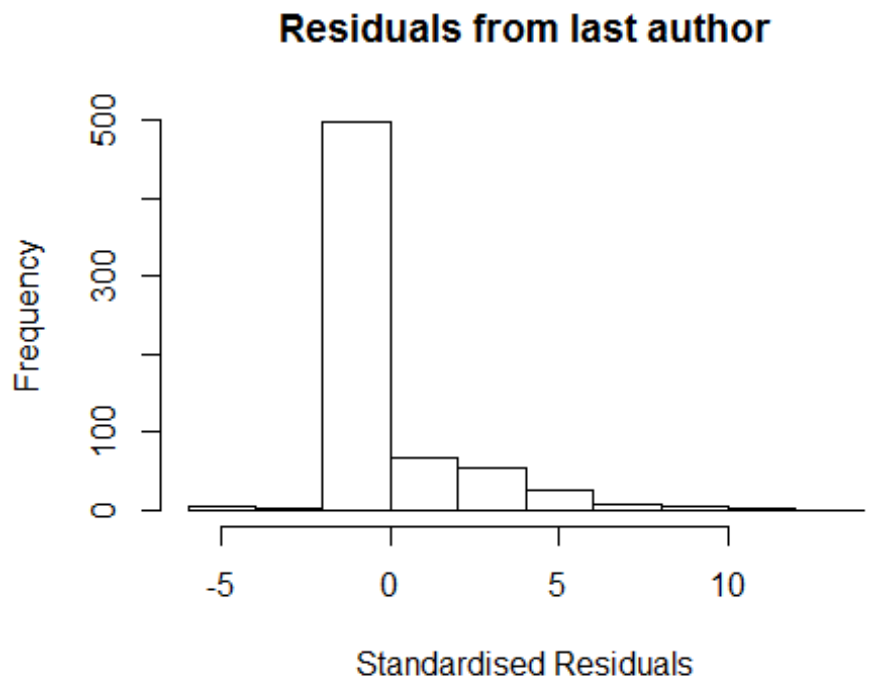
```
## 493 77951215761  8.698 2010     2915      1     8.698
## 501 77951883752  3.725 2010     2915      1     3.725
## 528 77956284940  6.076 2010     2915      1     6.076
## 556 79952110823  3.725 2010     2915      1     3.725
## 563 79952111770  7.808 2010     2915      1     7.808
## 575 79952999731  4.701 2010     2915      1     4.701
## 610 80155186701  2.596 2011     2915      1     2.596
## 617 79960070053  2.596 2011     2915      1     2.596
## 629 79954434291  2.596 2011     2915      1     2.596
## 635 79952271257  2.596 2011     2915      1     2.596
## 641 79954450264  2.596 2011     2915      1     2.596
## 658 79960063998  2.596 2011     2915      1     2.596
## 666 79960956243  2.596 2011     2915      1     2.596
## 668 80053308320  2.596 2011     2915      1     2.596
## 682 80053531269  2.596 2011     2915      1     2.596
## 698 84855167539  2.596 2011     2915      1     2.596
## 703 84855179061  4.115 2011     2915      1     4.115
## 707 84856144361  4.115 2011     2915      1     4.115
## 709 84856161379  2.596 2011     2915      1     2.596
## 731 84872593614 11.903 2012     2915      1    11.903
## 758 84864116706  2.596 2012     2915      1     2.596
## 759 84863206357  2.596 2012     2915      1     2.596
## 761 84863219078  4.115 2012     2915      1     4.115
## 772 84860786544  2.596 2012     2915      1     2.596
## 773 84861143519  6.711 2012     2915      1     6.711
## 783 84859378234  6.028 2012     2915      1     6.028
## 786 84859387819  2.596 2012     2915      1     2.596
## 788 84859408406  2.596 2012     2915      1     2.596
## 789 84859409650  4.115 2012     2915      1     4.115
## 790 84859410305  4.115 2012     2915      1     4.115
## 799 84860813506  2.596 2012     2915      1     2.596
## 800 84860819544  8.624 2012     2915      1     8.624
## 807 84861137684  4.115 2012     2915      1     4.115
## 816 84864118197  2.596 2012     2915      1     2.596
## 817 84864124334  6.711 2012     2915      1     6.711
## 819 84864146277  4.115 2012     2915      1     4.115
## 824 84866454565  4.115 2012     2915      1     4.115
## 825 84866458027  5.192 2012     2915      1     5.192
## 826 84866458789  2.596 2012     2915      1     2.596
## 828 84867034981  2.596 2012     2915      1     2.596
## 831 84867077526  2.596 2012     2915      1     2.596
## 839 84871890823  7.788 2012     2915      1     7.788
## 840 84871900701  2.596 2012     2915      1     2.596
## 842 84871915967  4.115 2012     2915      1     4.115
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "S"
```

```
## Residuals:
##    Min     1Q Median     3Q    Max
##  -4.73   0.00   0.00   0.00  12.71
##
## Exact fit detected
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)            0.000      0.000      NA       NA
## FirstAuthorFemale1     0.000      0.000      NA       NA
## Year1999               0.000      0.000      NA       NA
## Year2001               0.000      0.000      NA       NA
## Year2002               4.726      0.000      NA       NA
## Year2003               3.870      0.000      NA       NA
## Year2004               0.000      0.000      NA       NA
## Year2005               0.000      0.000      NA       NA
## Year2006               0.000      0.000      NA       NA
## Year2007               0.000      0.000      NA       NA
## Year2008               0.000      0.000      NA       NA
## Year2009               0.000      0.000      NA       NA
## Year2010               0.707      0.000      NA       NA
## Year2011               0.000      0.000      NA       NA
## Year2012               0.000      0.000      NA       NA
##
## Robustness weights:
##  253 observations
c(4,8,23,24,25,26,27,28,29,30,32,34,38,40,41,47,52,55,58,62,63,64,75,79,87,89
,90,92,93,97,99,100,102,105,108,110,117,122,124,127,129,130,132,134,144,146,1
48,164,166,169,181,183,192,193,199,209,213,214,221,222,231,234,249,250,251,25
3,257,261,266,269,270,275,277,278,282,291,293,294,295,296,300,302,303,310,312
,314,315,317,321,329,334,336,338,346,347,349,357,358,359,360,361,362,363,364,
365,366,367,368,369,370,371,373,374,376,377,378,379,380,381,382,383,385,386,3
88,389,390,391,392,393,394,395,398,399,400,402,403,404,405,406,407,408,409,41
0,411,412,413,414,415,416,417,418,419,420,421,422,423,424,425,426,427,428,429
,430,431,432,433,434,435,436,437,438,439,440,441,442,443,444,445,446,447,448,
449,450,451,452,453,454,455,456,457,458,459,460,461,462,463,464,465,467,469,4
70,475,479,480,483,486,491,496,501,506,509,513,526,530,532,539,547,550,552,55
4,563,565,568,581,585,586,588,589,591,596,600,601,602,611,613,614,615,616,622
,623,628,635,636,638,643,644,645,646,649,654,655,656,667)
##   are outliers with |weight| = 0 ( < 0.00015);
##  418 weights are ~= 1.
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.49e-04          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
```

```
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                    cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"

## Warning in lmrob.S(x, y, control = control, mf = mf): find_scale() did not
## converge in 'maxit.scale' (= 200) iterations

## Warning in lmrob.S(x, y, control = control, mf = mf): S-estimated scale ==
## 0: Probably exact fit; check your data

## Warning in cov2cor(v): diag(.) had 0 or NA entries; non-finite result is
## doubtful
```




```
##                   GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale   NaN  1             NaN
## Year               NaN 13             NaN
```



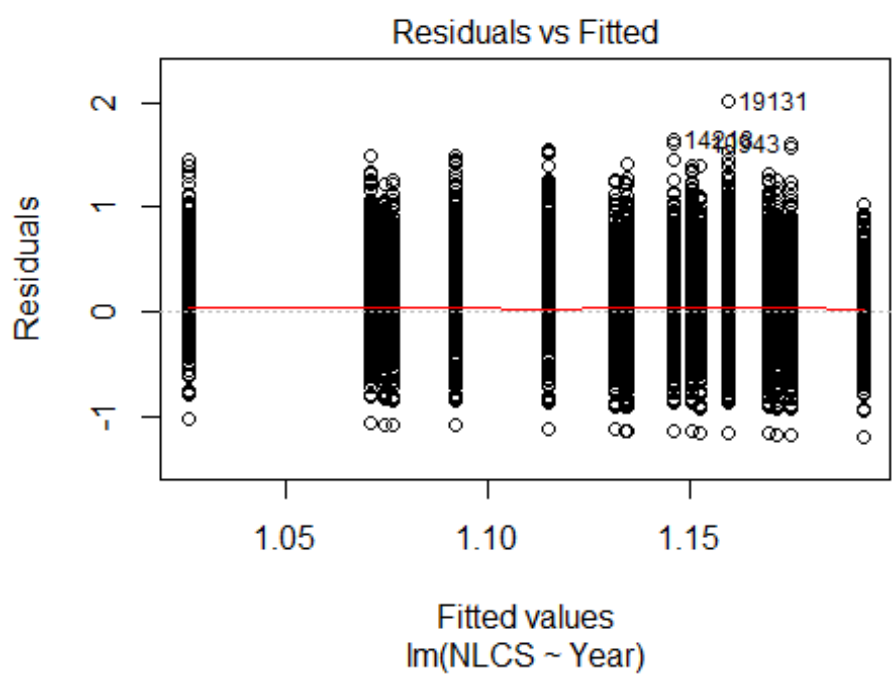

```
## [1] "List of  97 outliers with residuals above  2.5"
##        ScopusId   NLCS Year OneField Fields residuals
## 1   53249095953  0.000 1997     2915      1    -5.702
## 2   53249095954  0.000 1997     2915      1    -5.702
## 3   53249098867  0.000 1997     2915      1    -5.702
## 5   53249117643  0.000 1997     2915      1    -5.702
## 6   53249121981  0.000 1997     2915      1    -5.702
## 8   53249125913  0.000 1997     2915      1    -5.702
## 47   0036358592  4.726 2002     2915      1     4.726
## 51   0036691206  2.637 2002     2915      1     2.637
## 59   1442359636  0.000 2003     2915      1    -3.870
## 62  16644377794  5.000 2004     2915      1     5.000
## 71  19644380535  4.957 2005     2915      1     4.957
## 72  19644386919  4.042 2005     2915      1     4.042
## 110 33645635094  8.722 2006     2915      1     6.820
## 117 33745267080  4.917 2006     2915      1     3.015
## 125 33746537334  4.417 2006     2915      1     2.515
## 141 37349063360  7.039 2006     2915      1     5.137
## 150 38449116798  2.809 2007     2915      1     2.809
## 157 33847639651  2.809 2007     2915      1     2.809
## 162 33947516595  2.809 2007     2915      1     2.809
## 164 34047195453  5.619 2007     2915      1     5.619
## 167 34047223899  4.453 2007     2915      1     4.453
## 169 34047223902  6.523 2007     2915      1     6.523
## 170 34047229521  2.809 2007     2915      1     2.809
## 172 34247854872  2.809 2007     2915      1     2.809
## 175 34447506593  5.619 2007     2915      1     5.619
```

```
## 185 35248818866  2.809 2007     2915      1     2.809
## 188 35248836604  2.809 2007     2915      1     2.809
## 192 36048931651  4.453 2007     2915      1     4.453
## 211 38449105733  4.453 2007     2915      1     4.453
## 213 38449113755  4.453 2007     2915      1     4.453
## 255 39749091154  9.372 2008     2915      1     9.372
## 266 43049134885  5.291 2008     2915      1     5.291
## 270 43849113261  3.338 2008     2915      1     3.338
## 271 44449085442  3.338 2008     2915      1     3.338
## 279 47249152188  3.338 2008     2915      1     3.338
## 280 51249085065 12.710 2008     2915      1    12.710
## 291 54449098418  3.338 2008     2915      1     3.338
## 294 55449135631 10.015 2008     2915      1    10.015
## 310 77954378540  3.161 2009     2915      1     3.161
## 311 80051915932  3.161 2009     2915      1     3.161
## 333 67649282818  5.155 2009     2915      1     5.155
## 357 63349105104  3.988 2009     2915      1     3.988
## 365 65349134659  3.988 2009     2915      1     3.988
## 369 65349168323  3.161 2009     2915      1     3.161
## 373 68049097133  5.155 2009     2915      1     5.155
## 384 70349305314  3.161 2009     2915      1     3.161
## 386 70349322617  9.260 2009     2915      1     9.260
## 388 70349332693  3.988 2009     2915      1     3.988
## 390 70350455625  3.161 2009     2915      1     3.161
## 394 70350490568  4.630 2009     2915      1     4.630
## 437 78649361325  3.725 2010     2915      1     3.725
## 471 77955497686  3.725 2010     2915      1     3.725
## 488 77951919793  3.725 2010     2915      1     3.725
## 493 77951215761  8.698 2010     2915      1     8.698
## 501 77951883752  3.725 2010     2915      1     3.725
## 528 77956284940  6.076 2010     2915      1     6.076
## 556 79952110823  3.725 2010     2915      1     3.725
## 563 79952111770  7.808 2010     2915      1     7.808
## 575 79952999731  4.701 2010     2915      1     4.701
## 610 80155186701  2.596 2011     2915      1     2.596
## 617 79960070053  2.596 2011     2915      1     2.596
## 629 79954434291  2.596 2011     2915      1     2.596
## 635 79952271257  2.596 2011     2915      1     2.596
## 641 79954450264  2.596 2011     2915      1     2.596
## 658 79960063998  2.596 2011     2915      1     2.596
## 666 79960956243  2.596 2011     2915      1     2.596
## 668 80053308320  2.596 2011     2915      1     2.596
## 682 80053531269  2.596 2011     2915      1     2.596
## 698 84855167539  2.596 2011     2915      1     2.596
## 703 84855179061  4.115 2011     2915      1     4.115
## 707 84856144361  4.115 2011     2915      1     4.115
## 709 84856161379  2.596 2011     2915      1     2.596
## 731 84872593614 11.903 2012     2915      1    11.903
## 758 84864116706  2.596 2012     2915      1     2.596
## 759 84863206357  2.596 2012     2915      1     2.596
```

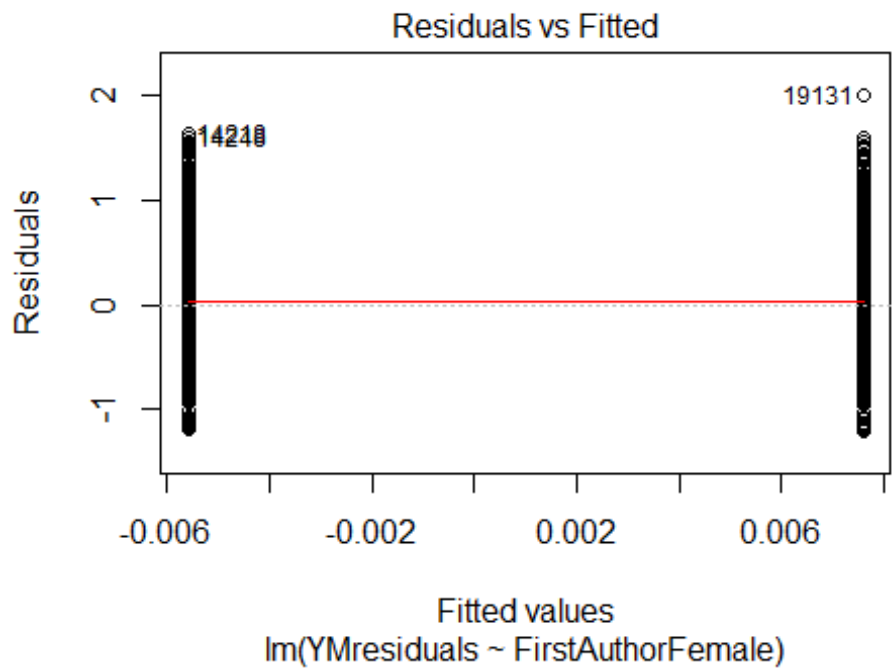
```
## 761 84863219078  4.115 2012     2915      1     4.115
## 772 84860786544  2.596 2012     2915      1     2.596
## 773 84861143519  6.711 2012     2915      1     6.711
## 783 84859378234  6.028 2012     2915      1     6.028
## 786 84859387819  2.596 2012     2915      1     2.596
## 788 84859408406  2.596 2012     2915      1     2.596
## 789 84859409650  4.115 2012     2915      1     4.115
## 790 84859410305  4.115 2012     2915      1     4.115
## 799 84860813506  2.596 2012     2915      1     2.596
## 800 84860819544  8.624 2012     2915      1     8.624
## 807 84861137684  4.115 2012     2915      1     4.115
## 816 84864118197  2.596 2012     2915      1     2.596
## 817 84864124334  6.711 2012     2915      1     6.711
## 819 84864146277  4.115 2012     2915      1     4.115
## 824 84866454565  4.115 2012     2915      1     4.115
## 825 84866458027  5.192 2012     2915      1     5.192
## 826 84866458789  2.596 2012     2915      1     2.596
## 828 84867034981  2.596 2012     2915      1     2.596
## 831 84867077526  2.596 2012     2915      1     2.596
## 839 84871890823  7.788 2012     2915      1     7.788
## 840 84871900701  2.596 2012     2915      1     2.596
## 842 84871915967  4.115 2012     2915      1     4.115
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "S"
## Residuals:
##    Min     1Q Median     3Q    Max
##   -5.0    0.0    0.0    0.0   11.9
##
## Exact fit detected
##
## Coefficients:
##                   Estimate Std. Error t value Pr(>|t|)
## (Intercept)           0.00       0.00      NA       NA
## LastAuthorFemale1     0.00       0.00      NA       NA
## Year1999              0.00       0.00      NA       NA
## Year2001              0.00       0.00      NA       NA
## Year2002              4.73       0.00      NA       NA
## Year2003              1.57       0.00      NA       NA
## Year2004              5.00       0.00      NA       NA
## Year2005              0.00       0.00      NA       NA
## Year2006              4.92       0.00      NA       NA
## Year2007              0.00       0.00      NA       NA
## Year2008              1.30       0.00      NA       NA
## Year2009              0.00       0.00      NA       NA
## Year2010              0.00       0.00      NA       NA
## Year2011              0.00       0.00      NA       NA
```

```
## Year2012              0.00       0.00      NA       NA
##
## Robustness weights:
##  272 observations
c(4,8,23,24,25,26,27,28,31,32,33,35,38,40,41,47,52,55,58,62,63,64,65,66,67,68
,69,70,71,72,73,74,75,76,77,78,80,81,82,83,84,85,86,87,88,89,90,91,92,93,94,9
5,96,97,98,99,100,101,102,103,104,105,106,107,108,110,117,122,124,127,129,130
,132,134,144,146,148,164,166,169,170,171,172,173,174,175,176,177,178,179,180,
182,183,184,185,186,187,188,189,190,191,192,193,194,195,196,197,198,199,200,2
01,202,203,204,205,206,207,208,209,210,211,212,213,214,215,216,217,218,219,22
0,221,222,223,224,225,226,227,228,229,230,231,232,233,234,235,236,237,238,239
,240,241,242,243,244,245,246,247,248,250,251,253,257,261,266,269,270,275,277,
278,282,291,293,294,295,296,300,302,303,310,312,314,315,317,321,329,334,336,3
38,346,347,349,360,372,373,375,379,384,387,390,393,396,397,400,401,402,403,40
4,407,409,414,422,434,448,449,450,454,461,466,468,471,475,479,480,483,486,491
,496,501,506,509,513,526,530,532,539,547,550,552,554,563,565,568,581,585,586,
588,589,591,596,600,601,602,611,613,614,615,616,622,623,628,635,636,638,643,6
44,645,646,649,654,655,656,667)
##   are outliers with |weight| = 0 ( < 0.00015);
##  399 weights are ~= 1.
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.49e-04          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Sample size for the above analysis:  671"
## [1] ""
## [1] ""
## [1] "####################################"
## [1] "Analysis of AJSC 2916"
## [1] "####################################"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  792  864  825  755  844  807  764  630  627  666  799  969 1052 1353 1390
## 2011 2012
## 1637 1621
##
```

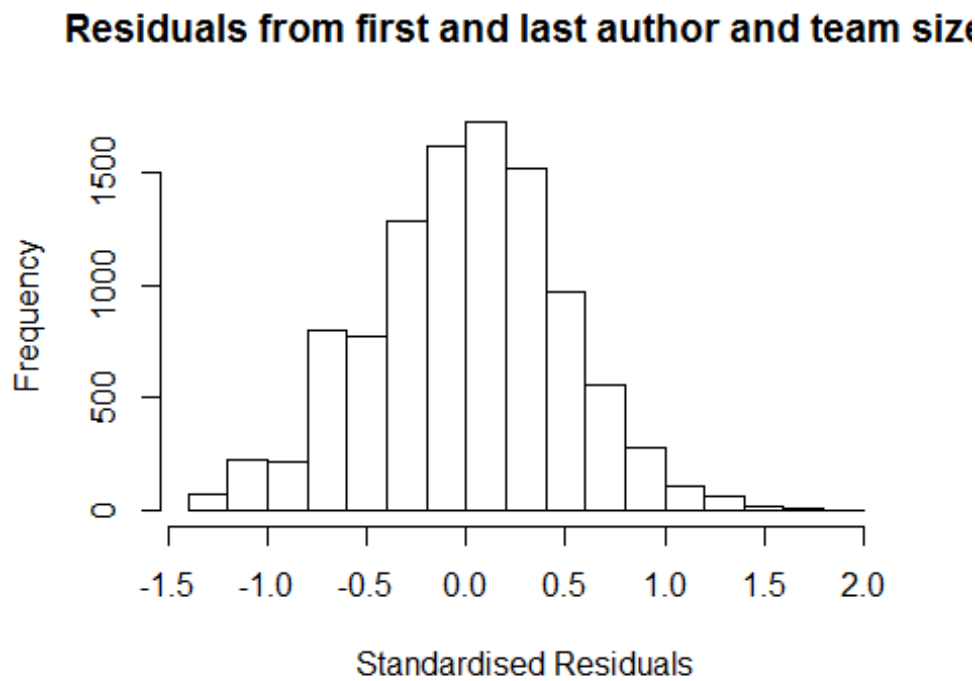
```
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  518  598  555  480  452  435  533  447  424  485  582  727  784  986 1066
## 2011 2012
## 1251 1250
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  455  542  503  407  404  388  475  405  378  423  507  642  676  857  947
## 2011 2012
## 1118 1099
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
##  Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 150, df = 16, p-value <2e-16
```




```
##
##  Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 2.4, df = 1, p-value = 0.1
```

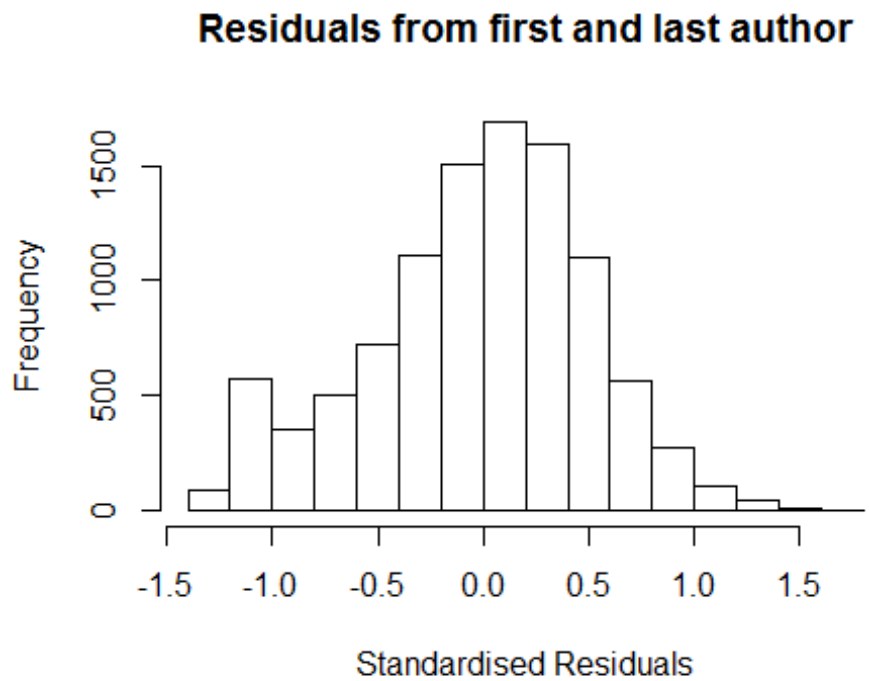



```
## [1] "Female first author team size 2018 geometric mean: 4.58165611893995"
## [1] "Male first author team size 2018 geometric mean: 4.48401475765137"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 110000, p-value = 1
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 4.60583846367206"
## [1] "Male last author team size 2018 geometric mean: 4.49192675480573"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 130000, p-value = 0.5
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.045  1           1.022
## LastAuthorFemale  1.044  1           1.022
## UniqueAuthors     1.089  4           1.011
## Year              1.099 16           1.003
```



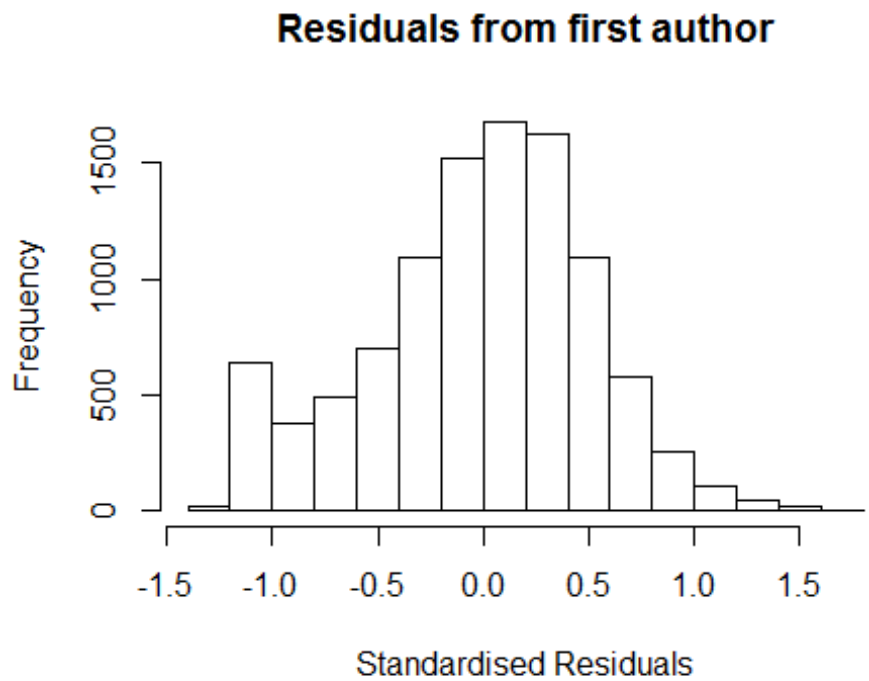

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
UniqueAuthors +
##     Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##     k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.3861 -0.3226  0.0132  0.3145  1.8363
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)         0.719291   0.034999   20.55  < 2e-16 ***
## FirstAuthorFemale1 -0.014798   0.010125   -1.46    0.144
## LastAuthorFemale1  -0.053444   0.010082   -5.30  1.2e-07 ***
## UniqueAuthors2      0.371400   0.024346   15.25  < 2e-16 ***
## UniqueAuthors3      0.465925   0.023261   20.03  < 2e-16 ***
## UniqueAuthors4      0.545553   0.022797   23.93  < 2e-16 ***
## UniqueAuthors5      0.631344   0.020694   30.51  < 2e-16 ***
## Year1997           -0.021961   0.039107   -0.56    0.574
## Year1998           -0.015602   0.038967   -0.40    0.689
## Year1999            0.035468   0.040713    0.87    0.384
```

```
## Year2000           -0.020959   0.038970   -0.54    0.591
## Year2001           -0.061414   0.041714   -1.47    0.141
## Year2002           -0.000124   0.037990    0.00    0.997
## Year2003           -0.041287   0.038926   -1.06    0.289
## Year2004            0.000854   0.038171    0.02    0.982
## Year2005           -0.082018   0.038016   -2.16    0.031 *
## Year2006           -0.004757   0.035652   -0.13    0.894
## Year2007            0.004233   0.035588    0.12    0.905
## Year2008           -0.029740   0.034705   -0.86    0.392
## Year2009           -0.003341   0.034516   -0.10    0.923
## Year2010           -0.021090   0.033833   -0.62    0.533
## Year2011           -0.025971   0.033770   -0.77    0.442
## Year2012           -0.012233   0.033739   -0.36    0.717
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.461
## Multiple R-squared:  0.162,  Adjusted R-squared:  0.16
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
##  903 weights are ~= 1. The remaining 9323 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0766  0.8510  0.9480  0.8970  0.9840  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          9.78e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.029  1           1.014
## LastAuthorFemale  1.030  1           1.015
## Year              1.015 16           1.000
```



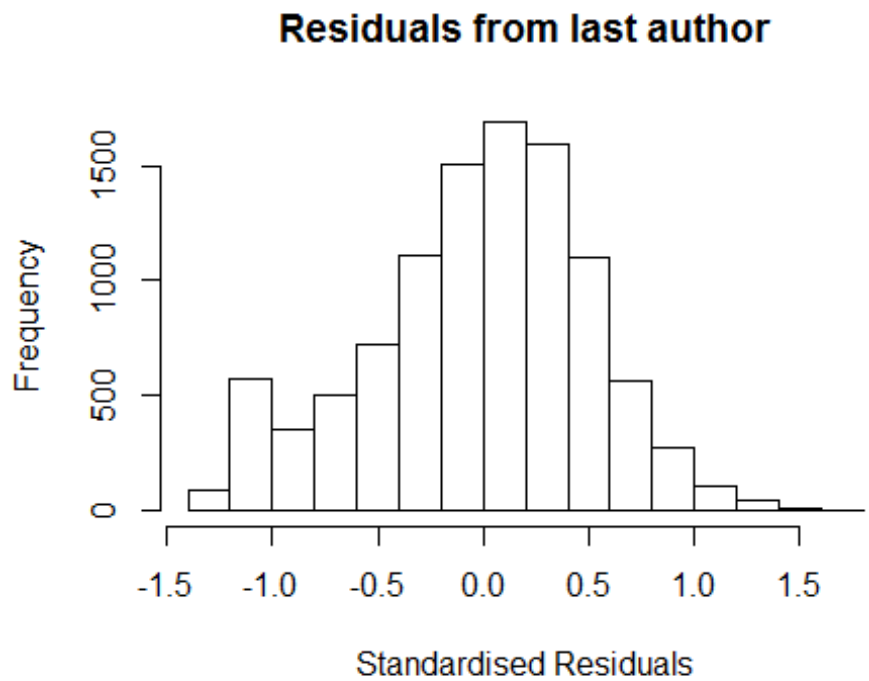

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.2525 -0.3326  0.0318  0.3347  1.6413
##
## Coefficients:
##                      Estimate Std. Error t value Pr(>|t|)
## (Intercept)          1.113126   0.033864   32.87  < 2e-16 ***
## FirstAuthorFemale1   0.022110   0.010804    2.05   0.0407 *
## LastAuthorFemale1   -0.082566   0.010908   -7.57  4.1e-14 ***
## Year1997            -0.006092   0.043709   -0.14   0.8891
## Year1998             0.000757   0.043482    0.02   0.9861
## Year1999             0.068557   0.044859    1.53   0.1265
## Year2000             0.003214   0.042174    0.08   0.9393
## Year2001            -0.049291   0.047454   -1.04   0.2990
## Year2002             0.067326   0.041404    1.63   0.1040
## Year2003             0.050799   0.042136    1.21   0.2280
## Year2004             0.070910   0.041872    1.69   0.0904 .
## Year2005            -0.014185   0.042544   -0.33   0.7388
```

```
## Year2006            0.117290   0.038482    3.05   0.0023 **
## Year2007            0.102942   0.038113    2.70   0.0069 **
## Year2008            0.083915   0.037682    2.23   0.0260 *
## Year2009            0.101568   0.037782    2.69   0.0072 **
## Year2010            0.087019   0.037194    2.34   0.0193 *
## Year2011            0.079063   0.036767    2.15   0.0315 *
## Year2012            0.078820   0.036958    2.13   0.0330 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.487
## Multiple R-squared:  0.0131, Adjusted R-squared:  0.0113
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
##  834 weights are ~= 1. The remaining 9392 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.232   0.856   0.948   0.895   0.985   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          9.78e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.006  1           1.003
## Year              1.006 16           1.000
```




```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##    Min     1Q Median     3Q    Max
## -1.208 -0.330  0.032  0.335  1.607
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.094898   0.033794   32.40   <2e-16 ***
## FirstAuthorFemale1 -0.000619   0.010774   -0.06   0.9542
## Year1997           -0.009665   0.043848   -0.22   0.8255
## Year1998           -0.002267   0.043596   -0.05   0.9585
## Year1999            0.063296   0.045113    1.40   0.1606
## Year2000           -0.006327   0.042411   -0.15   0.8814
## Year2001           -0.059744   0.047779   -1.25   0.2112
## Year2002            0.066783   0.041508    1.61   0.1077
## Year2003            0.046375   0.042139    1.10   0.2711
## Year2004            0.068450   0.042045    1.63   0.1035
## Year2005           -0.016452   0.042517   -0.39   0.6988
## Year2006            0.113081   0.038518    2.94   0.0033 **
```

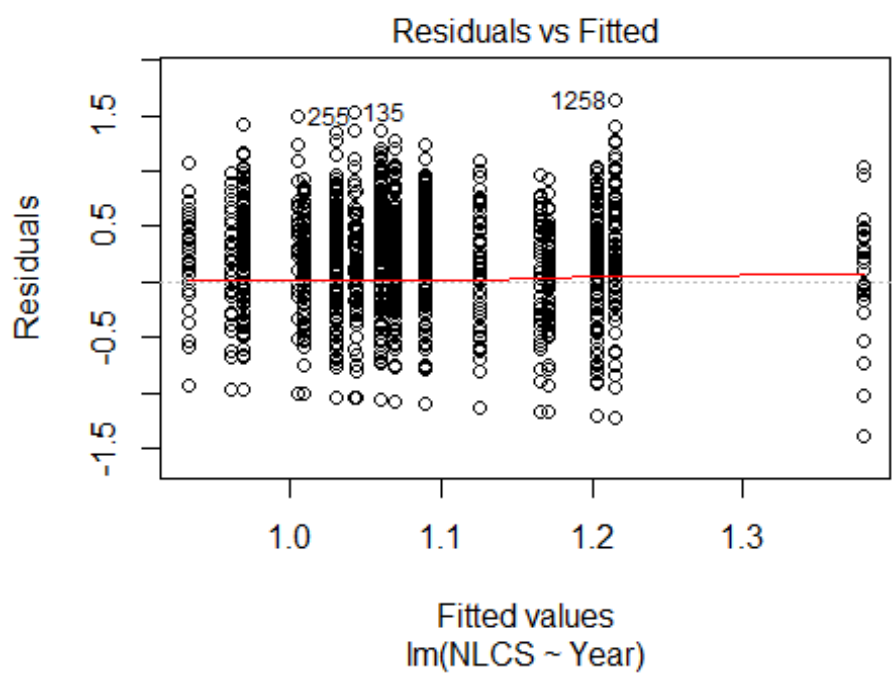
```
## Year2007            0.100564   0.038210    2.63   0.0085 **
## Year2008            0.081345   0.037837    2.15   0.0316 *
## Year2009            0.100711   0.037822    2.66   0.0078 **
## Year2010            0.080118   0.037291    2.15   0.0317 *
## Year2011            0.072966   0.036904    1.98   0.0480 *
## Year2012            0.074089   0.037075    2.00   0.0457 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.487
## Multiple R-squared:  0.00755,    Adjusted R-squared:  0.0059
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
##  852 weights are ~= 1. The remaining 9374 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.253   0.856   0.949   0.894   0.985   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          9.78e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##                   GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.008  1           1.004
## Year             1.008 16           1.000
```



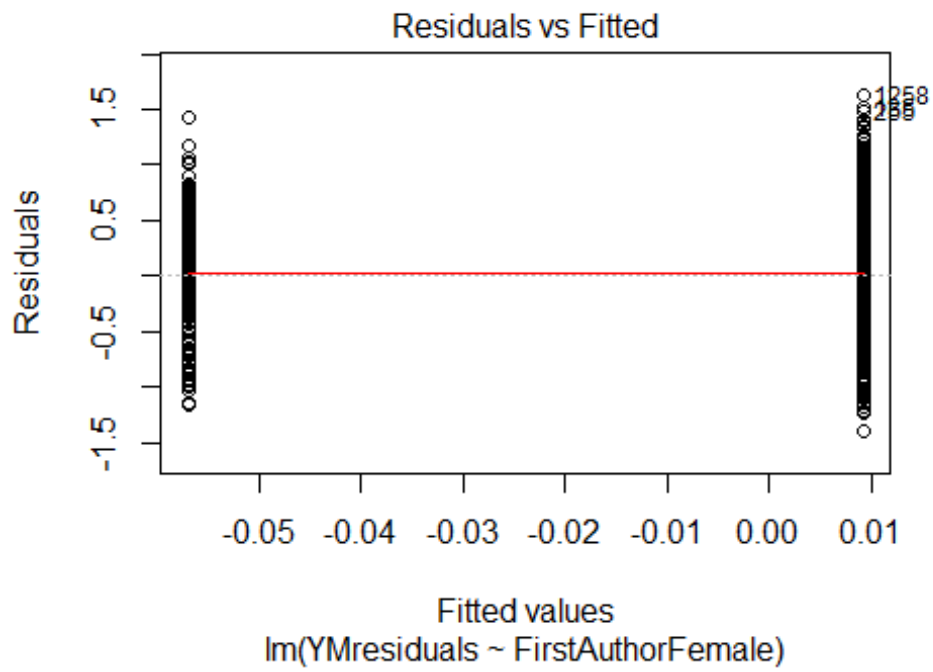

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.2406 -0.3305  0.0306  0.3347  1.6471
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)        1.12e+00   3.35e-02   33.50   <2e-16 ***
## LastAuthorFemale1 -7.65e-02   1.09e-02   -7.04    2e-12 ***
## Year1997          -6.31e-03   4.36e-02   -0.14   0.8851
## Year1998          -8.53e-05   4.34e-02    0.00   0.9984
## Year1999           6.89e-02   4.48e-02    1.54   0.1240
## Year2000           3.69e-03   4.21e-02    0.09   0.9301
## Year2001          -4.87e-02   4.74e-02   -1.03   0.3037
## Year2002           6.88e-02   4.14e-02    1.66   0.0962 .
## Year2003           5.11e-02   4.21e-02    1.21   0.2250
## Year2004           7.28e-02   4.18e-02    1.74   0.0816 .
## Year2005          -1.33e-02   4.25e-02   -0.31   0.7543
## Year2006           1.18e-01   3.84e-02    3.07   0.0021 **
```

```
## Year2007           1.03e-01   3.81e-02    2.71   0.0068 **
## Year2008           8.42e-02   3.77e-02    2.24   0.0254 *
## Year2009           1.03e-01   3.77e-02    2.73   0.0064 **
## Year2010           8.80e-02   3.72e-02    2.37   0.0178 *
## Year2011           8.12e-02   3.67e-02    2.21   0.0270 *
## Year2012           8.10e-02   3.69e-02    2.20   0.0280 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.487
## Multiple R-squared:  0.0128, Adjusted R-squared:  0.0111
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
##  835 weights are ~= 1. The remaining 9391 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.229   0.856   0.948   0.895   0.985   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          9.78e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Sample size for the above analysis:  10226"
## [1] ""
## [1] ""
## [1] "####################################"
## [1] "Analysis of AJSC 2917"
## [1] "####################################"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  139  138  116  146  127  127  117   91   79   80  129  162  238  251  237
## 2011 2012
##  323  319
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##   36   40   27   44   49   42  104   87   69   73  103  145  216  216  214
## 2011 2012
```

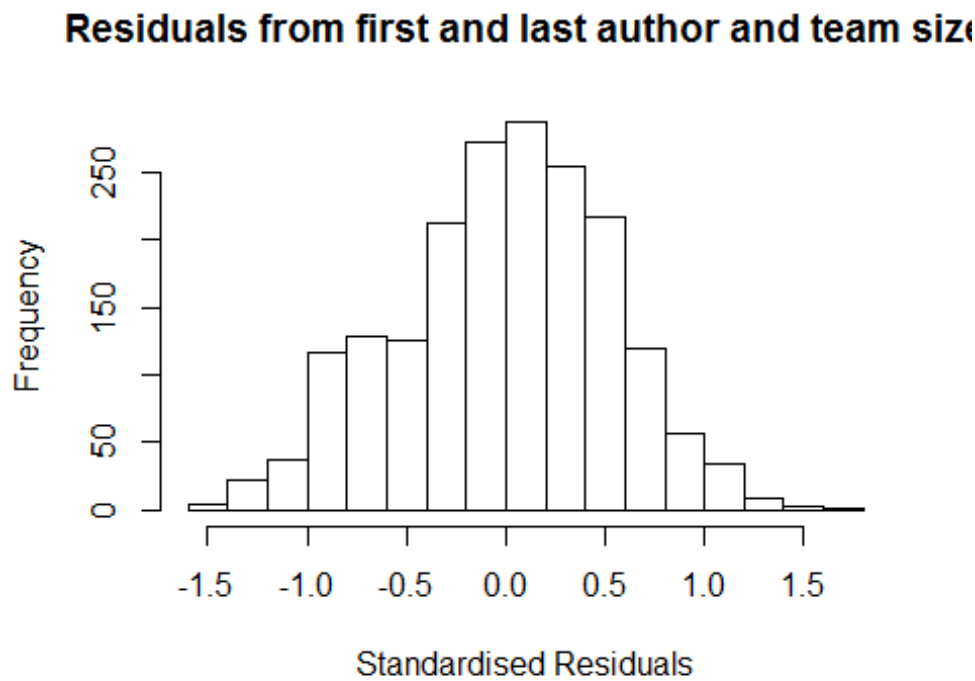
```
##  277  282
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##   36   38   27   44   49   39   93   82   65   70   95  135  201  201  194
## 2011 2012
##  264  264
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
##  Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 75, df = 16, p-value = 1e-09
```




```
##
##  Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 0.1, df = 1, p-value = 0.8
```

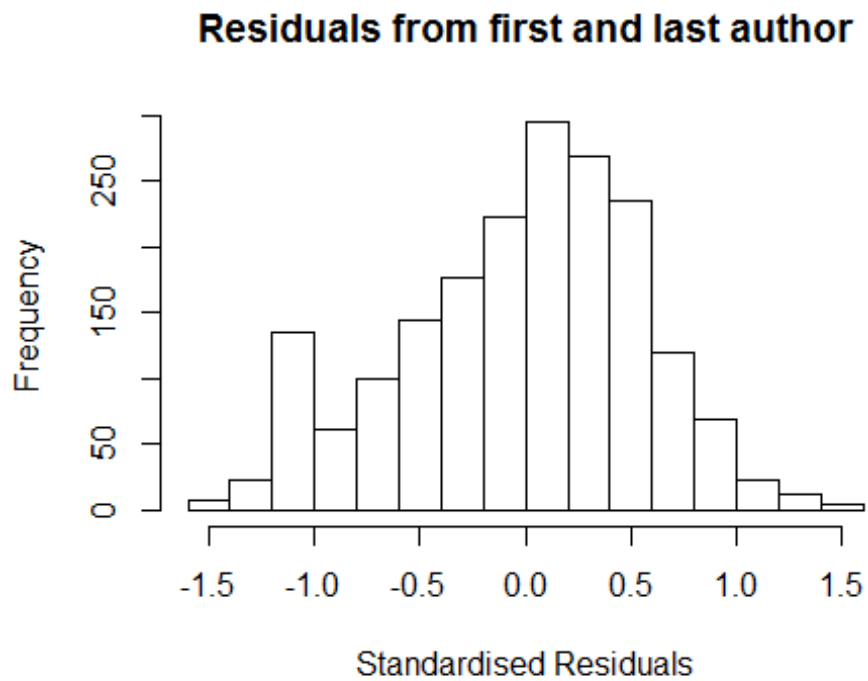



```
## [1] "Female first author team size 2018 geometric mean: 3.90317320075644"
## [1] "Male first author team size 2018 geometric mean: 5.06170685649937"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 3000, p-value = 0.08
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 3.5760697357714"
## [1] "Male last author team size 2018 geometric mean: 5.61966798037977"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 3600, p-value = 1e-04
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##                     GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.163  1           1.078
## LastAuthorFemale  1.216  1           1.103
## UniqueAuthors     1.321  4           1.035
## Year              1.328 16           1.009
```



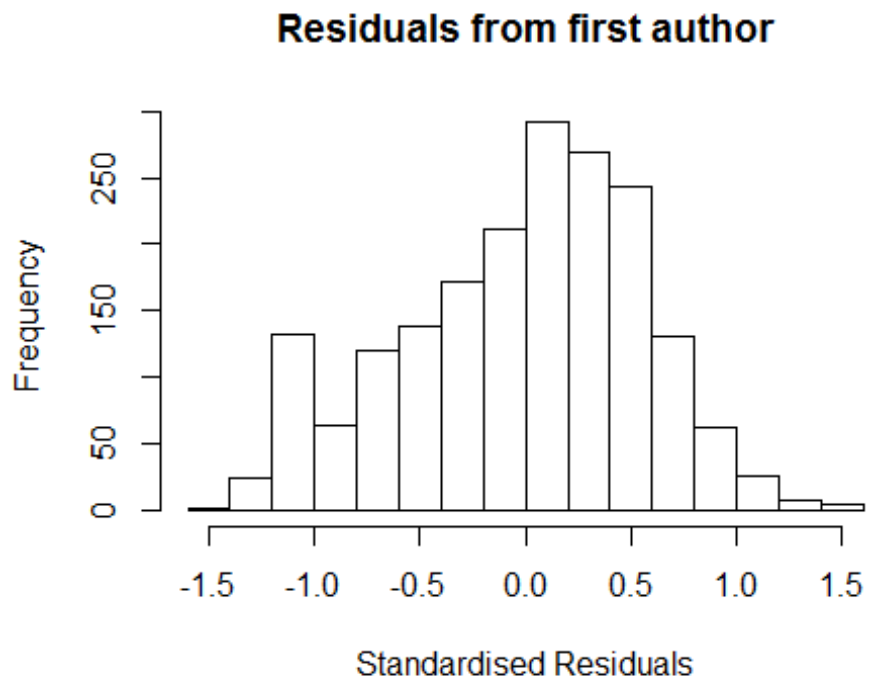

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
UniqueAuthors +
##     Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##     k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.5816 -0.3594  0.0195  0.3701  1.6264
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)          0.8599     0.1981    4.34  1.5e-05 ***
## FirstAuthorFemale1   0.1046     0.0417    2.51    0.012 *
## LastAuthorFemale1   -0.0229     0.0328   -0.70    0.486
## UniqueAuthors2       0.1858     0.0438    4.24  2.3e-05 ***
## UniqueAuthors3       0.3191     0.0450    7.09  1.9e-12 ***
## UniqueAuthors4       0.4361     0.0429   10.17  < 2e-16 ***
## UniqueAuthors5       0.4898     0.0346   14.15  < 2e-16 ***
## Year1997            -0.0738     0.2634   -0.28    0.779
## Year1998             0.4128     0.2219    1.86    0.063 .
## Year1999             0.1273     0.2136    0.60    0.551
```

```
## Year2000            -0.1720     0.2135   -0.81    0.421
## Year2001            -0.0573     0.2244   -0.26    0.798
## Year2002            -0.1753     0.2009   -0.87    0.383
## Year2003            -0.1301     0.2007   -0.65    0.517
## Year2004            -0.0863     0.1990   -0.43    0.665
## Year2005            -0.0709     0.2039   -0.35    0.728
## Year2006             0.1112     0.2117    0.53    0.600
## Year2007             0.0420     0.1985    0.21    0.832
## Year2008            -0.1228     0.1969   -0.62    0.533
## Year2009            -0.1687     0.1962   -0.86    0.390
## Year2010            -0.1123     0.1956   -0.57    0.566
## Year2011            -0.1197     0.1944   -0.62    0.538
## Year2012            -0.1589     0.1952   -0.81    0.416
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.529
## Multiple R-squared:  0.139,  Adjusted R-squared:  0.129
## Convergence in 21 IRWLS iterations
##
## Robustness weights:
##  159 weights are ~= 1. The remaining 1738 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.324   0.853   0.948   0.904   0.985   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          5.27e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.171  1           1.082
## LastAuthorFemale  1.165  1           1.079
## Year              1.083 16           1.002
```



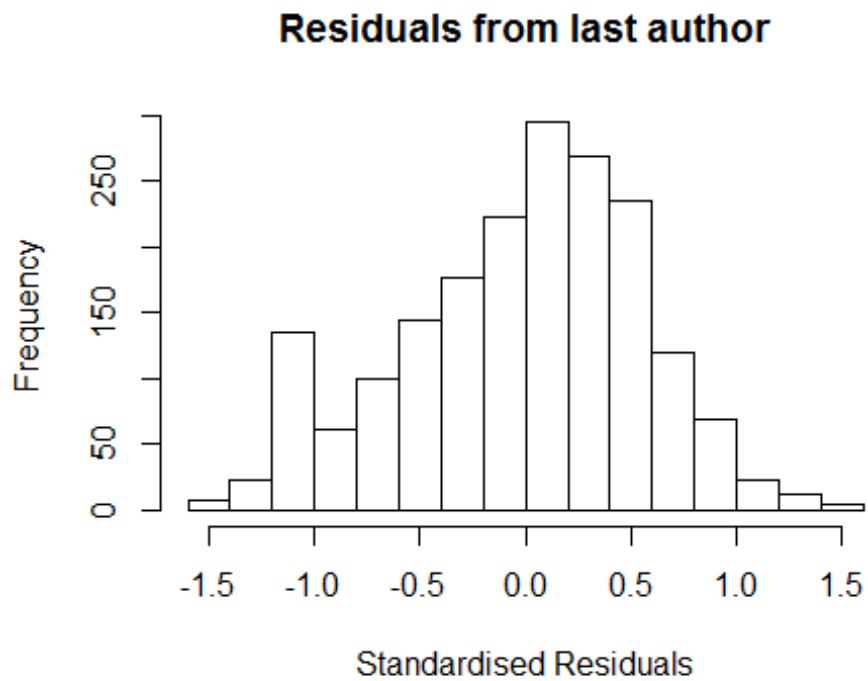

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.4264 -0.3910  0.0546  0.3940  1.5652
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)           1.0346     0.2102    4.92  9.3e-07 ***
## FirstAuthorFemale1    0.1336     0.0448    2.98  0.00291 **
## LastAuthorFemale1    -0.1223     0.0335   -3.65  0.00027 ***
## Year1997             -0.0335     0.2733   -0.12  0.90236
## Year1998              0.3805     0.2321    1.64  0.10132
## Year1999              0.1380     0.2253    0.61  0.54014
## Year2000             -0.0985     0.2261   -0.44  0.66330
## Year2001             -0.0359     0.2304   -0.16  0.87605
## Year2002             -0.0364     0.2153   -0.17  0.86570
## Year2003              0.0182     0.2127    0.09  0.93198
## Year2004              0.0984     0.2125    0.46  0.64329
## Year2005              0.0504     0.2166    0.23  0.81603
```

```
## Year2006             0.2389     0.2250    1.06  0.28831
## Year2007             0.1936     0.2135    0.91  0.36465
## Year2008            -0.0132     0.2119   -0.06  0.95046
## Year2009            -0.0765     0.2108   -0.36  0.71679
## Year2010             0.0173     0.2101    0.08  0.93427
## Year2011             0.0551     0.2088    0.26  0.79169
## Year2012             0.0148     0.2097    0.07  0.94381
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.55
## Multiple R-squared:  0.0338, Adjusted R-squared:  0.0245
## Convergence in 21 IRWLS iterations
##
## Robustness weights:
##  149 weights are ~= 1. The remaining 1748 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.397   0.861   0.945   0.901   0.983   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          5.27e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.052  1           1.026
## Year              1.052 16           1.002
```




```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.4347 -0.4002  0.0603  0.3990  1.5493
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)         0.973959   0.209470    4.65  3.6e-06 ***
## FirstAuthorFemale1  0.077686   0.041897    1.85    0.064 .
## Year1997           -0.026655   0.272186   -0.10    0.922
## Year1998            0.383054   0.233238    1.64    0.101
## Year1999            0.144138   0.226488    0.64    0.525
## Year2000           -0.086407   0.226062   -0.38    0.702
## Year2001           -0.022481   0.230564   -0.10    0.922
## Year2002           -0.020440   0.216217   -0.09    0.925
## Year2003            0.026060   0.213024    0.12    0.903
## Year2004            0.123516   0.212608    0.58    0.561
## Year2005            0.071789   0.216877    0.33    0.741
## Year2006            0.249053   0.226251    1.10    0.271
```

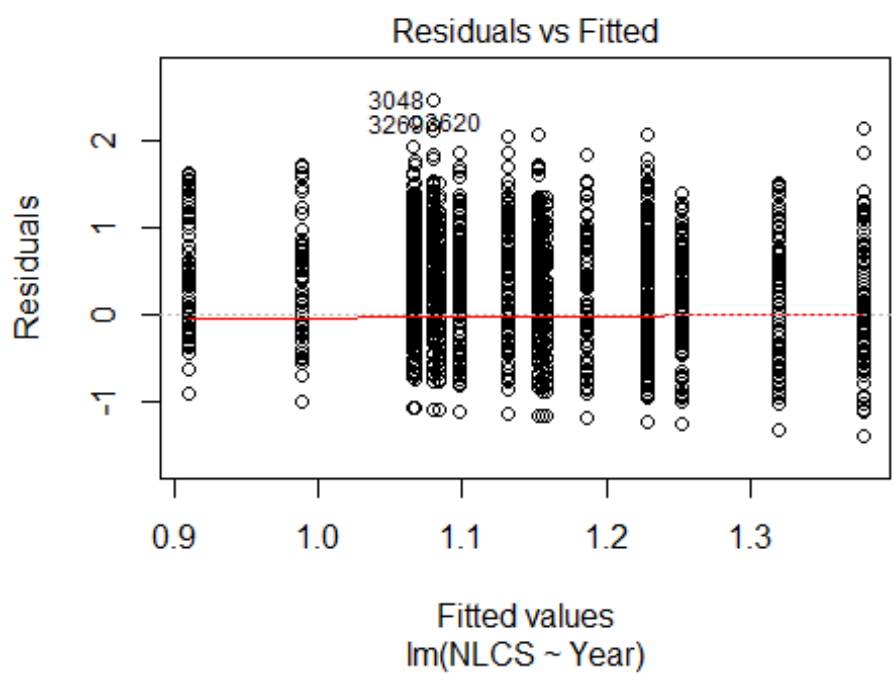
```
## Year2007            0.210268   0.213863    0.98    0.326
## Year2008           -0.000659   0.212194    0.00    0.998
## Year2009           -0.061792   0.210944   -0.29    0.770
## Year2010            0.036791   0.210290    0.17    0.861
## Year2011            0.074788   0.209082    0.36    0.721
## Year2012            0.029864   0.209907    0.14    0.887
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.553
## Multiple R-squared:  0.0264, Adjusted R-squared:  0.0176
## Convergence in 21 IRWLS iterations
##
## Robustness weights:
##  145 weights are ~= 1. The remaining 1752 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.413   0.863   0.945   0.902   0.984   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          5.27e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##                   GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.041  1           1.020
## Year             1.041 16           1.001
```



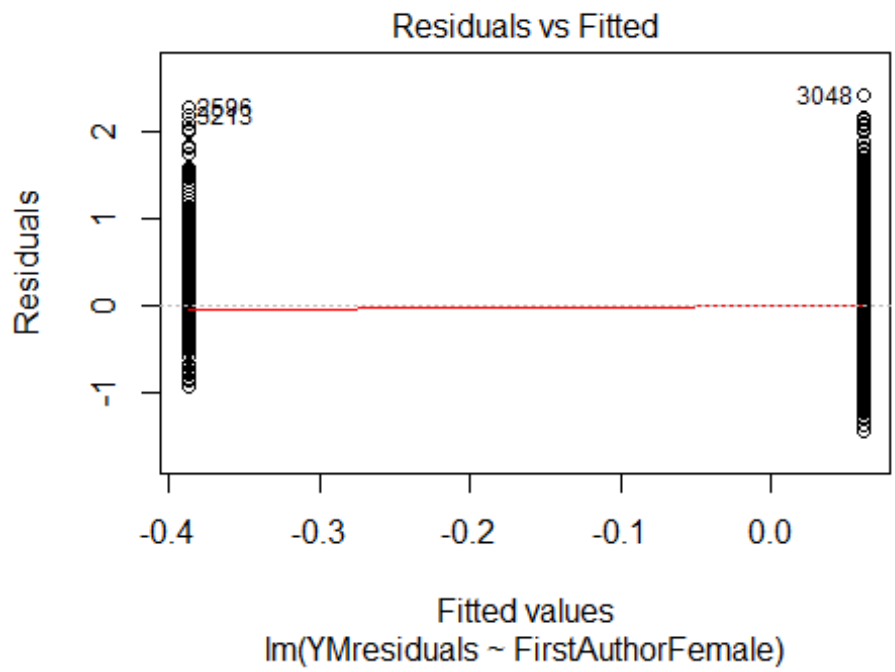

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##    Min     1Q Median     3Q    Max
## -1.423 -0.393  0.061  0.402  1.566
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.1211     0.2046    5.48  4.8e-08 ***
## LastAuthorFemale1  -0.0839     0.0323   -2.60   0.0094 **
## Year1997           -0.0220     0.2706   -0.08   0.9351
## Year1998            0.3859     0.2287    1.69   0.0917 .
## Year1999            0.1376     0.2214    0.62   0.5344
## Year2000           -0.0911     0.2225   -0.41   0.6822
## Year2001           -0.0359     0.2274   -0.16   0.8744
## Year2002           -0.0260     0.2120   -0.12   0.9025
## Year2003            0.0228     0.2091    0.11   0.9132
## Year2004            0.1121     0.2088    0.54   0.5913
## Year2005            0.0626     0.2131    0.29   0.7691
## Year2006            0.2470     0.2220    1.11   0.2659
```

```
## Year2007            0.1831     0.2101    0.87   0.3835
## Year2008           -0.0192     0.2085   -0.09   0.9267
## Year2009           -0.0777     0.2072   -0.37   0.7077
## Year2010            0.0183     0.2064    0.09   0.9293
## Year2011            0.0483     0.2051    0.24   0.8138
## Year2012            0.0117     0.2061    0.06   0.9546
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.552
## Multiple R-squared:  0.0288, Adjusted R-squared:  0.02
## Convergence in 21 IRWLS iterations
##
## Robustness weights:
##  140 weights are ~= 1. The remaining 1757 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.401   0.863   0.943   0.902   0.984   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          5.27e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Sample size for the above analysis:  1897"
## [1] ""
## [1] ""
## [1] "####################################"
## [1] "Analysis of AJSC 2919"
## [1] "####################################"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  216  224  196  187  183  190  166  140  143  148  206  237  265  330  380
## 2011 2012
##  362  410
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##   93  116   76   81  108  129  146  127  124  136  177  218  234  295  327
## 2011 2012
```

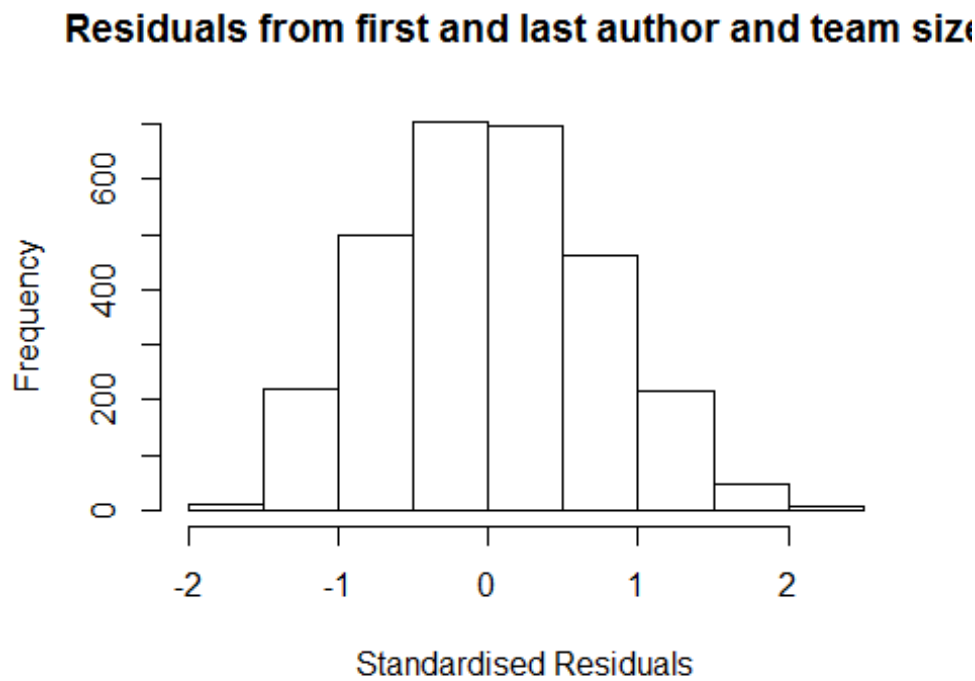
```
##  318  364
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##   89  108   74   80  104  121  135  119  112  125  170  198  222  274  308
## 2011 2012
##  298  327
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
##  Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 23, df = 16, p-value = 0.1
```




```
##
##  Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 1.6, df = 1, p-value = 0.2
```

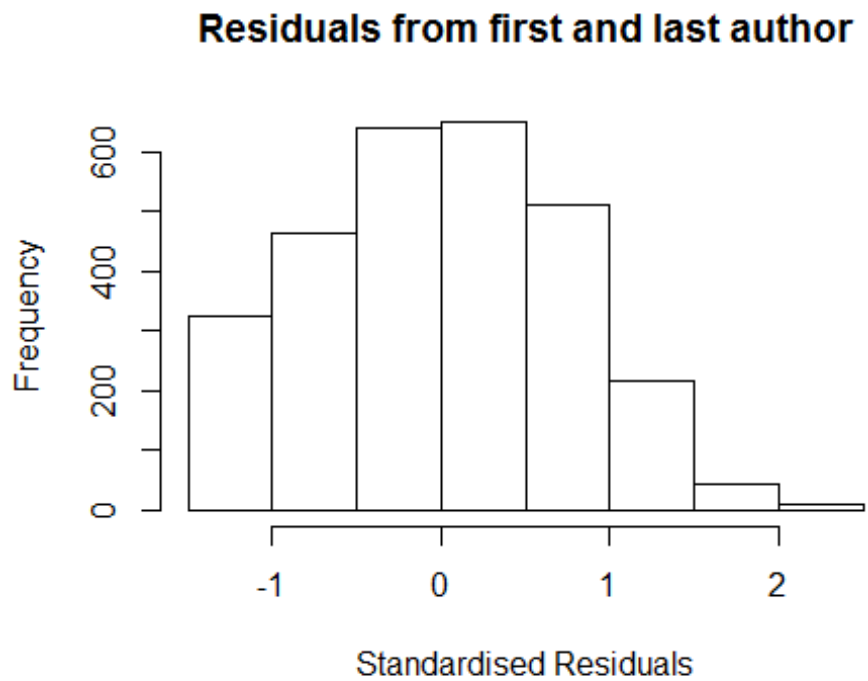



```
## [1] "Female first author team size 2018 geometric mean: 2.73828897583865"
## [1] "Male first author team size 2018 geometric mean: 2.40845473994476"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 2500, p-value = 0.4
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 2.60495293334932"
## [1] "Male last author team size 2018 geometric mean: 3.14236198805633"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 2900, p-value = 0.1
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.662  1           1.289
## LastAuthorFemale  1.627  1           1.276
## UniqueAuthors     1.164  4           1.019
## Year              1.187 16           1.005
```



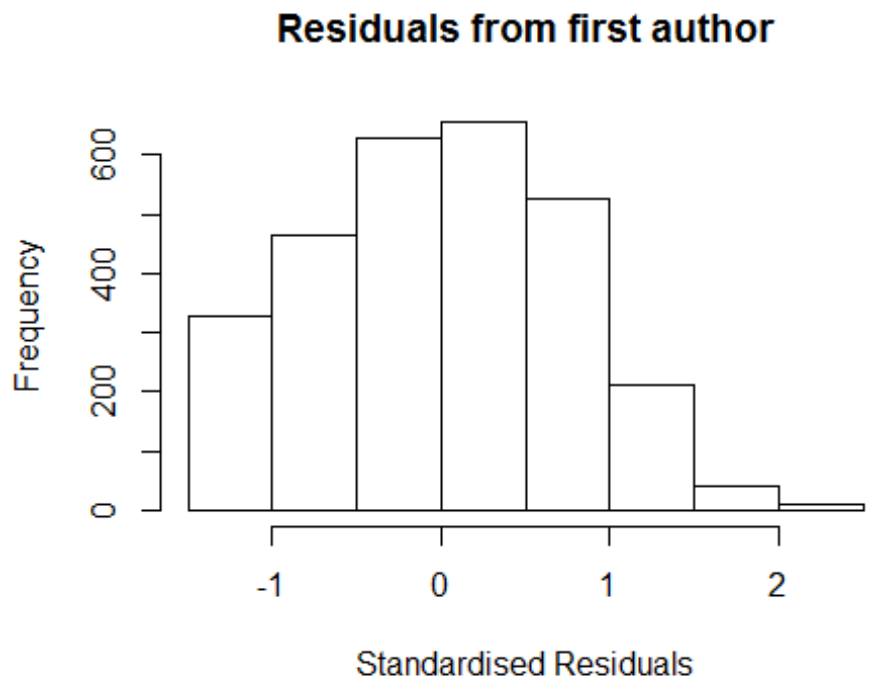

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
UniqueAuthors +
##     Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##     k.max = 1000))
##  \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.74203 -0.51593 -0.00224  0.51806  2.46401
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)          0.3770     0.0993    3.80  0.00015 ***
## FirstAuthorFemale1   0.3781     0.0479    7.89  4.2e-15 ***
## LastAuthorFemale1    0.1388     0.0426    3.26  0.00113 **
## UniqueAuthors2       0.1464     0.0361    4.06  5.1e-05 ***
## UniqueAuthors3       0.3644     0.0450    8.09  8.8e-16 ***
## UniqueAuthors4       0.4118     0.0546    7.55  6.0e-14 ***
## UniqueAuthors5       0.5537     0.0462   11.99  < 2e-16 ***
## Year1997            -0.0152     0.1193   -0.13  0.89828
## Year1998             0.1995     0.1213    1.64  0.10030
## Year1999             0.2272     0.1204    1.89  0.05936 .
```

```
## Year2000             0.2255     0.1252    1.80  0.07185 .
## Year2001             0.2567     0.1210    2.12  0.03398 *
## Year2002             0.1855     0.1194    1.55  0.12032
## Year2003             0.3363     0.1186    2.83  0.00462 **
## Year2004             0.2944     0.1178    2.50  0.01246 *
## Year2005             0.2243     0.1106    2.03  0.04273 *
## Year2006             0.0947     0.1060    0.89  0.37198
## Year2007             0.1625     0.1076    1.51  0.13109
## Year2008             0.0831     0.1053    0.79  0.43040
## Year2009             0.1515     0.1021    1.48  0.13792
## Year2010             0.0447     0.1020    0.44  0.66098
## Year2011             0.0363     0.1014    0.36  0.72069
## Year2012            -0.0219     0.0989   -0.22  0.82463
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.756
## Multiple R-squared:  0.121,  Adjusted R-squared:  0.115
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
##  207 weights are ~= 1. The remaining 2657 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.267   0.868   0.952   0.917   0.985   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          3.49e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.534  1           1.238
## LastAuthorFemale  1.535  1           1.239
## Year              1.055 16           1.002
```



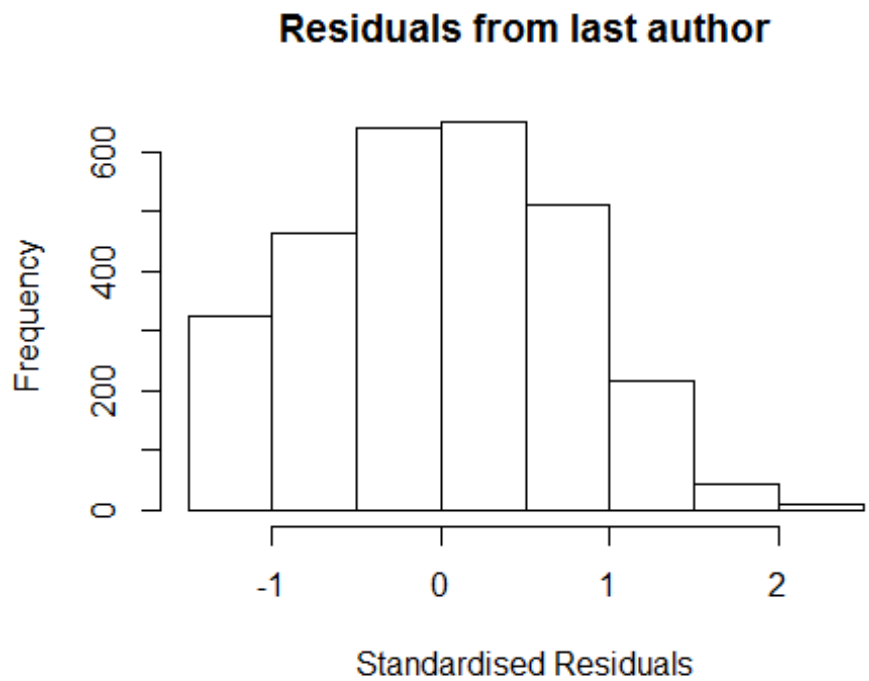

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.38838 -0.58598  0.00356  0.55675  2.42383
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)           0.4854     0.0995    4.88  1.1e-06 ***
## FirstAuthorFemale1    0.4715     0.0498    9.46  < 2e-16 ***
## LastAuthorFemale1     0.0525     0.0444    1.18   0.2373
## Year1997             -0.0293     0.1184   -0.25   0.8049
## Year1998              0.2107     0.1220    1.73   0.0841 .
## Year1999              0.2228     0.1191    1.87   0.0614 .
## Year2000              0.2261     0.1249    1.81   0.0704 .
## Year2001              0.2471     0.1189    2.08   0.0378 *
## Year2002              0.2180     0.1210    1.80   0.0716 .
## Year2003              0.3790     0.1180    3.21   0.0013 **
## Year2004              0.3358     0.1186    2.83   0.0047 **
## Year2005              0.2676     0.1099    2.43   0.0150 *
```

```
## Year2006             0.1453     0.1070    1.36   0.1747
## Year2007             0.2286     0.1080    2.12   0.0344 *
## Year2008             0.1787     0.1063    1.68   0.0928 .
## Year2009             0.2191     0.1030    2.13   0.0334 *
## Year2010             0.1158     0.1034    1.12   0.2629
## Year2011             0.1222     0.1024    1.19   0.2328
## Year2012             0.1006     0.1000    1.01   0.3142
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.793
## Multiple R-squared:  0.0647, Adjusted R-squared:  0.0587
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  227 weights are ~= 1. The remaining 2637 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.330   0.874   0.947   0.919   0.983   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          3.49e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.032  1           1.016
## Year              1.032 16           1.001
```




```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.38624 -0.59933  0.00614  0.55635  2.42990
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)           0.5034     0.0983    5.12  3.2e-07 ***
## FirstAuthorFemale1    0.5041     0.0411   12.26  < 2e-16 ***
## Year1997             -0.0329     0.1181   -0.28   0.7808
## Year1998              0.2098     0.1216    1.72   0.0846 .
## Year1999              0.2234     0.1189    1.88   0.0605 .
## Year2000              0.2206     0.1247    1.77   0.0771 .
## Year2001              0.2450     0.1186    2.07   0.0389 *
## Year2002              0.2145     0.1207    1.78   0.0758 .
## Year2003              0.3787     0.1178    3.22   0.0013 **
## Year2004              0.3287     0.1182    2.78   0.0055 **
## Year2005              0.2646     0.1097    2.41   0.0159 *
## Year2006              0.1404     0.1068    1.31   0.1887
```

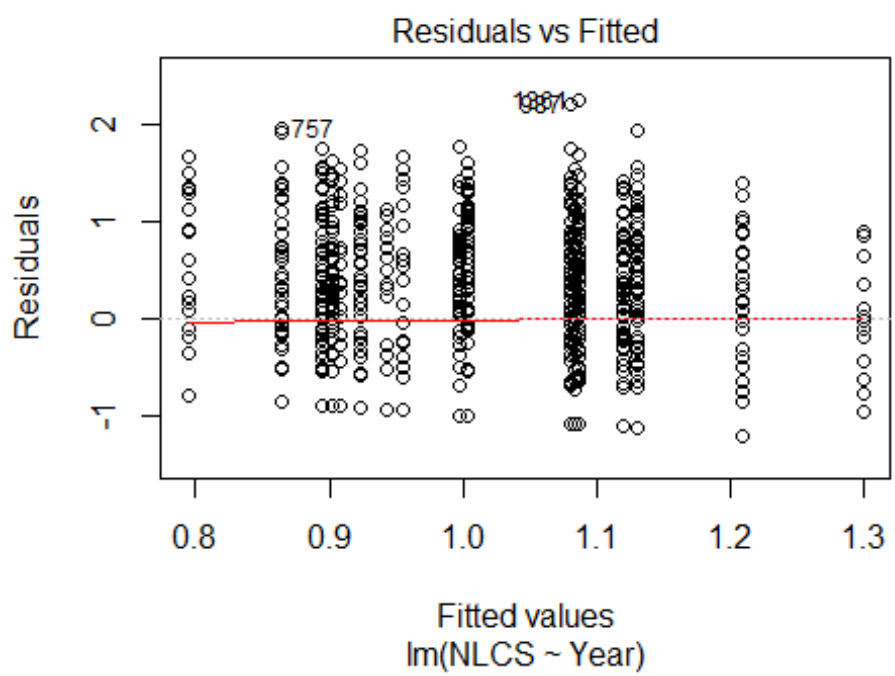
```
## Year2007             0.2209     0.1077    2.05   0.0403 *
## Year2008             0.1703     0.1058    1.61   0.1076
## Year2009             0.2148     0.1028    2.09   0.0367 *
## Year2010             0.1116     0.1032    1.08   0.2796
## Year2011             0.1191     0.1022    1.17   0.2441
## Year2012             0.0959     0.0997    0.96   0.3363
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.794
## Multiple R-squared:  0.0641, Adjusted R-squared:  0.0586
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  233 weights are ~= 1. The remaining 2631 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.328   0.873   0.946   0.919   0.984   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          3.49e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##                   GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.022  1           1.011
## Year             1.022 16           1.001
```




```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.3860 -0.5952  0.0131  0.5674  2.4525
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)          0.7044     0.0947    7.44  1.3e-13 ***
## LastAuthorFemale1    0.2680     0.0394    6.80  1.3e-11 ***
## Year1997            -0.0397     0.1156   -0.34  0.73096
## Year1998             0.2219     0.1228    1.81  0.07097 .
## Year1999             0.2170     0.1158    1.87  0.06109 .
## Year2000             0.2592     0.1211    2.14  0.03246 *
## Year2001             0.2885     0.1156    2.50  0.01262 *
## Year2002             0.2481     0.1179    2.10  0.03551 *
## Year2003             0.4136     0.1156    3.58  0.00035 ***
## Year2004             0.3853     0.1169    3.30  0.00099 ***
## Year2005             0.2880     0.1075    2.68  0.00745 **
## Year2006             0.1524     0.1047    1.46  0.14574
```

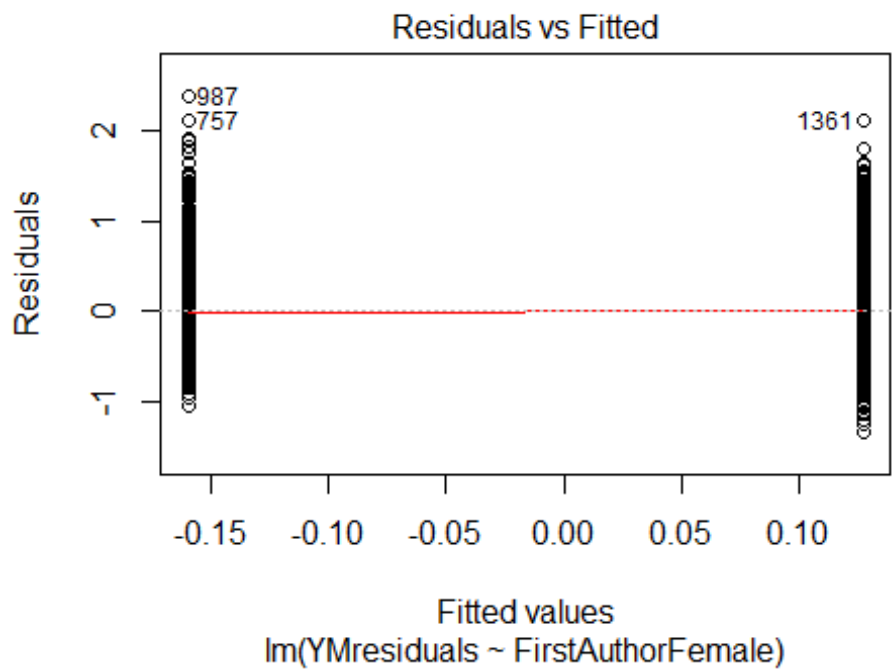
```
## Year2007            0.2335     0.1075    2.17  0.02998 *
## Year2008            0.1965     0.1041    1.89  0.05918 .
## Year2009            0.2365     0.1002    2.36  0.01835 *
## Year2010            0.1241     0.1003    1.24  0.21630
## Year2011            0.1284     0.0994    1.29  0.19684
## Year2012            0.1192     0.0965    1.23  0.21697
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.823
## Multiple R-squared:  0.0351, Adjusted R-squared:  0.0293
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  237 weights are ~= 1. The remaining 2627 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.354   0.884   0.948   0.922   0.984   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          3.49e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Sample size for the above analysis:  2864"
## [1] ""
## [1] ""
## [1] "####################################"
## [1] "Analysis of AJSC 2920"
## [1] "####################################"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##   52   79   51   68   89   59   39   72   71   83   89  117  124  139  110
## 2011 2012
##  147  144
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##   25   48   31   53   40   17   37   65   59   75   83   99  105  120   97
## 2011 2012
```

```
##  124  122
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##   24   46   31   50   38   15   37   61   56   69   79   94  101  103   84
## 2011 2012
##  114  107
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
##  Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 8.1, df = 16, p-value = 0.9
```

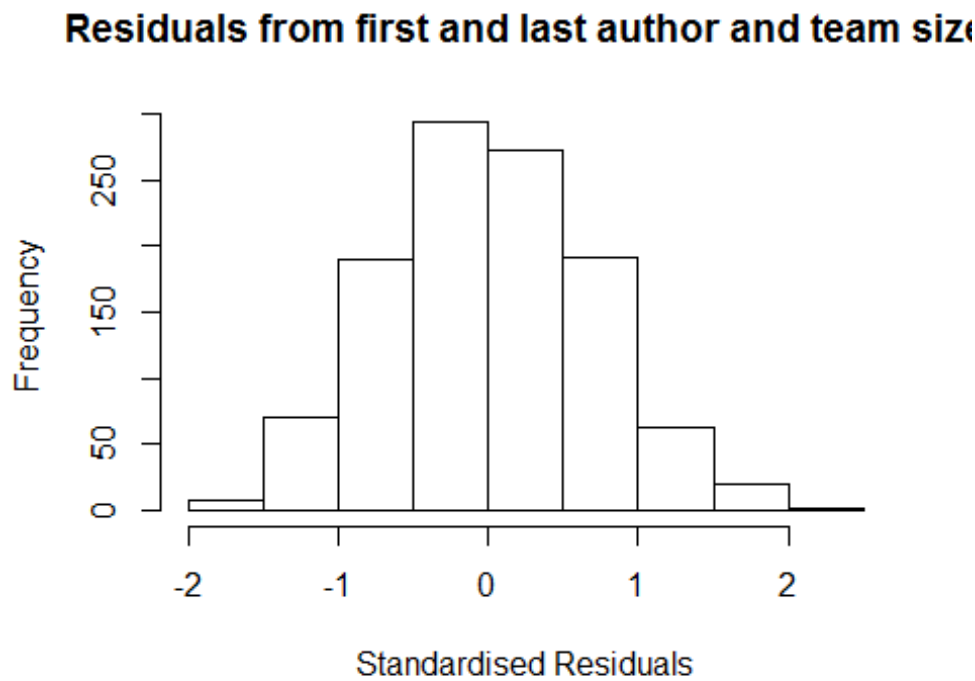



```
##
##  Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 6.8, df = 1, p-value = 0.009

## [1] "Female first author team size 2018 geometric mean: 2.12758662352506"
## [1] "Male first author team size 2018 geometric mean: 3.12031850602576"

## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties
```

```
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 54, p-value = 0.1
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 2.07793161366424"
## [1] "Male last author team size 2018 geometric mean: 3.19284598267482"

## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties
```

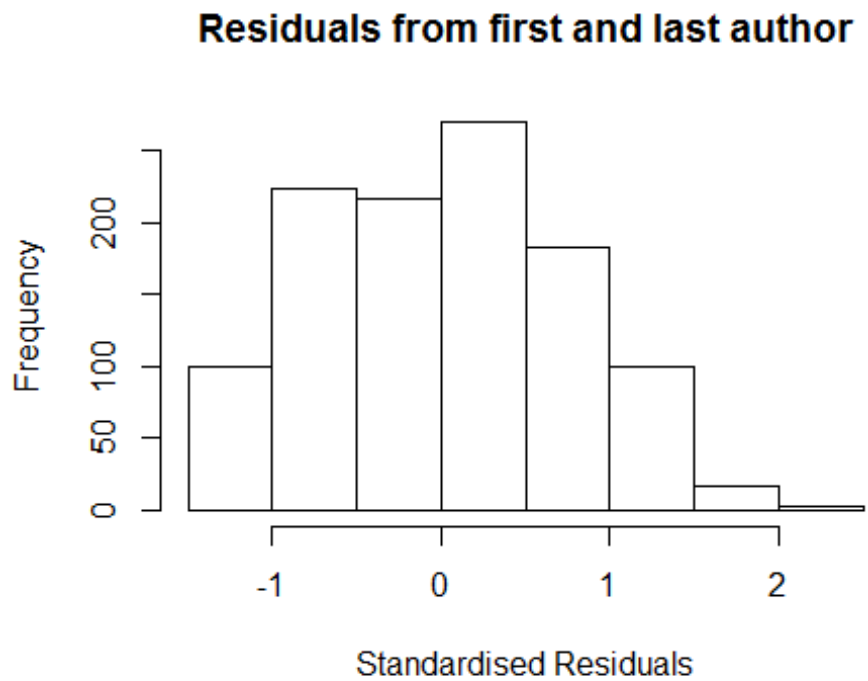



```
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 56, p-value = 0.08
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##                     GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.617  1           1.271
## LastAuthorFemale  1.615  1           1.271
## UniqueAuthors     1.395  4           1.043
## Year              1.444 16           1.012
```



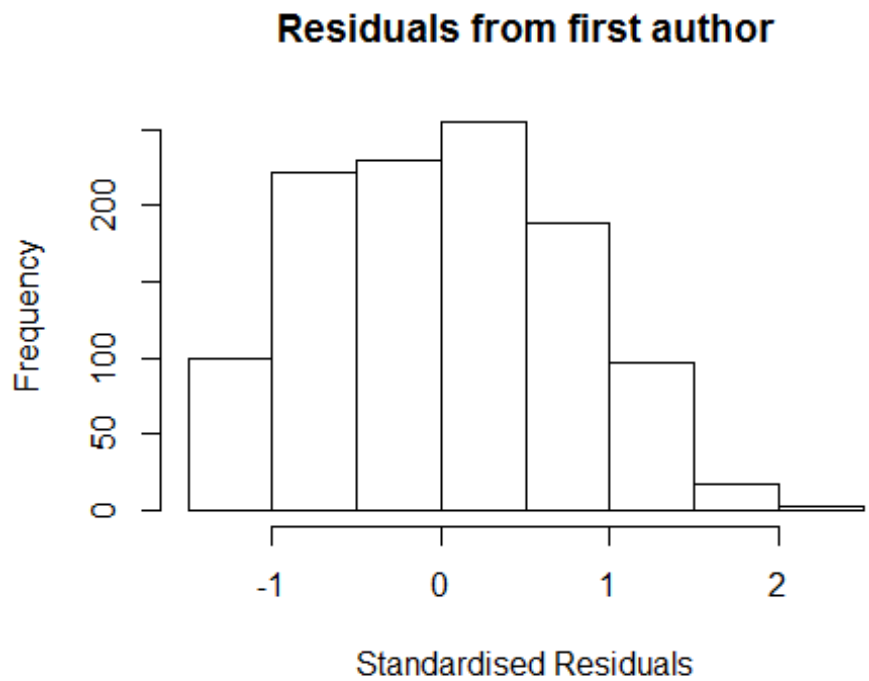

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
UniqueAuthors +
##     Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##     k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.7361 -0.4670 -0.0131  0.4922  2.2628
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         0.24431    0.22772    1.07    0.284
## FirstAuthorFemale1  0.26280    0.05188    5.06  4.8e-07 ***
## LastAuthorFemale1   0.06607    0.05251    1.26    0.209
## UniqueAuthors2      0.59853    0.06243    9.59  < 2e-16 ***
## UniqueAuthors3      0.60513    0.06781    8.92  < 2e-16 ***
## UniqueAuthors4      0.72415    0.07443    9.73  < 2e-16 ***
## UniqueAuthors5      0.89223    0.07169   12.45  < 2e-16 ***
## Year1997            0.00592    0.25429    0.02    0.981
## Year1998            0.23726    0.26811    0.88    0.376
## Year1999            0.33762    0.24491    1.38    0.168
```

```
## Year2000            0.44063    0.26309    1.67    0.094 .
## Year2001            0.53518    0.26979    1.98    0.048 *
## Year2002            0.10706    0.25281    0.42    0.672
## Year2003            0.04249    0.23908    0.18    0.859
## Year2004            0.09165    0.23846    0.38    0.701
## Year2005            0.15794    0.23448    0.67    0.501
## Year2006            0.06279    0.23586    0.27    0.790
## Year2007           -0.07259    0.23500   -0.31    0.757
## Year2008            0.12928    0.23672    0.55    0.585
## Year2009            0.27070    0.23778    1.14    0.255
## Year2010            0.13218    0.24232    0.55    0.586
## Year2011            0.11900    0.23795    0.50    0.617
## Year2012            0.05695    0.23509    0.24    0.809
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.65
## Multiple R-squared:  0.251,  Adjusted R-squared:  0.236
## Convergence in 16 IRWLS iterations
##
## Robustness weights:
##  89 weights are ~= 1. The remaining 1020 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.200   0.866   0.943   0.899   0.982   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          9.02e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.367  1           1.169
## LastAuthorFemale  1.406  1           1.186
## Year              1.120 16           1.004
```



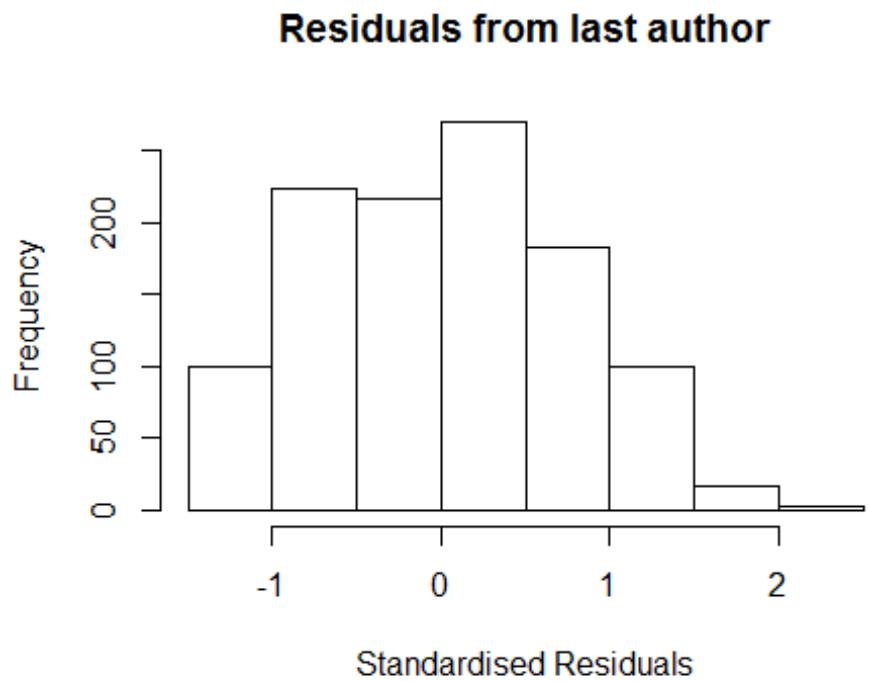

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##    Min     1Q Median     3Q    Max
## -1.334 -0.641  0.023  0.532  2.403
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)           0.6029     0.1877    3.21   0.0014 **
## FirstAuthorFemale1    0.2819     0.0560    5.03  5.7e-07 ***
## LastAuthorFemale1     0.0745     0.0562    1.33   0.1851
## Year1997             -0.0272     0.2189   -0.12   0.9012
## Year1998              0.0790     0.2346    0.34   0.7364
## Year1999              0.3165     0.2132    1.48   0.1380
## Year2000              0.3748     0.2241    1.67   0.0947 .
## Year2001              0.4738     0.2347    2.02   0.0438 *
## Year2002              0.0663     0.2253    0.29   0.7687
## Year2003              0.1901     0.2062    0.92   0.3566
## Year2004              0.1223     0.2154    0.57   0.5705
## Year2005              0.1675     0.2096    0.80   0.4244
```

```
## Year2006             0.0384     0.2118    0.18   0.8560
## Year2007             0.0756     0.2031    0.37   0.7097
## Year2008             0.2242     0.2027    1.11   0.2688
## Year2009             0.3391     0.2024    1.68   0.0941 .
## Year2010             0.2769     0.2050    1.35   0.1769
## Year2011             0.2566     0.2033    1.26   0.2072
## Year2012             0.1827     0.2035    0.90   0.3696
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.82
## Multiple R-squared:  0.0626, Adjusted R-squared:  0.0471
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
##  82 weights are ~= 1. The remaining 1027 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.370   0.883   0.945   0.923   0.983   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          9.02e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.044  1           1.022
## Year              1.044 16           1.001
```




```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##    Min     1Q Median     3Q    Max
## -1.325 -0.638  0.027  0.537  2.459
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)           0.6174     0.1875    3.29    0.001 **
## FirstAuthorFemale1    0.3212     0.0488    6.58  7.4e-11 ***
## Year1997             -0.0116     0.2169   -0.05    0.958
## Year1998              0.0914     0.2333    0.39    0.695
## Year1999              0.3195     0.2130    1.50    0.134
## Year2000              0.3869     0.2230    1.73    0.083 .
## Year2001              0.4897     0.2340    2.09    0.037 *
## Year2002              0.0772     0.2250    0.34    0.731
## Year2003              0.1920     0.2058    0.93    0.351
## Year2004              0.1238     0.2154    0.57    0.566
## Year2005              0.1684     0.2091    0.81    0.421
## Year2006              0.0421     0.2112    0.20    0.842
```

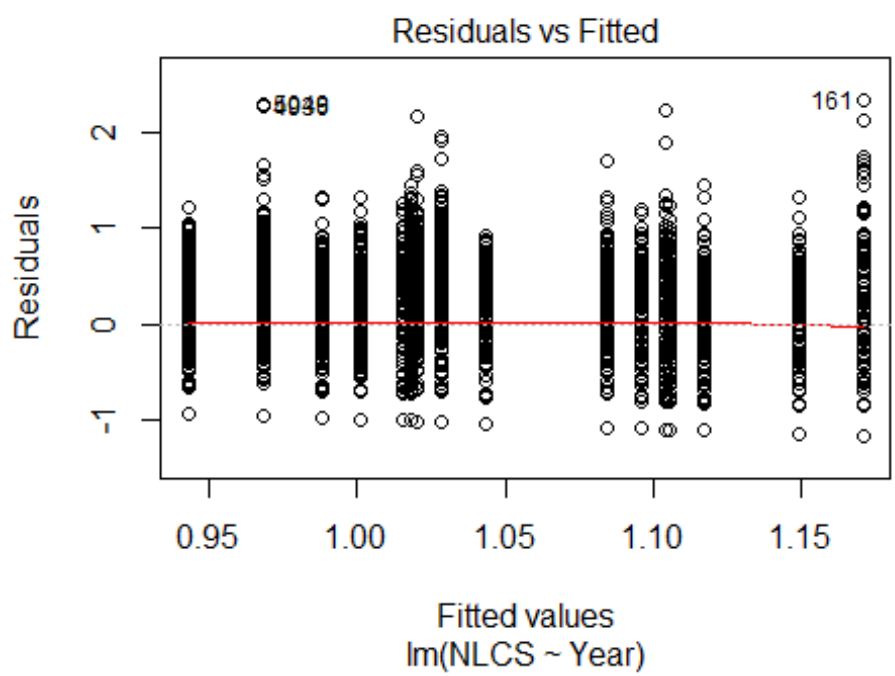
```
## Year2007             0.0743     0.2027    0.37    0.714
## Year2008             0.2288     0.2021    1.13    0.258
## Year2009             0.3369     0.2018    1.67    0.095 .
## Year2010             0.2798     0.2046    1.37    0.172
## Year2011             0.2639     0.2024    1.30    0.192
## Year2012             0.1837     0.2031    0.90    0.366
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.827
## Multiple R-squared:  0.0609, Adjusted R-squared:  0.0462
## Convergence in 10 IRWLS iterations
##
## Robustness weights:
##  79 weights are ~= 1. The remaining 1030 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.356   0.884   0.947   0.924   0.984   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          9.02e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##                   GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.066  1           1.032
## Year             1.066 16           1.002
```



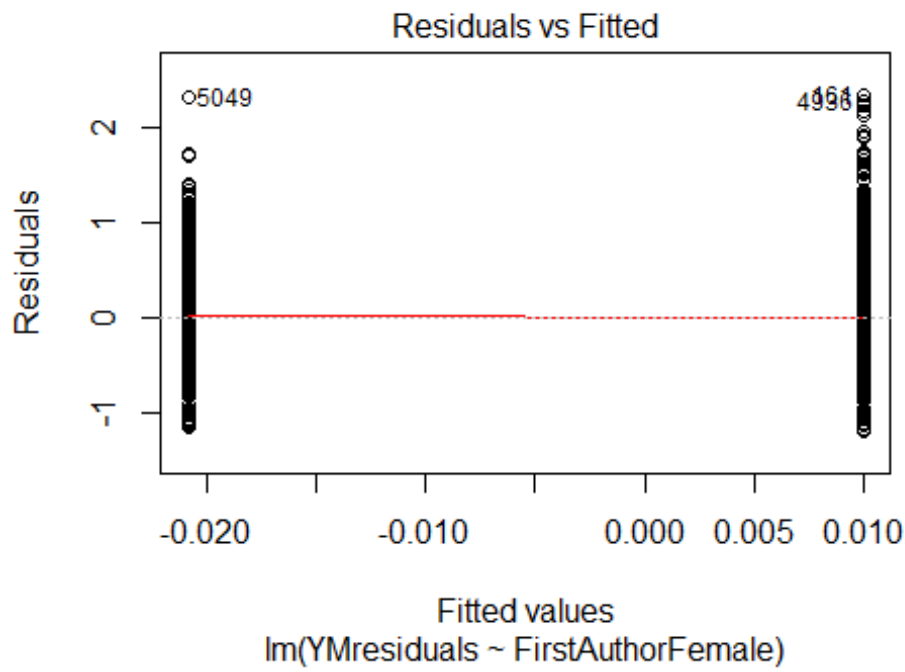

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.27595 -0.63920  0.00826  0.57026  2.17563
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)          0.7207     0.1832    3.93  8.9e-05 ***
## LastAuthorFemale1    0.2240     0.0493    4.55  6.1e-06 ***
## Year1997            -0.0853     0.2172   -0.39    0.695
## Year1998             0.0569     0.2312    0.25    0.806
## Year1999             0.2695     0.2114    1.27    0.203
## Year2000             0.3312     0.2230    1.49    0.138
## Year2001             0.3954     0.2339    1.69    0.091 .
## Year2002             0.0348     0.2223    0.16    0.876
## Year2003             0.1391     0.2044    0.68    0.496
## Year2004             0.0598     0.2119    0.28    0.778
## Year2005             0.0862     0.2047    0.42    0.674
## Year2006            -0.0287     0.2074   -0.14    0.890
```

```
## Year2007            0.0406     0.1996    0.20    0.839
## Year2008            0.1846     0.2001    0.92    0.356
## Year2009            0.3128     0.2016    1.55    0.121
## Year2010            0.2469     0.2026    1.22    0.223
## Year2011            0.2186     0.2003    1.09    0.275
## Year2012            0.1364     0.2009    0.68    0.497
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.849
## Multiple R-squared:  0.041,  Adjusted R-squared:  0.026
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  90 weights are ~= 1. The remaining 1019 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.491   0.890   0.946   0.925   0.981   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          9.02e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Sample size for the above analysis:  1109"
## [1] ""
## [1] ""
## [1] "####################################"
## [1] "Analysis of AJSC 2921"
## [1] "####################################"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  229  258  228  236  285  206  247  165  166  200  303  346  369  518  487
## 2011 2012
##  459  432
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  189  210  189  202  239  161  206  146  143  180  255  298  317  464  419
## 2011 2012
```

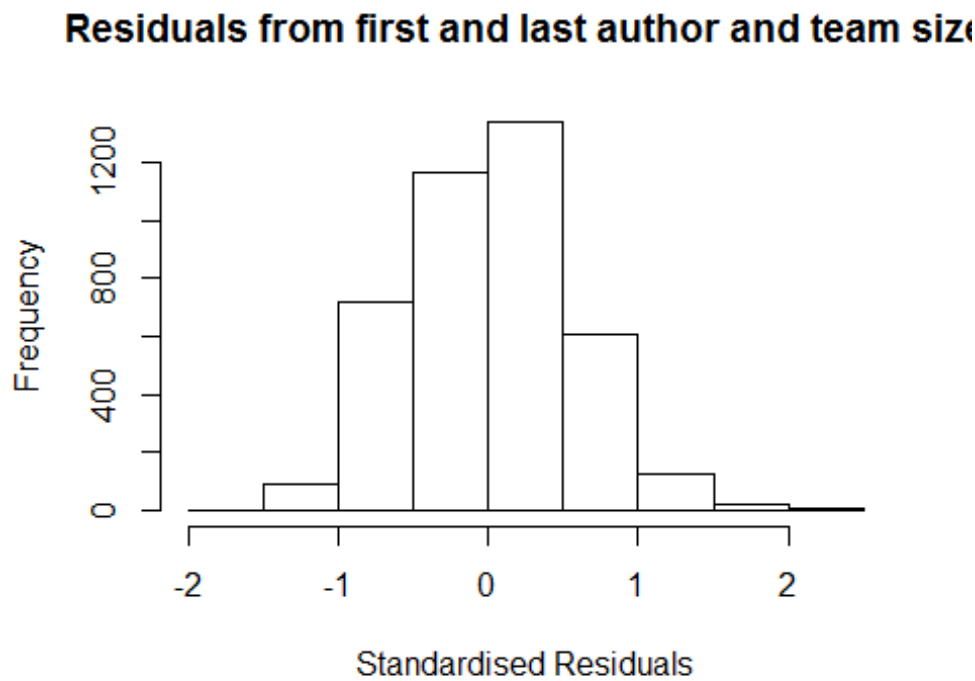
```
##  406  375
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  183  195  170  195  223  151  195  137  125  161  233  271  287  434  386
## 2011 2012
##  373  350
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
##  Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 78, df = 16, p-value = 4e-10
```




```
##
##  Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 2.1, df = 1, p-value = 0.1
```

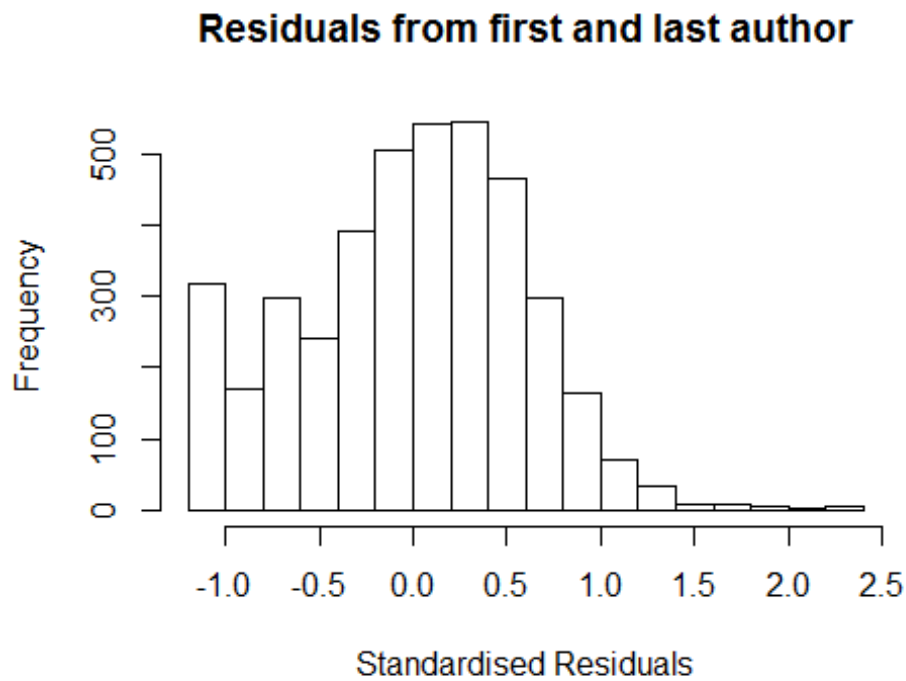



```
## [1] "Female first author team size 2018 geometric mean: 2.78393145459949"
## [1] "Male first author team size 2018 geometric mean: 2.62967136465921"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 16000, p-value = 0.7
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 2.6486942705534"
## [1] "Male last author team size 2018 geometric mean: 2.86986902236217"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 17000, p-value = 0.3
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.181  1           1.087
## LastAuthorFemale  1.208  1           1.099
## UniqueAuthors     1.181  4           1.021
## Year              1.159 16           1.005
```



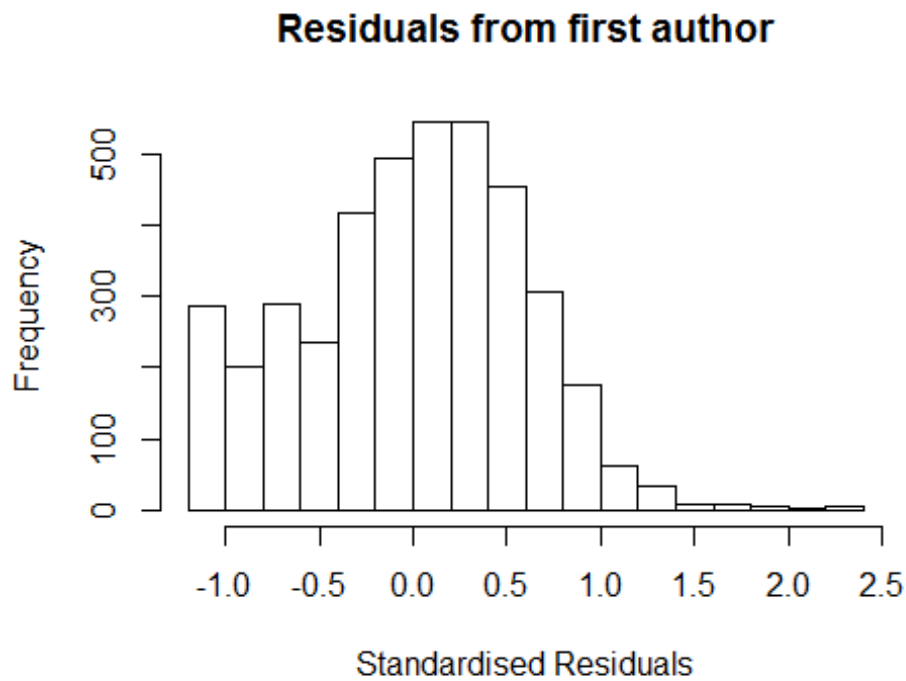

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
UniqueAuthors +
##     Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##     k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.5076 -0.3804  0.0184  0.3815  2.4176
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)          0.8971     0.0640   14.02  < 2e-16 ***
## FirstAuthorFemale1   0.0585     0.0207    2.82  0.00481 **
## LastAuthorFemale1    0.0305     0.0199    1.54  0.12448
## UniqueAuthors2       0.2402     0.0262    9.17  < 2e-16 ***
## UniqueAuthors3       0.4047     0.0281   14.40  < 2e-16 ***
## UniqueAuthors4       0.4629     0.0305   15.19  < 2e-16 ***
## UniqueAuthors5       0.5520     0.0256   21.58  < 2e-16 ***
## Year1997            -0.1741     0.0781   -2.23  0.02593 *
## Year1998            -0.0517     0.0782   -0.66  0.50843
## Year1999            -0.0715     0.0722   -0.99  0.32230
```

```
## Year2000            -0.0705     0.0699   -1.01  0.31374
## Year2001            -0.1535     0.0851   -1.80  0.07136 .
## Year2002            -0.1870     0.0731   -2.56  0.01053 *
## Year2003            -0.1827     0.0748   -2.44  0.01467 *
## Year2004            -0.1255     0.0756   -1.66  0.09703 .
## Year2005            -0.1359     0.0722   -1.88  0.06001 .
## Year2006            -0.2704     0.0705   -3.83  0.00013 ***
## Year2007            -0.2320     0.0687   -3.38  0.00075 ***
## Year2008            -0.2463     0.0673   -3.66  0.00026 ***
## Year2009            -0.2059     0.0681   -3.02  0.00251 **
## Year2010            -0.2264     0.0681   -3.32  0.00089 ***
## Year2011            -0.2703     0.0679   -3.98    7e-05 ***
## Year2012            -0.2055     0.0688   -2.99  0.00283 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.567
## Multiple R-squared:  0.13,   Adjusted R-squared:  0.125
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
##  359 weights are ~= 1. The remaining 3710 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0295  0.8720  0.9500  0.9100  0.9840  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          2.46e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.132  1           1.064
## LastAuthorFemale  1.134  1           1.065
## Year              1.038 16           1.001
```



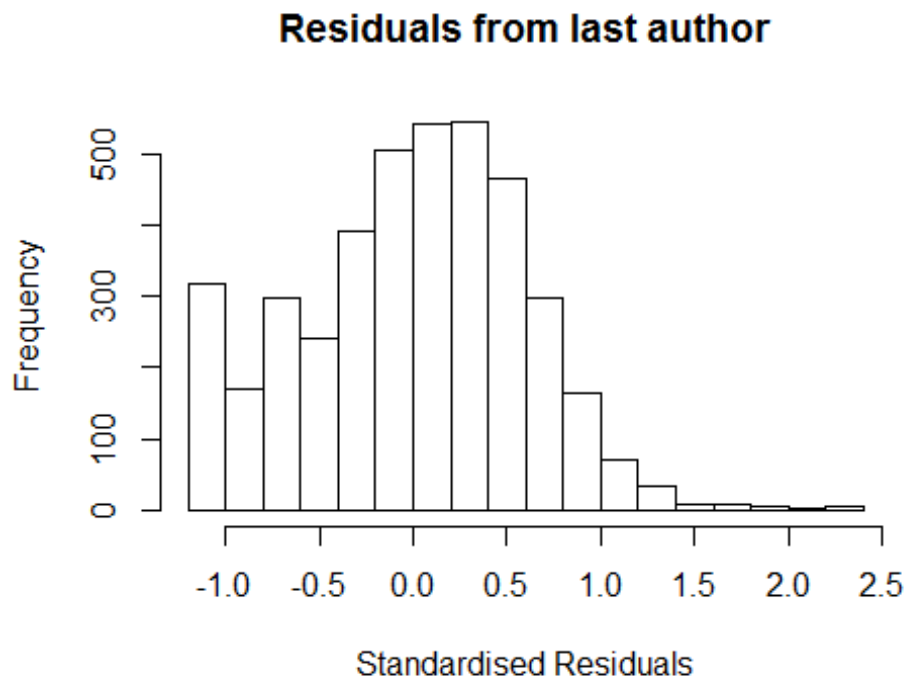

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.1648 -0.4064  0.0422  0.4132  2.3921
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)          1.11116    0.05980   18.58   <2e-16 ***
## FirstAuthorFemale1  0.05363    0.02255    2.38   0.0174 *
## LastAuthorFemale1  -0.04088    0.02114   -1.93   0.0532 .
## Year1997           -0.11913    0.07647   -1.56   0.1194
## Year1998           -0.02410    0.07668   -0.31   0.7533
## Year1999           -0.00409    0.06973   -0.06   0.9532
## Year2000           -0.02850    0.06868   -0.41   0.6782
## Year2001           -0.09829    0.08355   -1.18   0.2395
## Year2002           -0.11415    0.07235   -1.58   0.1147
## Year2003           -0.07333    0.07487   -0.98   0.3274
## Year2004           -0.00782    0.07314   -0.11   0.9148
## Year2005           -0.01213    0.07102   -0.17   0.8644
```

```
## Year2006           -0.18676    0.06987   -2.67   0.0075 **
## Year2007           -0.12080    0.06612   -1.83   0.0678 .
## Year2008           -0.13736    0.06601   -2.08   0.0375 *
## Year2009           -0.10776    0.06616   -1.63   0.1034
## Year2010           -0.11271    0.06638   -1.70   0.0896 .
## Year2011           -0.16201    0.06700   -2.42   0.0157 *
## Year2012           -0.03115    0.06622   -0.47   0.6381
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.593
## Multiple R-squared:  0.0106, Adjusted R-squared:  0.00618
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
##  341 weights are ~= 1. The remaining 3728 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0675  0.8690  0.9470  0.9080  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          2.46e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.014  1           1.007
## Year              1.014 16           1.000
```




```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.1312 -0.3962  0.0407  0.4107  2.3848
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)          1.09628    0.05895   18.60   <2e-16 ***
## FirstAuthorFemale1  0.03493    0.02143    1.63   0.1032
## Year1997           -0.11796    0.07640   -1.54   0.1227
## Year1998           -0.02343    0.07663   -0.31   0.7598
## Year1999           -0.00573    0.06966   -0.08   0.9345
## Year2000           -0.02992    0.06863   -0.44   0.6629
## Year2001           -0.09994    0.08362   -1.20   0.2321
## Year2002           -0.11272    0.07237   -1.56   0.1194
## Year2003           -0.06940    0.07479   -0.93   0.3535
## Year2004           -0.00577    0.07306   -0.08   0.9371
## Year2005           -0.01129    0.07087   -0.16   0.8734
## Year2006           -0.18408    0.06972   -2.64   0.0083 **
```

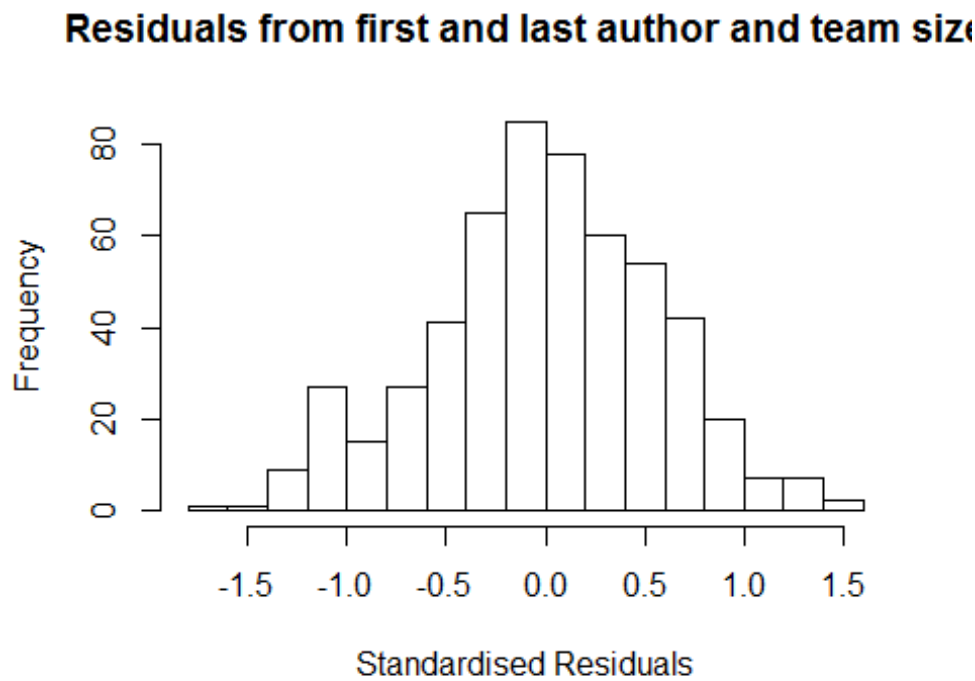
```
## Year2007           -0.11814    0.06603   -1.79   0.0737 .
## Year2008           -0.13464    0.06592   -2.04   0.0412 *
## Year2009           -0.10489    0.06600   -1.59   0.1121
## Year2010           -0.10955    0.06624   -1.65   0.0983 .
## Year2011           -0.15774    0.06688   -2.36   0.0184 *
## Year2012           -0.02712    0.06605   -0.41   0.6814
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.595
## Multiple R-squared:  0.00977,    Adjusted R-squared:  0.00562
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
##  336 weights are ~= 1. The remaining 3733 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0718  0.8700  0.9470  0.9090  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          2.46e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##                   GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.013  1           1.007
## Year             1.013 16           1.000
```




```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.1339 -0.4028  0.0429  0.4147  2.4003
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.13393    0.05911   19.18   <2e-16 ***
## LastAuthorFemale1 -0.01820    0.02028   -0.90   0.3694
## Year1997          -0.11440    0.07633   -1.50   0.1340
## Year1998          -0.02163    0.07680   -0.28   0.7783
## Year1999          -0.00354    0.06985   -0.05   0.9596
## Year2000          -0.02752    0.06884   -0.40   0.6893
## Year2001          -0.09790    0.08347   -1.17   0.2409
## Year2002          -0.11472    0.07263   -1.58   0.1143
## Year2003          -0.07478    0.07484   -1.00   0.3178
## Year2004          -0.00981    0.07331   -0.13   0.8935
## Year2005          -0.01511    0.07114   -0.21   0.8318
## Year2006          -0.19136    0.07015   -2.73   0.0064 **
```

```
## Year2007          -0.12262    0.06629   -1.85   0.0644 .
## Year2008          -0.14125    0.06621   -2.13   0.0330 *
## Year2009          -0.10796    0.06638   -1.63   0.1039
## Year2010          -0.11223    0.06651   -1.69   0.0916 .
## Year2011          -0.16139    0.06713   -2.40   0.0163 *
## Year2012          -0.02893    0.06642   -0.44   0.6631
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.593
## Multiple R-squared:  0.00919,    Adjusted R-squared:  0.00503
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
##  333 weights are ~= 1. The remaining 3736 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0647  0.8680  0.9480  0.9080  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          2.46e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Sample size for the above analysis:  4069"
## [1] ""
## [1] ""
## [1] "####################################"
## [1] "Analysis of AJSC 2922"
## [1] "####################################"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011
##    1    1   23   23   25   24   28   34   27   46   56   63   78   82   72
## 2012
##   85
##
## 1996 1997 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011
##    1    1   21   18   20   20   24   28   27   42   50   54   70   71   64
## 2012
```

```
##   76
##
## 1996 1997 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011
##    1    1   20   16   18   17   23   23   26   39   46   51   64   66   60
## 2012
##   70
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
## [1] "Female first author team size 2018 geometric mean: 2.42873635620953"
## [1] "Male first author team size 2018 geometric mean: 1.73205080756888"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 250, p-value = 0.3
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 2.32180575523557"
## [1] "Male last author team size 2018 geometric mean: 3.11716895429367"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 380, p-value = 0.2
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##                        GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.590e+00  1           1.261
## LastAuthorFemale  2.112e+00  1           1.453
## UniqueAuthors     5.612e+12  4          39.232
## Year              7.051e+12 15           2.681
```

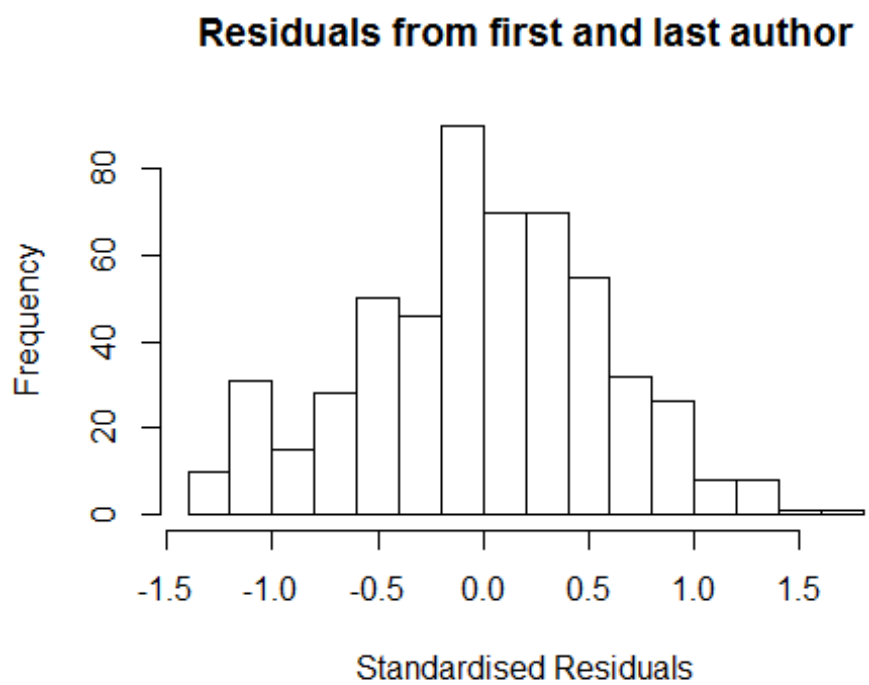



```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
UniqueAuthors +
##     Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##     k.max = 1000))
##  \--> method = "MM"
## Residuals:
##       Min        1Q    Median        3Q       Max
## -1.61e+00 -3.71e-01  8.88e-16  3.87e-01  1.43e+00
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)           1.0049     0.0818   12.29  < 2e-16 ***
## FirstAuthorFemale1    0.1905     0.0828    2.30  0.02185 *
## LastAuthorFemale1     0.0326     0.0603    0.54  0.58896
## UniqueAuthors2        0.0303     0.0761    0.40  0.69080
## UniqueAuthors3        0.2069     0.0751    2.76  0.00607 **
## UniqueAuthors4        0.2521     0.0814    3.10  0.00206 **
## UniqueAuthors5        0.2866     0.0905    3.17  0.00162 **
## Year1997             -0.7951     0.0814   -9.77  < 2e-16 ***
## Year1999             -0.4276     0.1159   -3.69  0.00025 ***
## Year2000              0.0921     0.1691    0.54  0.58625
```

```
## Year2001            -0.2406     0.1109   -2.17  0.03041 *
## Year2002            -0.1887     0.0875   -2.16  0.03151 *
## Year2003            -0.1602     0.1155   -1.39  0.16589
## Year2004            -0.0939     0.0731   -1.28  0.19966
## Year2005            -0.1182     0.1517   -0.78  0.43632
## Year2006            -0.3107     0.0928   -3.35  0.00088 ***
## Year2007            -0.0360     0.1150   -0.31  0.75457
## Year2008            -0.2367     0.1087   -2.18  0.02991 *
## Year2009            -0.2432     0.0934   -2.60  0.00949 **
## Year2010            -0.1527     0.0845   -1.81  0.07126 .
## Year2011            -0.2036     0.0769   -2.65  0.00836 **
## Year2012            -0.1534     0.0954   -1.61  0.10868
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.56
## Multiple R-squared:  0.0809, Adjusted R-squared:  0.0438
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
##  54 weights are ~= 1. The remaining 487 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.390   0.865   0.949   0.902   0.986   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.85e-04          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.390  1           1.179
## LastAuthorFemale  1.751  1           1.323
## Year              1.806 15           1.020

## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
```

```
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.38304 -0.38763  0.00402  0.37729  1.64991
##
## Coefficients:
##                     Estimate Std. Error   t value Pr(>|t|)
## (Intercept)         1.02e+00   8.84e-02  1.16e+01   <2e-16 ***
## FirstAuthorFemale1  2.44e-01   8.62e-02  2.83e+00   0.0049 **
## LastAuthorFemale1  -3.78e-02   5.88e-02 -6.40e-01   0.5204
## Year1997           -5.43e-01   1.76e-07 -3.08e+06   <2e-16 ***
## Year1999           -3.60e-01   1.25e-01 -2.88e+00   0.0042 **
## Year2000            1.55e-01   1.67e-01  9.30e-01   0.3539
## Year2001           -1.63e-01   1.09e-01 -1.49e+00   0.1361
## Year2002           -9.54e-02   6.98e-02 -1.37e+00   0.1724
## Year2003           -5.99e-02   1.17e-01 -5.10e-01   0.6076
## Year2004           -8.66e-04   7.65e-02 -1.00e-02   0.9910
## Year2005            2.17e-02   1.48e-01  1.50e-01   0.8833
## Year2006           -1.81e-01   8.58e-02 -2.11e+00   0.0352 *
## Year2007            4.47e-02   1.08e-01  4.10e-01   0.6791
## Year2008           -1.57e-01   9.86e-02 -1.60e+00   0.1112
## Year2009           -1.42e-01   9.24e-02 -1.54e+00   0.1244
## Year2010           -4.83e-02   7.85e-02 -6.20e-01   0.5385
## Year2011           -1.02e-01   6.72e-02 -1.52e+00   0.1295
## Year2012           -4.24e-02   8.83e-02 -4.80e-01   0.6316
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.565
## Multiple R-squared:  0.0448, Adjusted R-squared:  0.0138
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
##  51 weights are ~= 1. The remaining 490 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.374   0.865   0.944   0.901   0.986   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.85e-04          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
```

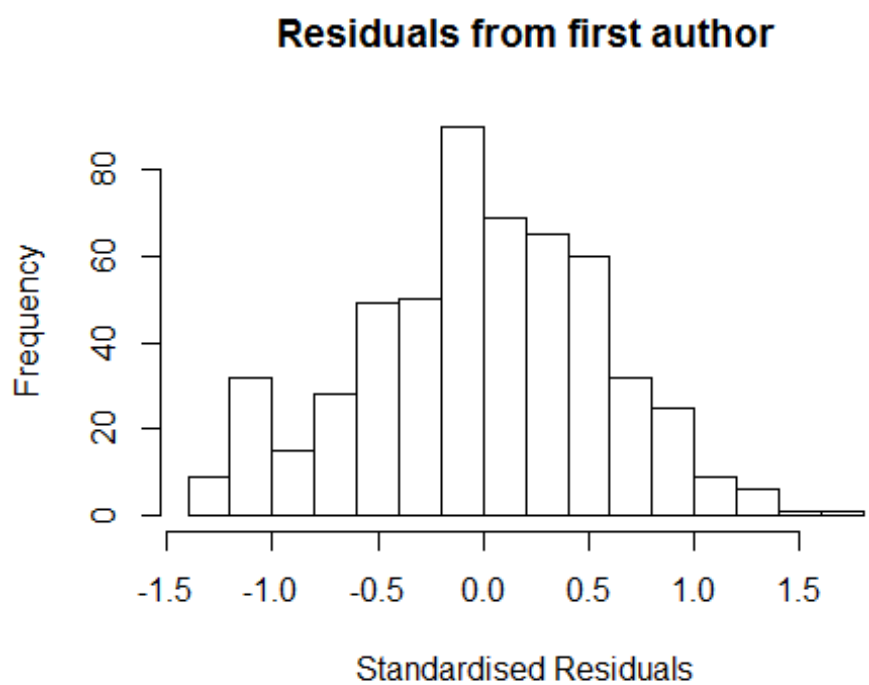
```
##           0        1000           0
##                    psi           subsampling                    cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                   "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"

## Warning in lf.cov(init, x = x): .vcov.avar1: negative diag(<vcov>) fixed
## up; consider 'cov=".vcov.w."' instead

## Warning in lf.cov(init, x = x): diag(.) had 0 or NA entries; non-finite
## result is doubtful
```



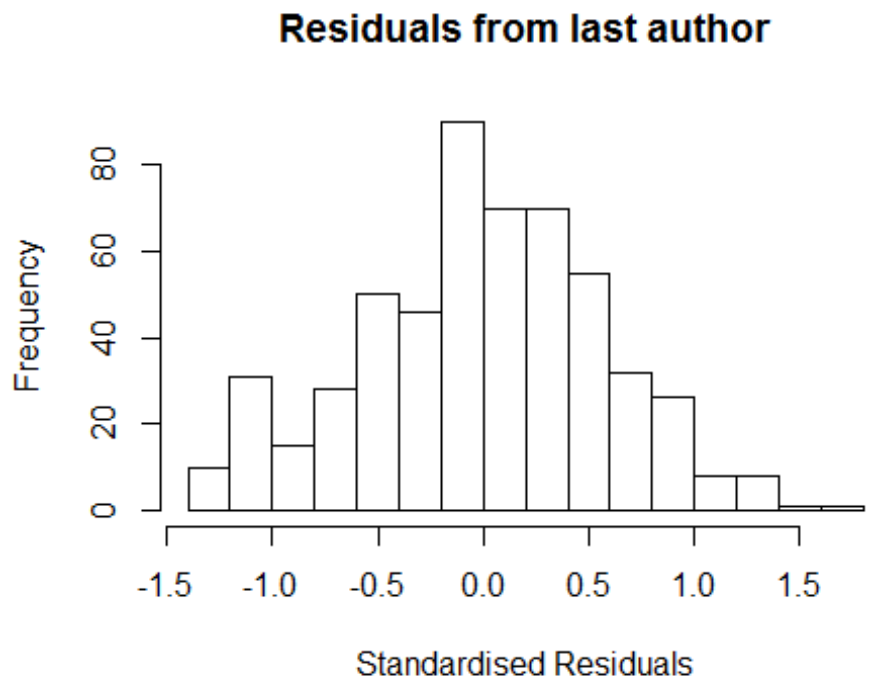

```
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale  NaN  1             NaN
## Year               NaN 15             NaN

## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
```

```
## Residuals:
##       Min        1Q    Median        3Q       Max
## -1.39e+00 -3.93e-01 -3.51e-05  3.87e-01  1.62e+00
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)          1.0033     0.0819   12.26   <2e-16 ***
## FirstAuthorFemale1   0.2247     0.0819    2.74   0.0063 **
## Year1997            -0.5430     0.0000    -Inf   <2e-16 ***
## Year1999            -0.3582     0.1251   -2.86   0.0044 **
## Year2000             0.1589     0.1666    0.95   0.3406
## Year2001            -0.1558     0.1095   -1.42   0.1556
## Year2002            -0.0896     0.0680   -1.32   0.1879
## Year2003            -0.0531     0.1146   -0.46   0.6433
## Year2004             0.0117     0.0752    0.16   0.8764
## Year2005             0.0300     0.1466    0.20   0.8378
## Year2006            -0.1730     0.0855   -2.02   0.0436 *
## Year2007             0.0526     0.1061    0.50   0.6201
## Year2008            -0.1510     0.0977   -1.55   0.1228
## Year2009            -0.1329     0.0892   -1.49   0.1370
## Year2010            -0.0403     0.0773   -0.52   0.6022
## Year2011            -0.0963     0.0666   -1.44   0.1491
## Year2012            -0.0348     0.0865   -0.40   0.6880
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.564
## Multiple R-squared:  0.0444, Adjusted R-squared:  0.0153
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
##  48 weights are ~= 1. The remaining 493 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.388   0.866   0.945   0.901   0.987   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.85e-04          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                    psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
```

```
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"

## Warning in lf.cov(init, x = x): .vcov.avar1: negative diag(<vcov>) fixed
## up; consider 'cov=".vcov.w."' instead

## Warning in lf.cov(init, x = x): diag(.) had 0 or NA entries; non-finite
## result is doubtful
```



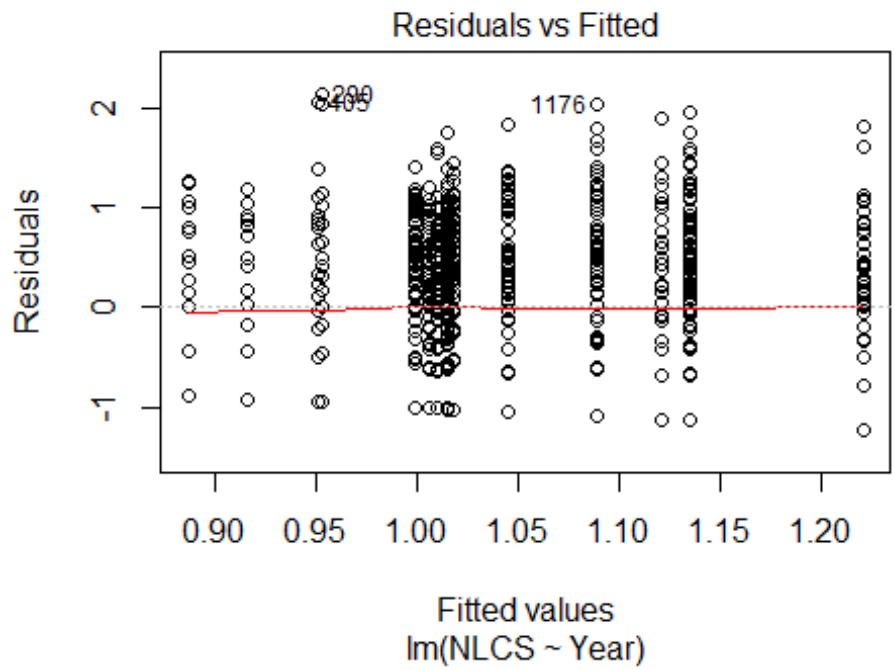

```
##                   GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale   NaN  1             NaN
## Year               NaN 15             NaN
```



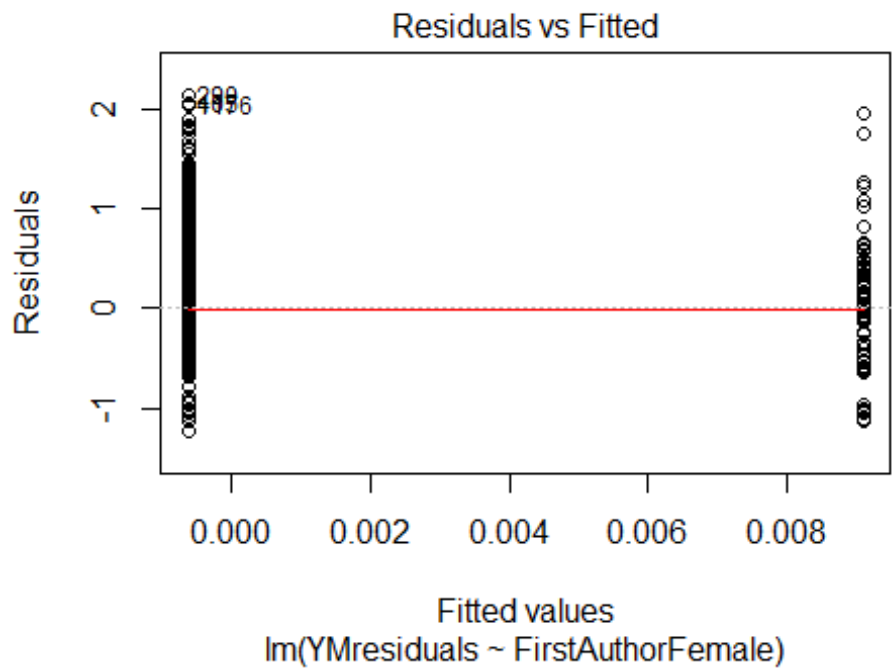

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.3585 -0.4137  0.0141  0.3935  1.4567
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.19166    0.05846   20.38   <2e-16 ***
## LastAuthorFemale1   0.03634    0.05846    0.62   0.5345
## Year1997           -0.54300    0.00000    -Inf   <2e-16 ***
## Year1999           -0.37237    0.12418   -3.00   0.0028 **
## Year2000            0.13055    0.18244    0.72   0.4746
## Year2001           -0.18097    0.11185   -1.62   0.1063
## Year2002           -0.12167    0.07128   -1.71   0.0884 .
## Year2003           -0.08226    0.11140   -0.74   0.4606
## Year2004           -0.00758    0.07532   -0.10   0.9199
## Year2005            0.00551    0.14685    0.04   0.9701
## Year2006           -0.20925    0.08631   -2.42   0.0157 *
## Year2007            0.03240    0.10465    0.31   0.7570
```

```
## Year2008          -0.16409    0.09746   -1.68   0.0929 .
## Year2009          -0.15792    0.09347   -1.69   0.0917 .
## Year2010          -0.06984    0.07903   -0.88   0.3773
## Year2011          -0.11700    0.06849   -1.71   0.0882 .
## Year2012          -0.04272    0.08887   -0.48   0.6309
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.586
## Multiple R-squared:  0.0271, Adjusted R-squared:  -0.00265
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  52 weights are ~= 1. The remaining 489 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.517   0.870   0.948   0.906   0.985   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.85e-04          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Sample size for the above analysis:  541"
## [1] ""
## [1] ""
## [1] "##################################"
## [1] "Analysis of AJSC 2923"
## [1] "##################################"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##   21   69  121  116  112   96  111   98  119  122   57   52  141  179  187
## 2011 2012
##  177  163
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    0    0   25   30   40   28  103   91  107  110   54   46  133  163  169
## 2011 2012
##  154  142
```

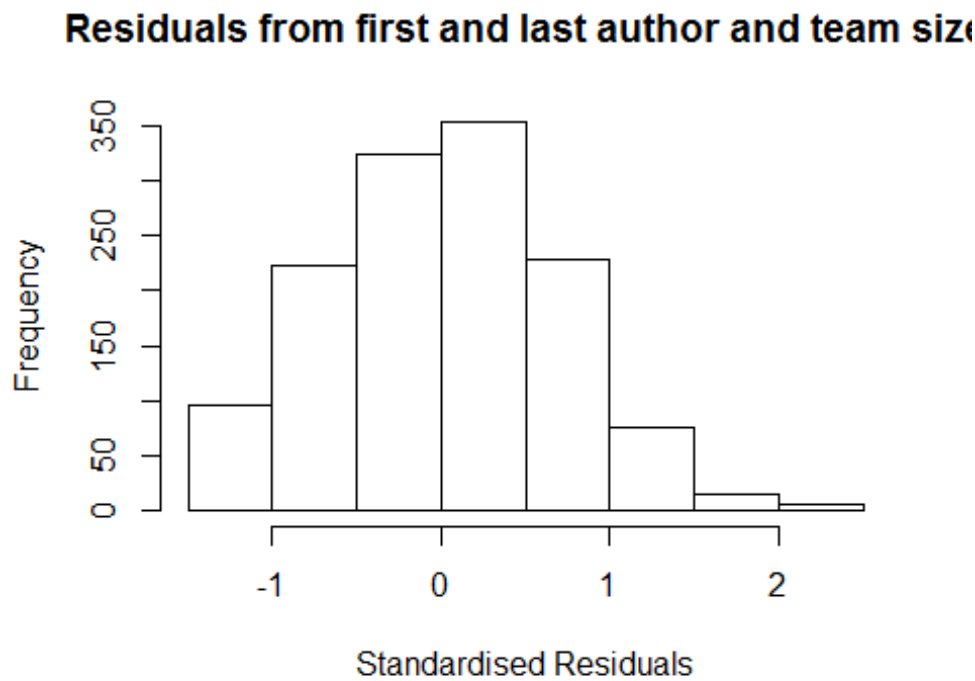
```
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    0    0   25   29   40   27   98   87  104  104   54   43  126  147  158
## 2011 2012
##  146  134
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
##  Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 37, df = 14, p-value = 8e-04
```




```
##
##  Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 0.61, df = 1, p-value = 0.4
```

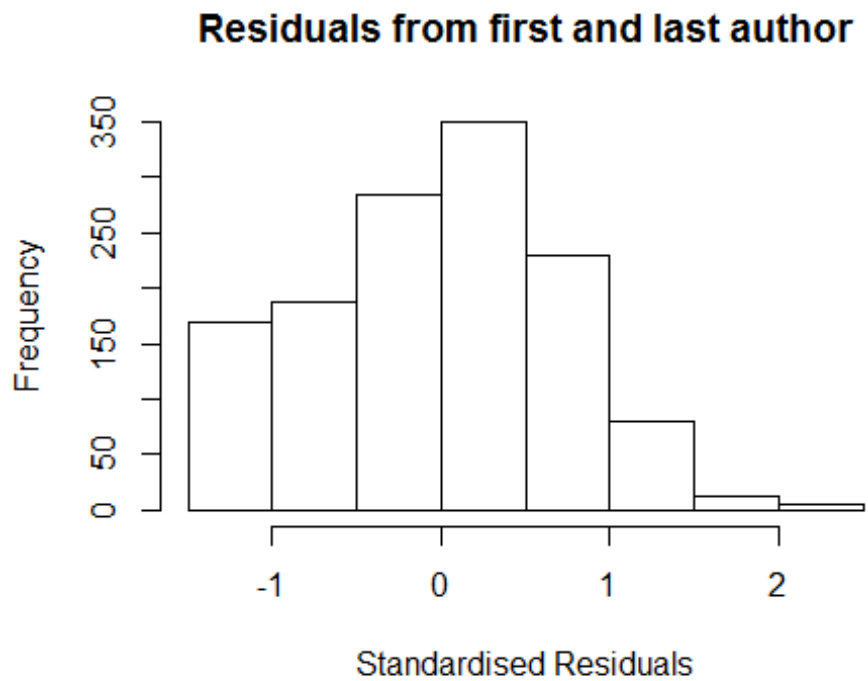



```
## [1] "Female first author team size 2018 geometric mean: 2.45618965918199"
## [1] "Male first author team size 2018 geometric mean: 2.66877524122045"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 1000, p-value = 0.6
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 2.48158054590142"
## [1] "Male last author team size 2018 geometric mean: 2.52509670400684"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 980, p-value = 0.8
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.507  1           1.228
## LastAuthorFemale  1.518  1           1.232
## UniqueAuthors     1.231  4           1.026
## Year              1.249 14           1.008
```



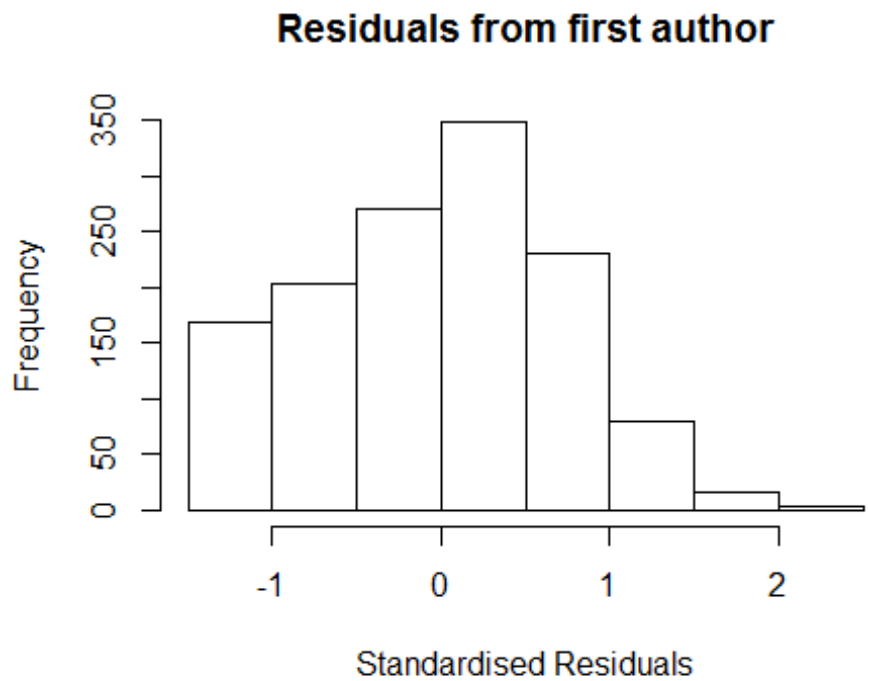

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
UniqueAuthors +
##     Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##     k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.3285 -0.4795  0.0149  0.4872  2.2642
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)          0.8168     0.1731    4.72  2.6e-06 ***
## FirstAuthorFemale1   0.0846     0.0865    0.98   0.3281
## LastAuthorFemale1   -0.1668     0.0690   -2.42   0.0157 *
## UniqueAuthors2       0.2100     0.0500    4.20  2.9e-05 ***
## UniqueAuthors3       0.3506     0.0605    5.80  8.3e-09 ***
## UniqueAuthors4       0.3208     0.0741    4.33  1.6e-05 ***
## UniqueAuthors5       0.2928     0.0751    3.90   0.0001 ***
## Year1999             0.0673     0.2120    0.32   0.7511
## Year2000             0.0319     0.2070    0.15   0.8774
## Year2001            -0.0671     0.2220   -0.30   0.7624
```

```
## Year2002             0.1470     0.1732    0.85   0.3962
## Year2003             0.1252     0.1756    0.71   0.4760
## Year2004             0.1282     0.1772    0.72   0.4695
## Year2005             0.1264     0.1706    0.74   0.4590
## Year2006             0.3147     0.2030    1.55   0.1214
## Year2007             0.1110     0.2070    0.54   0.5917
## Year2008             0.1342     0.1782    0.75   0.4515
## Year2009             0.2162     0.1750    1.24   0.2167
## Year2010             0.0882     0.1709    0.52   0.6058
## Year2011             0.0756     0.1715    0.44   0.6594
## Year2012             0.1059     0.1711    0.62   0.5362
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.734
## Multiple R-squared:  0.0502, Adjusted R-squared:  0.0356
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
##  115 weights are ~= 1. The remaining 1207 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.321   0.879   0.950   0.918   0.984   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          7.56e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.504  1           1.226
## LastAuthorFemale  1.483  1           1.218
## Year              1.080 14           1.003
```



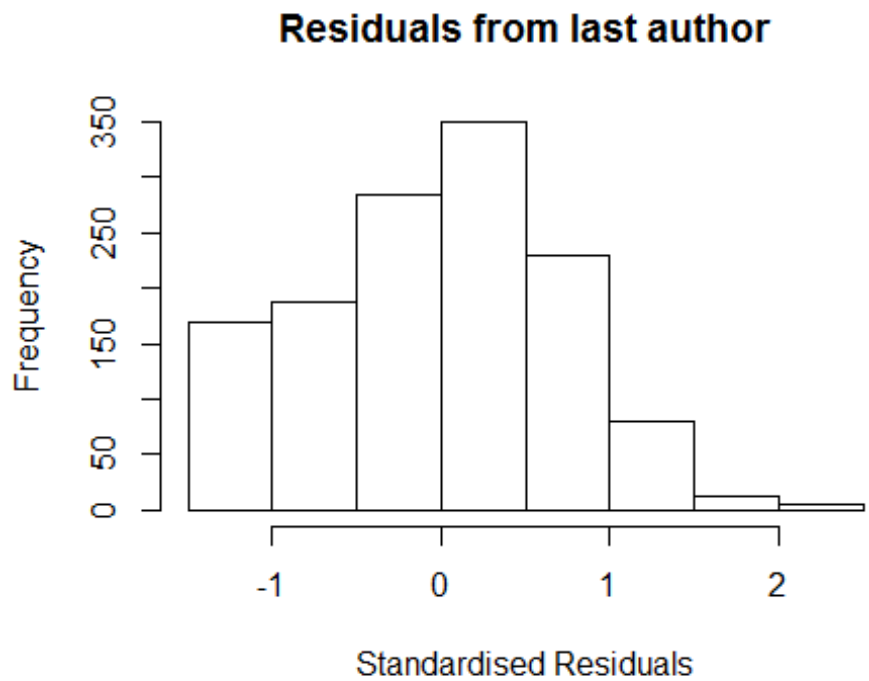

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.2566 -0.5506  0.0315  0.4964  2.1845
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)           0.9778     0.1639    5.97  3.1e-09 ***
## FirstAuthorFemale1    0.1326     0.0867    1.53   0.1262
## LastAuthorFemale1    -0.2229     0.0698   -3.20   0.0014 **
## Year1999              0.0140     0.2167    0.06   0.9485
## Year2000              0.0173     0.2026    0.09   0.9318
## Year2001             -0.0839     0.2096   -0.40   0.6892
## Year2002              0.1303     0.1683    0.77   0.4390
## Year2003              0.1254     0.1708    0.73   0.4632
## Year2004              0.1328     0.1710    0.78   0.4377
## Year2005              0.1130     0.1653    0.68   0.4943
## Year2006              0.3208     0.1993    1.61   0.1078
## Year2007              0.1167     0.2092    0.56   0.5771
```

```
## Year2008             0.1462     0.1718    0.85   0.3951
## Year2009             0.2034     0.1697    1.20   0.2308
## Year2010             0.1170     0.1648    0.71   0.4778
## Year2011             0.0825     0.1659    0.50   0.6189
## Year2012             0.1220     0.1649    0.74   0.4597
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.763
## Multiple R-squared:  0.0138, Adjusted R-squared:  0.00167
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  109 weights are ~= 1. The remaining 1213 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.393   0.858   0.951   0.922   0.985   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          7.56e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.055  1           1.027
## Year              1.055 14           1.002
```




```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##    Min     1Q Median     3Q    Max
## -1.224 -0.558  0.031  0.501  2.181
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)          0.91500    0.16682    5.48    5e-08 ***
## FirstAuthorFemale1 -0.00647    0.07219   -0.09     0.93
## Year1999           -0.00364    0.21978   -0.02     0.99
## Year2000            0.00473    0.20716    0.02     0.98
## Year2001           -0.08216    0.21263   -0.39     0.70
## Year2002            0.11919    0.17142    0.70     0.49
## Year2003            0.11306    0.17366    0.65     0.52
## Year2004            0.12321    0.17431    0.71     0.48
## Year2005            0.10399    0.16873    0.62     0.54
## Year2006            0.31587    0.20326    1.55     0.12
## Year2007            0.12877    0.21045    0.61     0.54
## Year2008            0.13862    0.17498    0.79     0.43
```

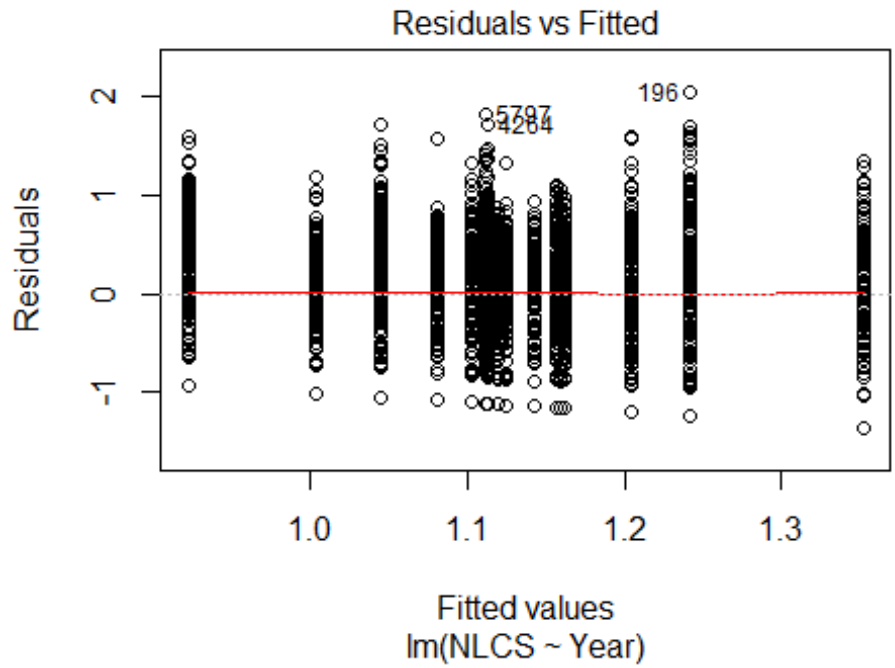
```
## Year2009            0.18717    0.17280    1.08     0.28
## Year2010            0.10745    0.16802    0.64     0.52
## Year2011            0.07823    0.16920    0.46     0.64
## Year2012            0.11745    0.16813    0.70     0.48
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.764
## Multiple R-squared:  0.00778,    Adjusted R-squared:  -0.00362
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  110 weights are ~= 1. The remaining 1212 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.396   0.857   0.950   0.922   0.985   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          7.56e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##                   GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.044  1           1.022
## Year             1.044 14           1.002
```



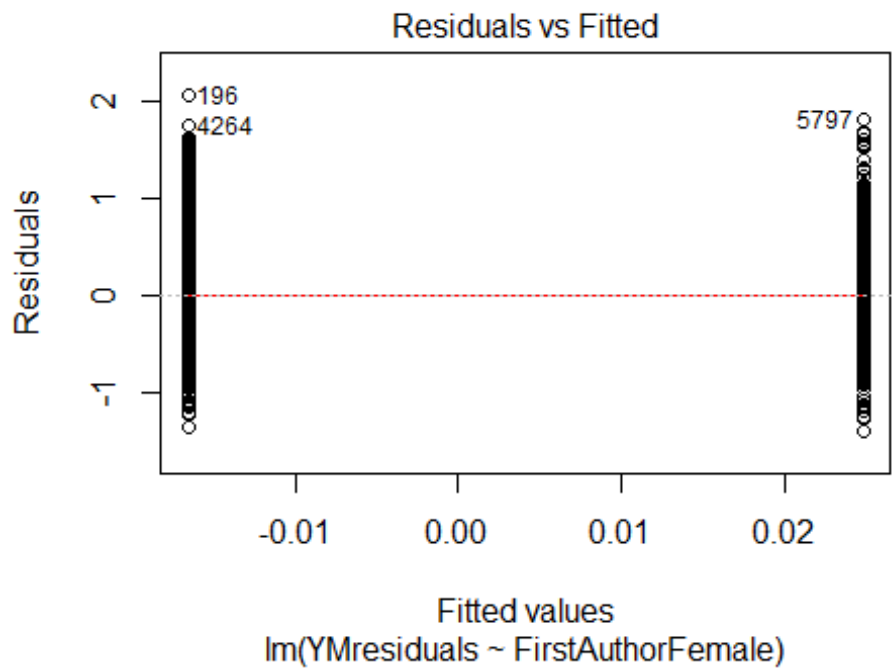

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.2053 -0.5458  0.0307  0.4864  2.1875
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)          1.0445     0.1597    6.54  8.9e-11 ***
## LastAuthorFemale1   -0.1649     0.0595   -2.77   0.0057 **
## Year1999             0.0189     0.2161    0.09   0.9302
## Year2000             0.0102     0.2024    0.05   0.9599
## Year2001            -0.0701     0.2095   -0.33   0.7382
## Year2002             0.1353     0.1686    0.80   0.4225
## Year2003             0.1301     0.1708    0.76   0.4463
## Year2004             0.1399     0.1711    0.82   0.4138
## Year2005             0.1196     0.1656    0.72   0.4701
## Year2006             0.3256     0.1996    1.63   0.1031
## Year2007             0.1220     0.2093    0.58   0.5602
## Year2008             0.1480     0.1720    0.86   0.3898
```

```
## Year2009            0.2087     0.1698    1.23   0.2193
## Year2010            0.1212     0.1648    0.73   0.4625
## Year2011            0.0863     0.1660    0.52   0.6035
## Year2012            0.1281     0.1650    0.78   0.4379
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.764
## Multiple R-squared:  0.0122, Adjusted R-squared:  0.000886
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  110 weights are ~= 1. The remaining 1212 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.392   0.860   0.952   0.922   0.985   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          7.56e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                  psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Sample size for the above analysis:  1322"
## [1] ""
## [1] ""
## [1] "####################################"
## [1] "Analysis of AJSC 3000"
## [1] "####################################"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  326  253  289  294  238  239  215  267  265  298  367  371  358  235  338
## 2011 2012
##  375  417
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  225  166  164  183  120   94  151  197  193  227  273  289  258  182  280
## 2011 2012
##  294  320
##
```

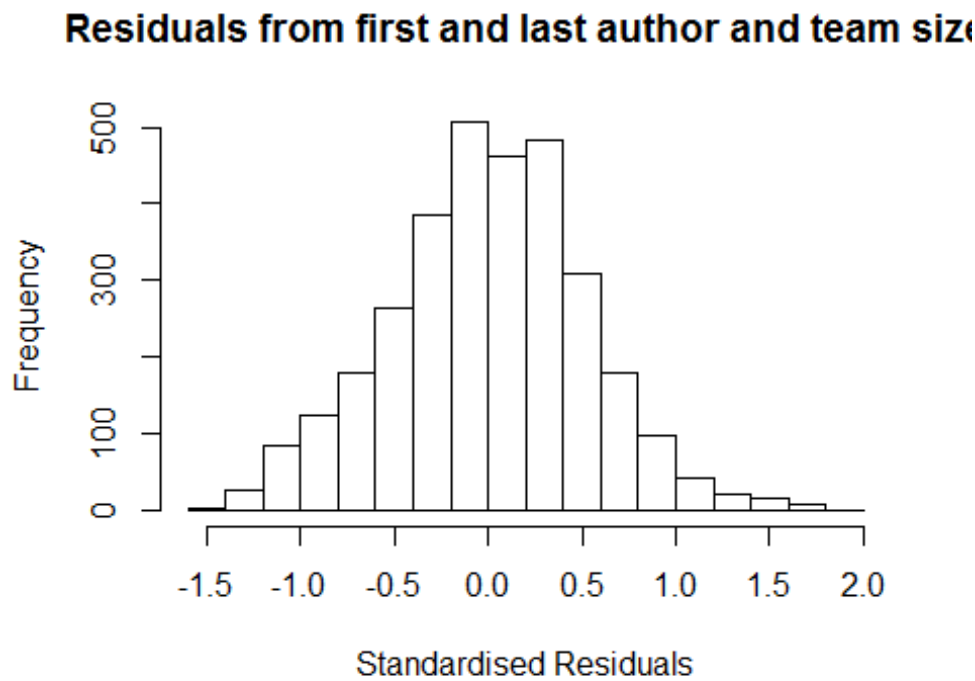
```
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  199  145  142  158  102   79  134  169  164  203  237  264  228  174  256
## 2011 2012
##  256  276
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
##  Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 240, df = 16, p-value <2e-16
```




```
##
##  Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 2.1, df = 1, p-value = 0.1
```

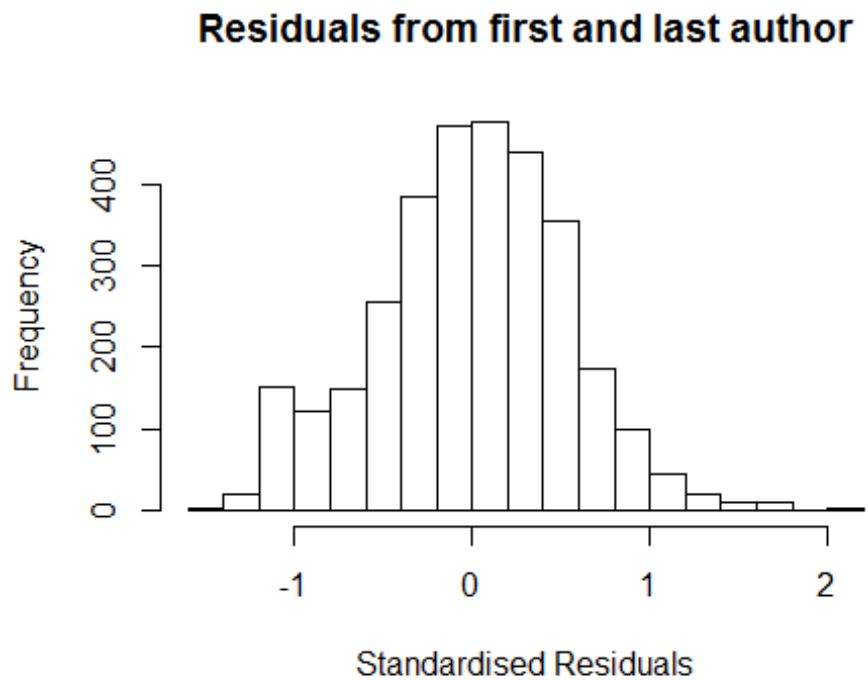



```
## [1] "Female first author team size 2018 geometric mean: 3.67649888132225"
## [1] "Male first author team size 2018 geometric mean: 3.47767065989217"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 13000, p-value = 0.7
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 3.63554844775662"
## [1] "Male last author team size 2018 geometric mean: 3.52540659010616"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 13000, p-value = 1
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.097  1           1.047
## LastAuthorFemale  1.082  1           1.040
## UniqueAuthors     1.126  4           1.015
## Year              1.184 16           1.005
```



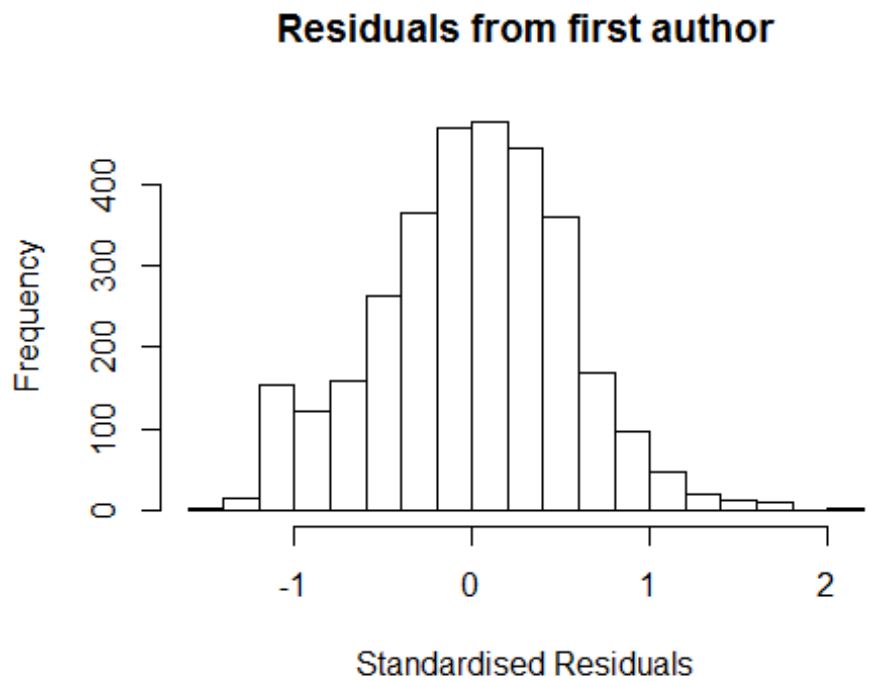

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
UniqueAuthors +
##     Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##     k.max = 1000))
##  \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.57010 -0.33744  0.00843  0.34379  1.94572
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)          0.9121     0.0614   14.86  < 2e-16 ***
## FirstAuthorFemale1   0.0191     0.0198    0.96   0.3362
## LastAuthorFemale1    0.0655     0.0208    3.15   0.0017 **
## UniqueAuthors2       0.2821     0.0347    8.12  6.4e-16 ***
## UniqueAuthors3       0.2881     0.0332    8.68  < 2e-16 ***
## UniqueAuthors4       0.3561     0.0360    9.91  < 2e-16 ***
## UniqueAuthors5       0.4171     0.0309   13.49  < 2e-16 ***
## Year1997             0.1754     0.0746    2.35   0.0189 *
## Year1998            -0.0947     0.0709   -1.34   0.1817
## Year1999            -0.0597     0.0672   -0.89   0.3741
```

```
## Year2000            -0.2089     0.0724   -2.89   0.0039 **
## Year2001            -0.0465     0.0752   -0.62   0.5360
## Year2002            -0.0917     0.0665   -1.38   0.1681
## Year2003            -0.0860     0.0650   -1.32   0.1861
## Year2004            -0.1016     0.0657   -1.55   0.1221
## Year2005            -0.1133     0.0630   -1.80   0.0723 .
## Year2006            -0.1480     0.0619   -2.39   0.0168 *
## Year2007            -0.0594     0.0610   -0.97   0.3300
## Year2008            -0.0308     0.0661   -0.47   0.6418
## Year2009            -0.1290     0.0706   -1.83   0.0676 .
## Year2010            -0.3090     0.0747   -4.13  3.6e-05 ***
## Year2011            -0.1935     0.0717   -2.70   0.0070 **
## Year2012            -0.1172     0.0667   -1.76   0.0787 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.505
## Multiple R-squared:  0.101,  Adjusted R-squared:  0.0945
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
##  261 weights are ~= 1. The remaining 2925 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.105   0.869   0.951   0.902   0.985   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          3.14e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.075  1           1.037
## LastAuthorFemale  1.075  1           1.037
## Year              1.077 16           1.002
```



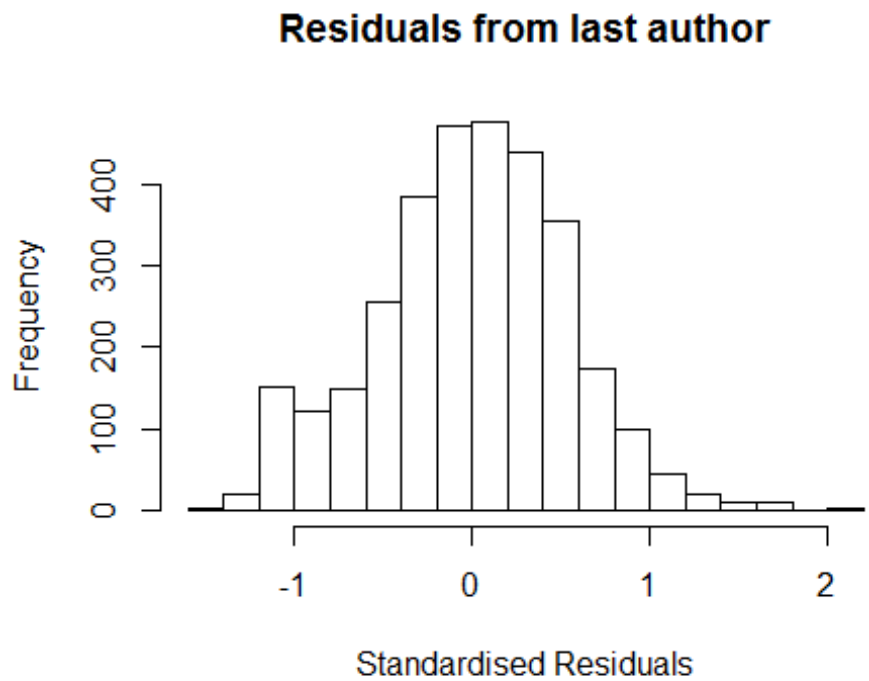

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.4502 -0.3452  0.0245  0.3622  2.0892
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)           1.1858     0.0604   19.64  < 2e-16 ***
## FirstAuthorFemale1    0.0321     0.0204    1.58  0.11513
## LastAuthorFemale1     0.0462     0.0214    2.16  0.03103 *
## Year1997              0.1860     0.0784    2.37  0.01774 *
## Year1998             -0.1229     0.0755   -1.63  0.10361
## Year1999             -0.0574     0.0706   -0.81  0.41613
## Year2000             -0.2077     0.0767   -2.71  0.00684 **
## Year2001             -0.0472     0.0794   -0.59  0.55231
## Year2002             -0.0533     0.0700   -0.76  0.44674
## Year2003             -0.0654     0.0685   -0.95  0.34006
## Year2004             -0.0647     0.0685   -0.94  0.34525
## Year2005             -0.0929     0.0662   -1.40  0.16017
```

```
## Year2006            -0.1373     0.0658   -2.09  0.03696 *
## Year2007            -0.0434     0.0642   -0.68  0.49869
## Year2008             0.0034     0.0692    0.05  0.96079
## Year2009            -0.1112     0.0749   -1.48  0.13787
## Year2010            -0.3019     0.0785   -3.85  0.00012 ***
## Year2011            -0.1643     0.0749   -2.19  0.02836 *
## Year2012            -0.1002     0.0711   -1.41  0.15868
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.519
## Multiple R-squared:  0.0332, Adjusted R-squared:  0.0277
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
##  266 weights are ~= 1. The remaining 2920 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0686  0.8690  0.9500  0.9000  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          3.14e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.039  1           1.020
## Year              1.039 16           1.001
```




```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.4251 -0.3478  0.0205  0.3580  2.0820
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)           1.1930     0.0602   19.80  < 2e-16 ***
## FirstAuthorFemale1    0.0435     0.0201    2.16  0.03068 *
## Year1997              0.1886     0.0784    2.41  0.01617 *
## Year1998             -0.1173     0.0755   -1.55  0.12016
## Year1999             -0.0568     0.0706   -0.80  0.42100
## Year2000             -0.2064     0.0767   -2.69  0.00718 **
## Year2001             -0.0453     0.0789   -0.57  0.56598
## Year2002             -0.0532     0.0702   -0.76  0.44876
## Year2003             -0.0647     0.0685   -0.94  0.34518
## Year2004             -0.0651     0.0685   -0.95  0.34222
## Year2005             -0.0906     0.0661   -1.37  0.17094
## Year2006             -0.1313     0.0656   -2.00  0.04536 *
```

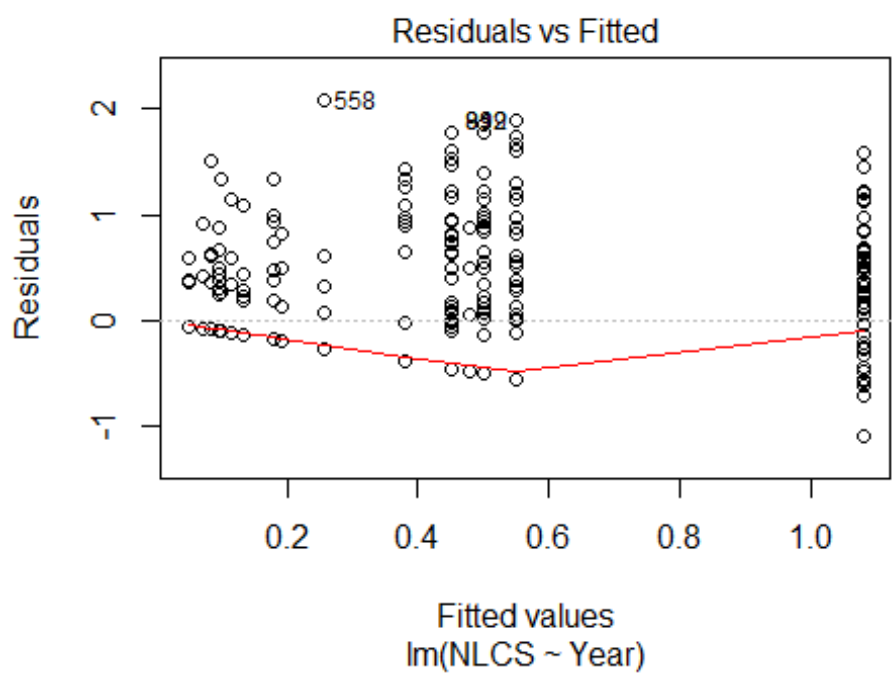
```
## Year2007            -0.0396     0.0641   -0.62  0.53662
## Year2008             0.0049     0.0692    0.07  0.94349
## Year2009            -0.1077     0.0748   -1.44  0.15031
## Year2010            -0.2986     0.0784   -3.81  0.00014 ***
## Year2011            -0.1547     0.0746   -2.08  0.03801 *
## Year2012            -0.0940     0.0710   -1.32  0.18606
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.52
## Multiple R-squared:  0.0318, Adjusted R-squared:  0.0266
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
##  259 weights are ~= 1. The remaining 2927 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0725  0.8710  0.9500  0.9010  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          3.14e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##                   GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.038  1           1.019
## Year             1.038 16           1.001
```




```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.4345 -0.3413  0.0211  0.3599  2.0789
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.19610    0.05933   20.16  < 2e-16 ***
## LastAuthorFemale1   0.05484    0.02106    2.60  0.00926 **
## Year1997            0.18360    0.07820    2.35  0.01895 *
## Year1998           -0.12718    0.07527   -1.69  0.09120 .
## Year1999           -0.06019    0.07039   -0.86  0.39251
## Year2000           -0.21141    0.07654   -2.76  0.00578 **
## Year2001           -0.04844    0.07896   -0.61  0.53961
## Year2002           -0.05505    0.06981   -0.79  0.43041
## Year2003           -0.06474    0.06853   -0.94  0.34486
## Year2004           -0.06485    0.06830   -0.95  0.34243
## Year2005           -0.09196    0.06606   -1.39  0.16404
## Year2006           -0.13841    0.06562   -2.11  0.03501 *
```

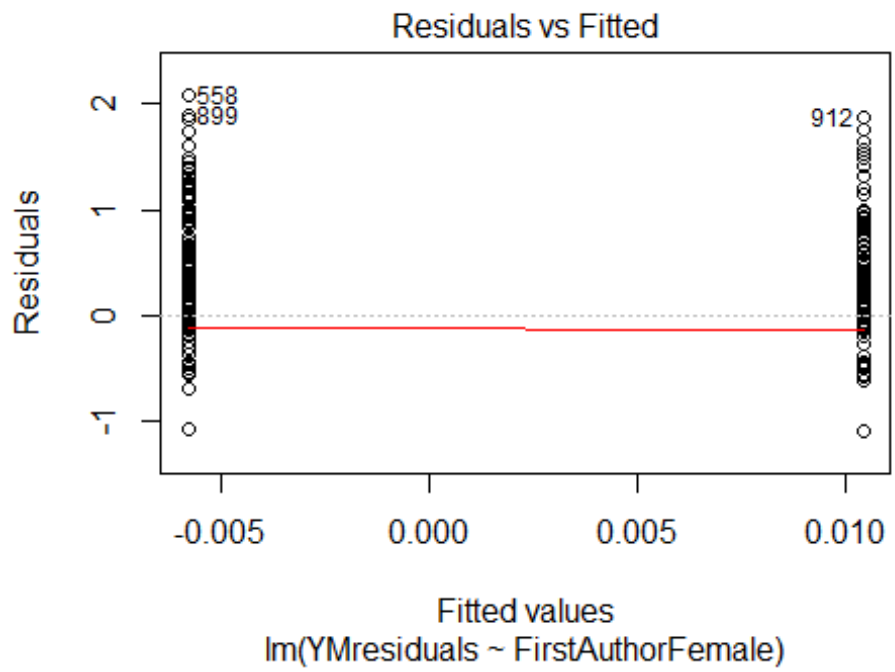
```
## Year2007          -0.04275    0.06415   -0.67  0.50522
## Year2008           0.00354    0.06914    0.05  0.95916
## Year2009          -0.10866    0.07485   -1.45  0.14669
## Year2010          -0.29936    0.07862   -3.81  0.00014 ***
## Year2011          -0.16277    0.07496   -2.17  0.02997 *
## Year2012          -0.09816    0.07105   -1.38  0.16717
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.519
## Multiple R-squared:  0.0325, Adjusted R-squared:  0.0273
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
##  260 weights are ~= 1. The remaining 2926 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0722  0.8690  0.9500  0.9000  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          3.14e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Sample size for the above analysis:  3186"
## [1] ""
## [1] ""
## [1] "####################################"
## [1] "Analysis of AJSC 3001"
## [1] "####################################"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##   92   81   60   72   56   52   39   26   29   17   18    8   46   63  105
## 2011 2012
##   92  113
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##   83   68   53   63   47   44   35   22   18   16   14    6   36   59   79
## 2011 2012
```

```
##   75   83
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##   78   64   53   59   45   42   34   21   17   15   13    5   35   53   69
## 2011 2012
##   64   60
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
##  Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 400, df = 16, p-value <2e-16
```




```
##
##  Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 5.6, df = 1, p-value = 0.02

## [1] "Female first author team size 2018 geometric mean: 4.57747808070587"
## [1] "Male first author team size 2018 geometric mean: 4.08754221130945"

## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties
```

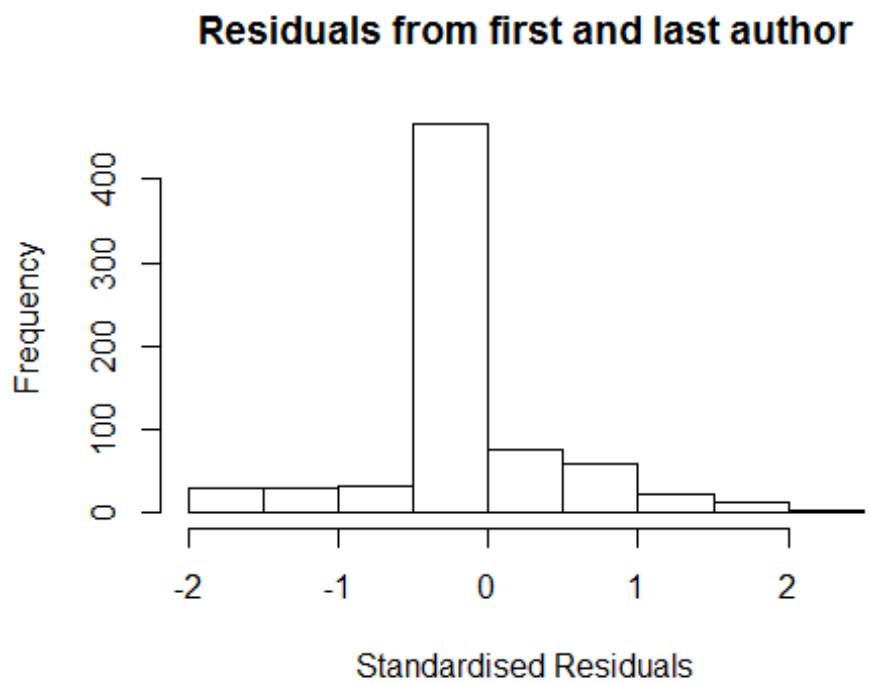
```
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 840, p-value = 0.4
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 4.58936998498914"
## [1] "Male last author team size 2018 geometric mean: 4.13995196247413"

## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties

##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 780, p-value = 0.6
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"

## Warning in lmrob.S(x, y, control = control, mf = mf): S-estimated scale ==
## 0: Probably exact fit; check your data

## Warning in lmrob.S(x, y, control = control, mf = mf): diag(.) had 0 or NA
## entries; non-finite result is doubtful
```




```
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale   NaN  1             NaN
## LastAuthorFemale    NaN  1             NaN
## Year                NaN 16             NaN

## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "S"
## Residuals:
##    Min     1Q Median     3Q    Max
##  -1.72   0.00   0.00   0.00   2.45
##
## Exact fit detected
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)            0.000      0.000      NA       NA
## FirstAuthorFemale1     0.000      0.000      NA       NA
## LastAuthorFemale1      0.000      0.000      NA       NA
## Year1997               0.000      0.000      NA       NA
## Year1998               0.000      0.000      NA       NA
```

```
## Year1999              0.000      0.000      NA       NA
## Year2000              0.000      0.000      NA       NA
## Year2001              0.000      0.000      NA       NA
## Year2002              0.000      0.000      NA       NA
## Year2003              0.000      0.000      NA       NA
## Year2004              0.000      0.000      NA       NA
## Year2005              0.587      0.000      NA       NA
## Year2006              0.329      0.000      NA       NA
## Year2007              0.541      0.000      NA       NA
## Year2008              1.034      0.000      NA       NA
## Year2009              1.720      0.000      NA       NA
## Year2010              0.364      0.000      NA       NA
## Year2011              0.000      0.000      NA       NA
## Year2012              0.955      0.000      NA       NA
##
## Robustness weights:
##  329 observations
c(8,10,11,31,32,36,39,54,62,79,85,89,99,118,120,133,136,138,150,162,165,169,1
71,174,176,184,186,189,196,200,207,209,215,216,218,251,258,260,277,292,296,30
0,304,305,306,308,312,331,335,339,342,343,344,345,346,347,367,376,377,385,396
,398,399,414,415,416,417,418,419,420,421,422,423,424,425,427,428,429,430,433,
435,436,437,438,439,440,441,442,443,445,446,447,448,450,451,452,453,454,455,4
56,457,458,459,460,461,462,463,464,465,466,467,468,469,470,471,472,473,474,47
5,476,477,478,479,480,481,482,483,484,485,486,487,488,489,490,491,492,493,494
,495,496,497,499,500,501,502,503,504,505,506,507,509,510,511,512,513,514,515,
516,517,518,519,520,521,522,523,524,525,526,527,528,529,530,531,532,533,534,5
35,536,537,538,539,540,541,542,543,544,546,547,548,549,550,551,552,553,554,55
5,556,557,558,559,560,561,562,563,564,566,567,568,569,570,571,572,573,575,576
,578,579,580,581,582,583,584,585,586,587,588,589,590,591,592,593,594,595,596,
597,598,599,600,601,602,603,605,606,607,608,609,610,611,614,616,620,621,623,6
26,629,631,634,638,639,640,642,643,644,646,648,649,650,651,652,657,663,668,66
9,670,671,672,673,674,675,676,677,678,679,680,681,682,683,684,685,686,687,688
,689,690,691,692,693,694,695,696,697,698,699,700,701,702,703,704,705,706,707,
708,709,710,711,712,713,714,715,716,717,718,719,722,723,724,725,726,727)
##   are outliers with |weight| = 0 ( < 0.00014);
##  398 weights are ~= 1.
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.38e-04          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
```

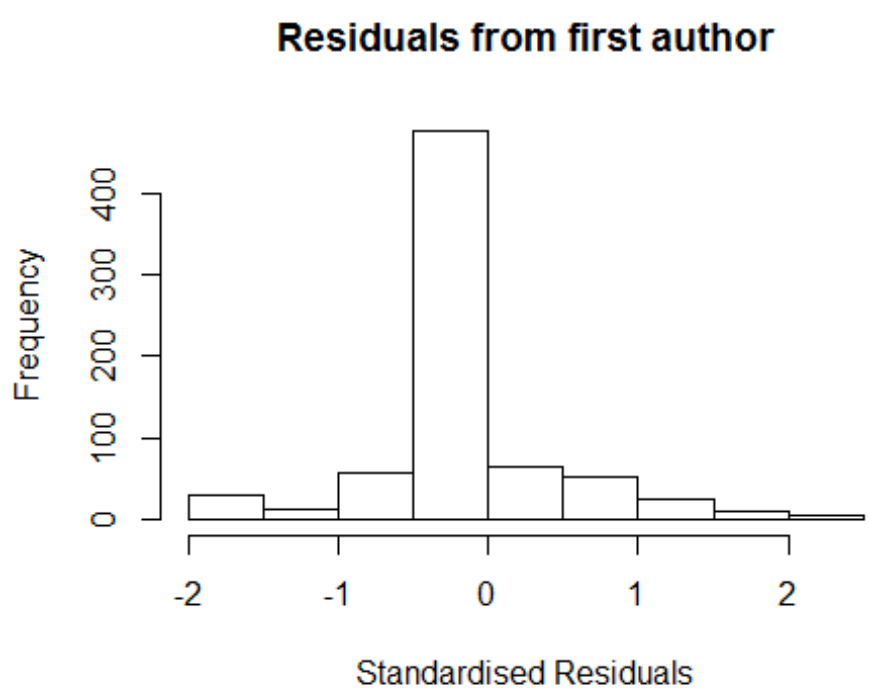
```
##                    "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"

## Warning in lmrob.S(x, y, control = control, mf = mf): S-estimated scale ==
## 0: Probably exact fit; check your data

## Warning in lmrob.S(x, y, control = control, mf = mf): diag(.) had 0 or NA
## entries; non-finite result is doubtful
```



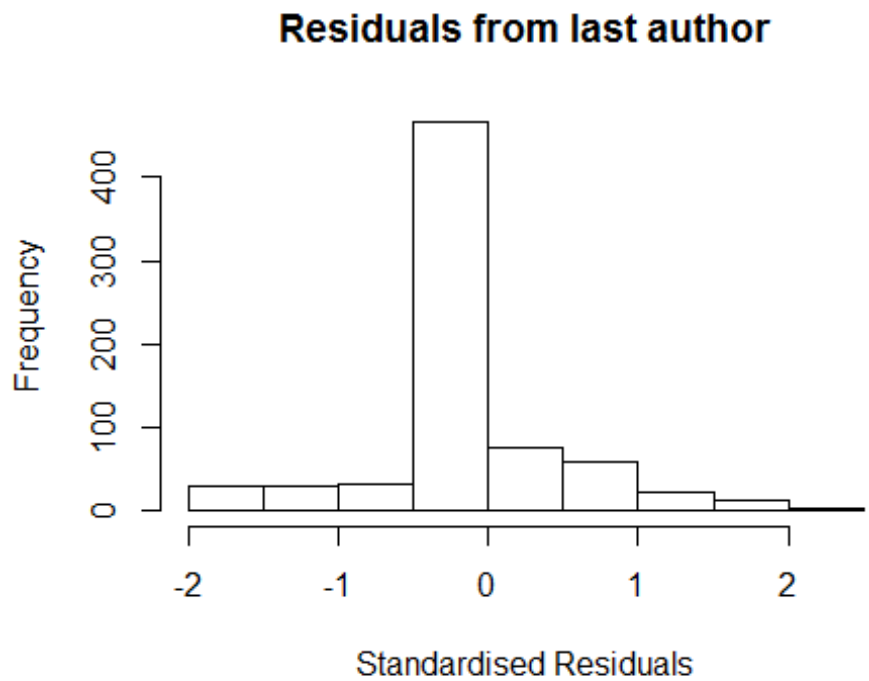

```
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale  NaN  1             NaN
## Year               NaN 16             NaN

## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "S"
## Residuals:
##    Min     1Q Median     3Q    Max
##  -1.51   0.00   0.00   0.00   2.45
##
```

```
## Exact fit detected
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)            0.000      0.000      NA       NA
## FirstAuthorFemale1     0.000      0.000      NA       NA
## Year1997               0.000      0.000      NA       NA
## Year1998               0.450      0.000      NA       NA
## Year1999               0.000      0.000      NA       NA
## Year2000               0.000      0.000      NA       NA
## Year2001               0.000      0.000      NA       NA
## Year2002               0.000      0.000      NA       NA
## Year2003               0.000      0.000      NA       NA
## Year2004               0.000      0.000      NA       NA
## Year2005               0.000      0.000      NA       NA
## Year2006               0.000      0.000      NA       NA
## Year2007               0.000      0.000      NA       NA
## Year2008               0.000      0.000      NA       NA
## Year2009               1.507      0.000      NA       NA
## Year2010               0.538      0.000      NA       NA
## Year2011               0.000      0.000      NA       NA
## Year2012               1.312      0.000      NA       NA
##
## Robustness weights:
##  327 observations
c(8,10,11,31,32,36,39,54,62,79,85,89,99,118,120,133,136,138,143,144,145,146,1
47,148,149,151,152,153,154,155,156,157,158,159,160,161,163,164,166,167,168,17
0,171,172,173,174,175,177,178,179,180,181,182,183,184,185,186,187,188,190,191
,192,193,194,195,196,200,207,209,215,216,218,251,258,260,277,292,296,300,304,
305,306,308,312,331,335,339,342,343,344,345,346,347,367,376,377,385,396,398,3
99,414,415,416,426,429,430,431,432,434,443,444,447,448,449,450,451,452,453,45
4,455,456,458,482,483,484,485,486,487,488,489,490,491,492,493,494,495,496,497
,498,499,500,502,503,504,505,506,507,508,509,510,511,512,513,514,515,516,517,
518,519,520,521,522,523,524,525,526,527,528,529,530,531,532,533,534,535,536,5
38,539,540,541,542,543,544,545,546,547,548,549,550,551,552,554,555,556,557,55
8,559,560,563,564,565,566,567,569,570,571,572,573,574,575,576,577,578,579,580
,581,582,583,584,585,586,587,588,589,590,591,592,593,594,595,596,597,598,599,
600,601,602,603,605,606,607,608,609,610,611,614,616,620,621,623,626,629,631,6
34,638,639,640,642,643,644,646,648,649,650,651,652,657,663,668,669,670,671,67
2,673,674,675,676,677,678,679,681,682,683,684,685,686,687,688,689,690,691,692
,693,694,695,696,697,698,699,700,701,702,703,704,705,706,707,708,709,710,711,
712,713,714,715,716,717,718,719,720,721,722,723,724,725,726,727)
##   are outliers with |weight| = 0 ( < 0.00014);
##  400 weights are ~= 1.
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.38e-04          1.82e-12
## warn.limit.reject warn.limit.meanrw
```

```
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"

## Warning in lmrob.S(x, y, control = control, mf = mf): S-estimated scale ==
## 0: Probably exact fit; check your data

## Warning in lmrob.S(x, y, control = control, mf = mf): diag(.) had 0 or NA
## entries; non-finite result is doubtful
```



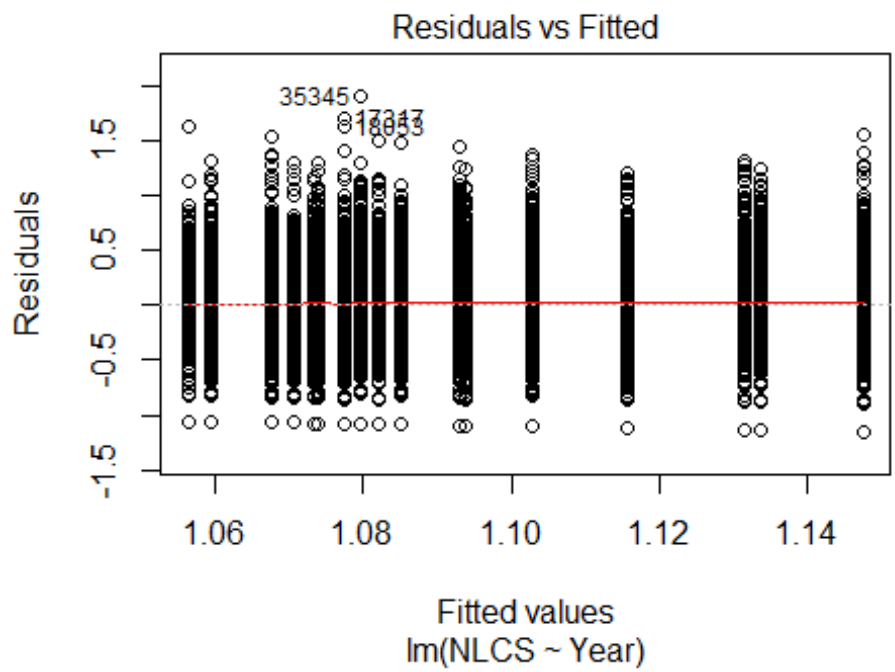

```
##                  GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale  NaN  1             NaN
## Year              NaN 16             NaN
```




```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "S"
## Residuals:
##    Min     1Q Median     3Q    Max
##  -2.05   0.00   0.00   0.00   2.35
##
## Exact fit detected
##
## Coefficients:
##                   Estimate Std. Error t value Pr(>|t|)
## (Intercept)           0.00       0.00      NA       NA
## LastAuthorFemale1     0.00       0.00      NA       NA
## Year1997              0.00       0.00      NA       NA
## Year1998              0.00       0.00      NA       NA
## Year1999              0.00       0.00      NA       NA
## Year2000              0.00       0.00      NA       NA
## Year2001              0.00       0.00      NA       NA
## Year2002              0.00       0.00      NA       NA
## Year2003              0.00       0.00      NA       NA
## Year2004              0.00       0.00      NA       NA
```

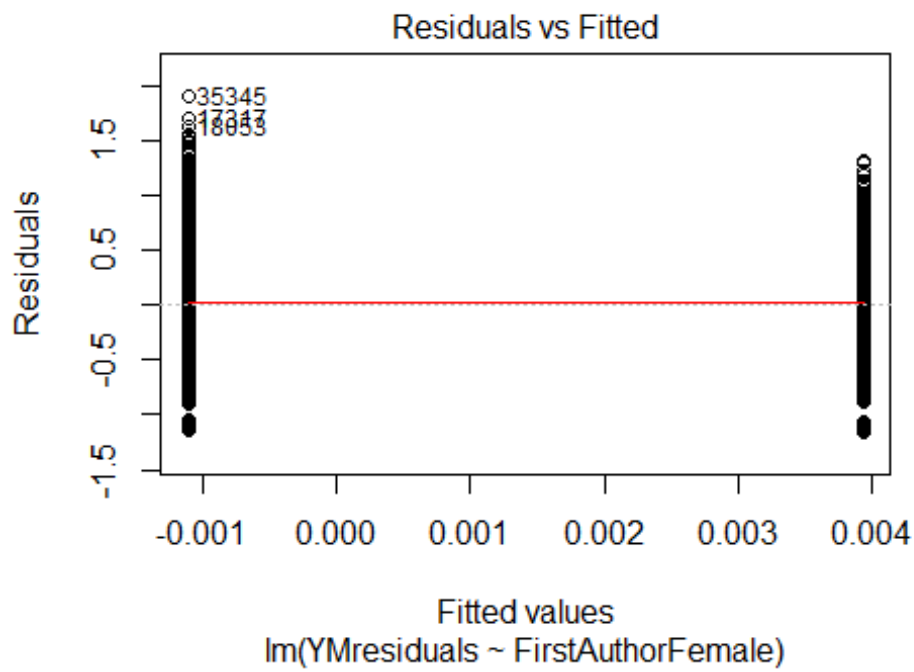
```
## Year2005              0.00       0.00      NA       NA
## Year2006              0.00       0.00      NA       NA
## Year2007              1.36       0.00      NA       NA
## Year2008              0.00       0.00      NA       NA
## Year2009              0.00       0.00      NA       NA
## Year2010              0.00       0.00      NA       NA
## Year2011              1.69       0.00      NA       NA
## Year2012              2.05       0.00      NA       NA
##
## Robustness weights:
##  265 observations
c(8,10,11,31,32,36,39,54,62,79,85,89,99,118,120,133,136,138,150,162,165,169,1
71,174,176,184,186,189,196,200,207,209,215,216,218,251,258,260,277,292,296,30
0,304,305,306,308,312,331,335,339,342,343,344,345,346,347,367,376,377,385,396
,398,399,414,415,416,426,429,430,431,432,434,442,444,445,446,447,448,449,450,
451,452,453,454,455,456,458,483,484,485,486,488,490,491,493,495,496,498,500,5
01,502,503,504,505,506,507,508,512,514,524,535,536,537,541,543,545,546,547,55
3,554,555,556,557,558,559,560,561,562,563,564,565,566,567,568,569,570,571,572
,573,574,577,583,586,601,604,605,606,607,609,610,611,612,613,614,615,616,617,
618,619,620,621,622,623,624,625,626,627,628,629,630,631,632,633,634,635,636,6
37,638,639,640,641,642,643,644,645,646,647,648,649,650,651,652,653,654,655,65
6,657,658,659,660,661,662,663,664,665,666,667,668,670,671,672,673,674,675,677
,678,679,680,681,682,683,684,685,686,687,688,689,690,691,692,693,694,695,696,
697,698,699,700,701,702,703,704,705,706,707,708,709,710,711,712,713,714,715,7
16,717,718,719,720,721,722,723,724,725,726,727)
##   are outliers with |weight| = 0 ( < 0.00014);
##  462 weights are ~= 1.
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.38e-04          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Sample size for the above analysis:  727"
## [1] ""
## [1] ""
## [1] "####################################"
## [1] "Analysis of AJSC 3002"
## [1] "####################################"
## [1] "Sample sizes for all years [All, first gendered, first & last
```

```
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 1933 1808 1884 1770 1738 1832 1959 1945 2030 2083 2236 2144 2242 2266 2380
## 2011 2012
## 2337 2260
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 1266 1204 1230 1157  872  942 1223 1209 1274 1333 1354 1380 1428 1439 1499
## 2011 2012
## 1462 1428
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 1014  956  999  919  711  781 1006 1012 1051 1069 1092 1114 1165 1161 1207
## 2011 2012
## 1180 1158
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
##  Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 240, df = 16, p-value <2e-16
```



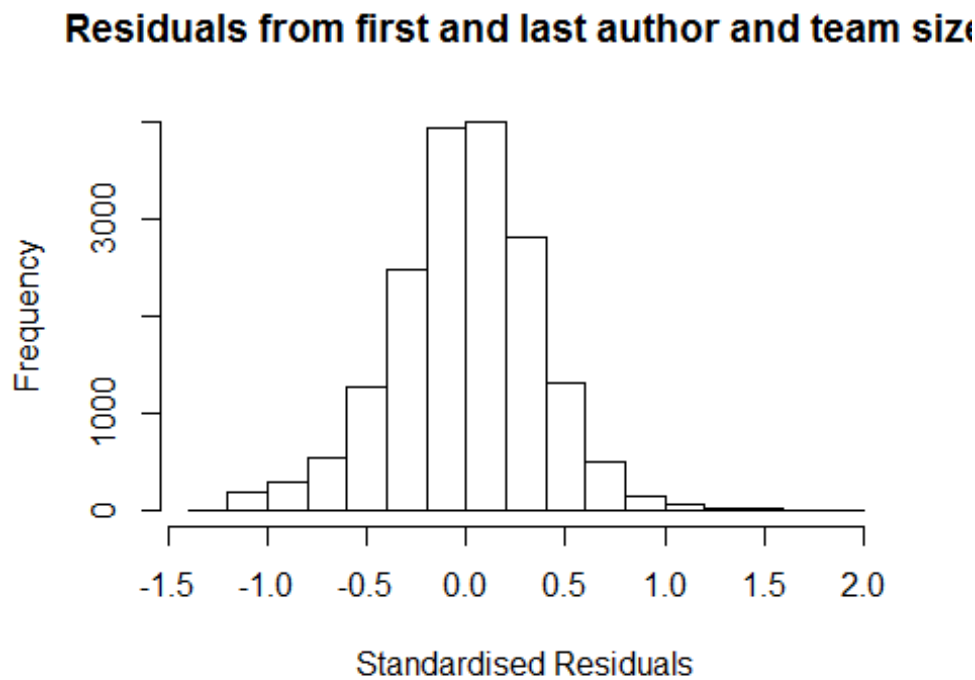

```
##
##  Bartlett test of homogeneity of variances
##
```

```
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 0.12, df = 1, p-value = 0.7
```




```
## [1] "Female first author team size 2018 geometric mean: 5.23326494894247"
## [1] "Male first author team size 2018 geometric mean: 5.91646854400563"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 73000, p-value = 0.02
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 5.00971395631651"
## [1] "Male last author team size 2018 geometric mean: 5.84860508287847"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 50000, p-value = 0.01
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.014  1           1.007
## LastAuthorFemale  1.006  1           1.003
```

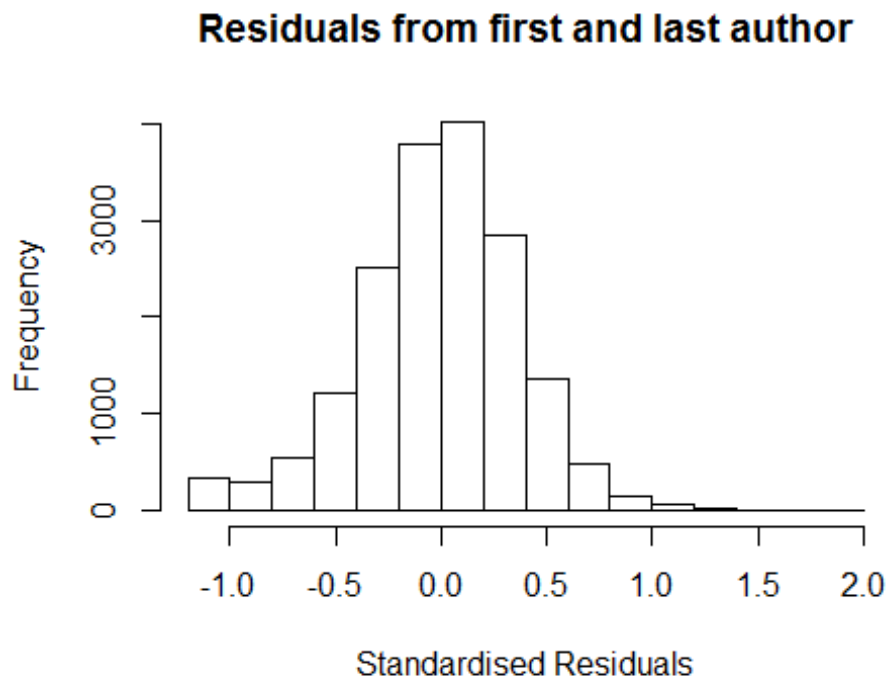
```
## UniqueAuthors     1.097  4           1.012
## Year              1.108 16           1.003
```




```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
UniqueAuthors +
##     Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##     k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.2507 -0.2244  0.0038  0.2264  1.9637
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         0.97764    0.02372   41.22  < 2e-16 ***
## FirstAuthorFemale1  0.00904    0.00654    1.38  0.16657
## LastAuthorFemale1  -0.01574    0.00758   -2.08  0.03797 *
## UniqueAuthors2      0.13915    0.02160    6.44  1.2e-10 ***
## UniqueAuthors3      0.15810    0.02169    7.29  3.2e-13 ***
## UniqueAuthors4      0.18269    0.02188    8.35  < 2e-16 ***
## UniqueAuthors5      0.26405    0.02100   12.58  < 2e-16 ***
```

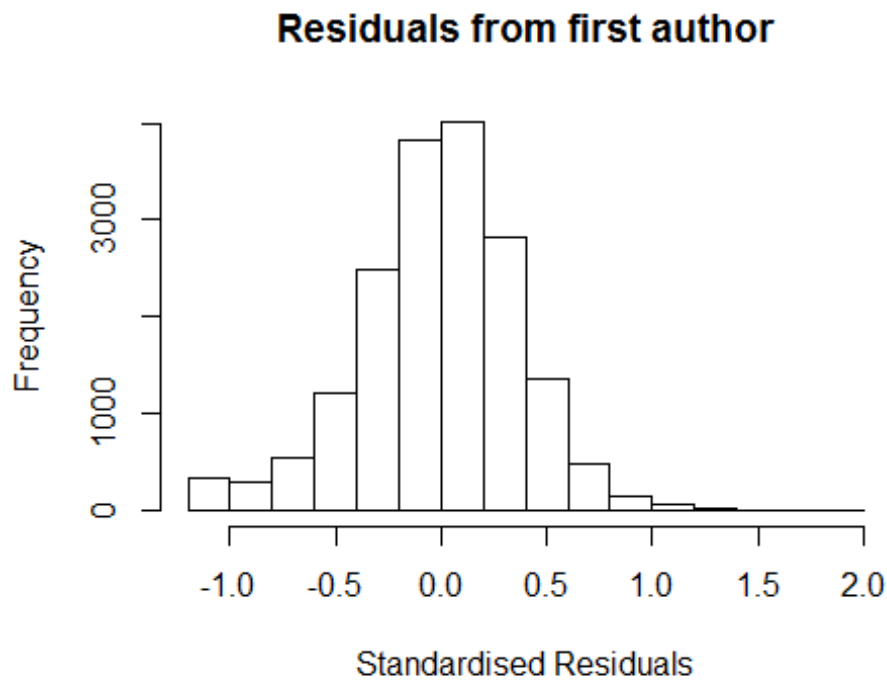
```
## Year1997           -0.00929    0.01888   -0.49  0.62279
## Year1998           -0.01182    0.01724   -0.69  0.49317
## Year1999           -0.04541    0.01742   -2.61  0.00915 **
## Year2000           -0.07382    0.01818   -4.06  4.9e-05 ***
## Year2001           -0.10210    0.01716   -5.95  2.7e-09 ***
## Year2002           -0.09807    0.01643   -5.97  2.4e-09 ***
## Year2003           -0.10479    0.01623   -6.46  1.1e-10 ***
## Year2004           -0.10112    0.01628   -6.21  5.4e-10 ***
## Year2005           -0.12098    0.01690   -7.16  8.5e-13 ***
## Year2006           -0.10748    0.01652   -6.51  7.8e-11 ***
## Year2007           -0.09168    0.01650   -5.56  2.8e-08 ***
## Year2008           -0.05689    0.01683   -3.38  0.00073 ***
## Year2009           -0.08531    0.01655   -5.16  2.6e-07 ***
## Year2010           -0.10656    0.01693   -6.29  3.2e-10 ***
## Year2011           -0.10449    0.01756   -5.95  2.7e-09 ***
## Year2012           -0.07695    0.01762   -4.37  1.3e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.337
## Multiple R-squared:  0.0457, Adjusted R-squared:  0.0445
## Convergence in 17 IRWLS iterations
##
## Robustness weights:
##  3 observations c(5003,7803,15824)
##   are outliers with |weight| = 0 ( < 5.7e-06);
##  1501 weights are ~= 1. The remaining 16091 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0001  0.8630  0.9510  0.8930  0.9860  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          5.68e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.015  1           1.007
```

```
## LastAuthorFemale  1.005  1           1.002
## Year              1.019 16           1.001
```



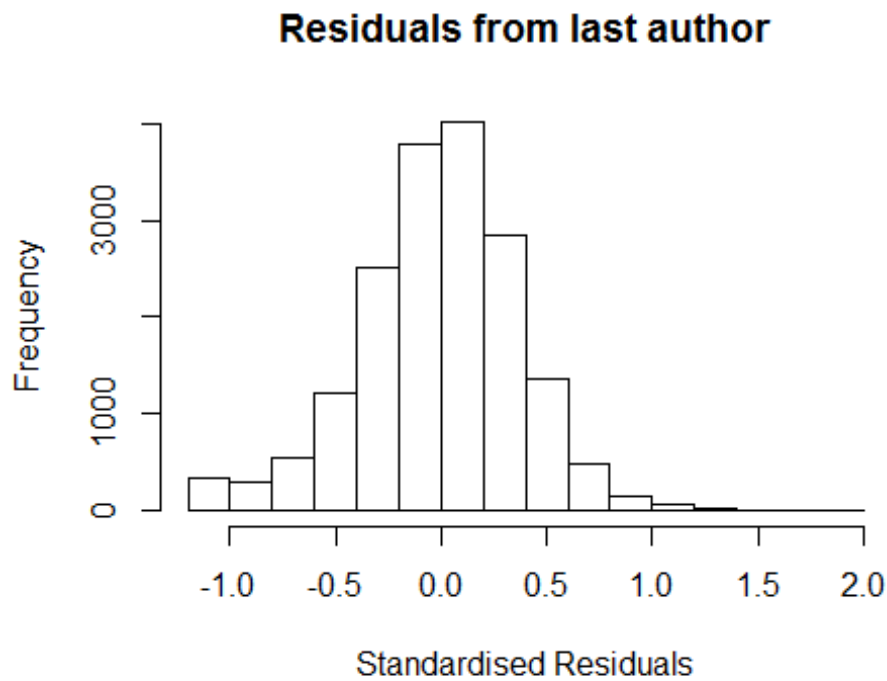

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.17196 -0.23042  0.00635  0.22941  1.89378
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.16006    0.01320   87.89  < 2e-16 ***
## FirstAuthorFemale1  0.01190    0.00661    1.80  0.07180 .
## LastAuthorFemale1  -0.01226    0.00772   -1.59  0.11231
## Year1997           -0.00811    0.01900   -0.43  0.66957
## Year1998           -0.01017    0.01753   -0.58  0.56188
## Year1999           -0.04241    0.01763   -2.41  0.01616 *
## Year2000           -0.06774    0.01842   -3.68  0.00024 ***
## Year2001           -0.09474    0.01739   -5.45  5.2e-08 ***
## Year2002           -0.08163    0.01662   -4.91  9.1e-07 ***
```

```
## Year2003           -0.07953    0.01644   -4.84  1.3e-06 ***
## Year2004           -0.07333    0.01645   -4.46  8.3e-06 ***
## Year2005           -0.09895    0.01721   -5.75  9.1e-09 ***
## Year2006           -0.08147    0.01674   -4.87  1.2e-06 ***
## Year2007           -0.05785    0.01674   -3.46  0.00055 ***
## Year2008           -0.02432    0.01697   -1.43  0.15176
## Year2009           -0.05006    0.01680   -2.98  0.00288 **
## Year2010           -0.07164    0.01713   -4.18  2.9e-05 ***
## Year2011           -0.07783    0.01772   -4.39  1.1e-05 ***
## Year2012           -0.04677    0.01783   -2.62  0.00873 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.342
## Multiple R-squared:  0.00711,    Adjusted R-squared:  0.00609
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
##  4 observations c(5003,7803,8168,15824)
##   are outliers with |weight| = 0 ( < 5.7e-06);
##  1508 weights are ~= 1. The remaining 16083 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0075  0.8650  0.9510  0.8920  0.9860  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          5.68e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.015  1           1.007
## Year              1.015 16           1.000
```




```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.16992 -0.23038  0.00687  0.22944  1.89562
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)          1.15840    0.01315   88.06  < 2e-16 ***
## FirstAuthorFemale1  0.01152    0.00662    1.74  0.08174 .
## Year1997           -0.00814    0.01901   -0.43  0.66862
## Year1998           -0.01015    0.01754   -0.58  0.56272
## Year1999           -0.04237    0.01763   -2.40  0.01627 *
## Year2000           -0.06796    0.01844   -3.69  0.00023 ***
## Year2001           -0.09490    0.01740   -5.45  5.0e-08 ***
## Year2002           -0.08165    0.01662   -4.91  9.1e-07 ***
## Year2003           -0.07971    0.01645   -4.85  1.3e-06 ***
## Year2004           -0.07346    0.01645   -4.47  8.0e-06 ***
## Year2005           -0.09878    0.01721   -5.74  9.6e-09 ***
## Year2006           -0.08141    0.01675   -4.86  1.2e-06 ***
```

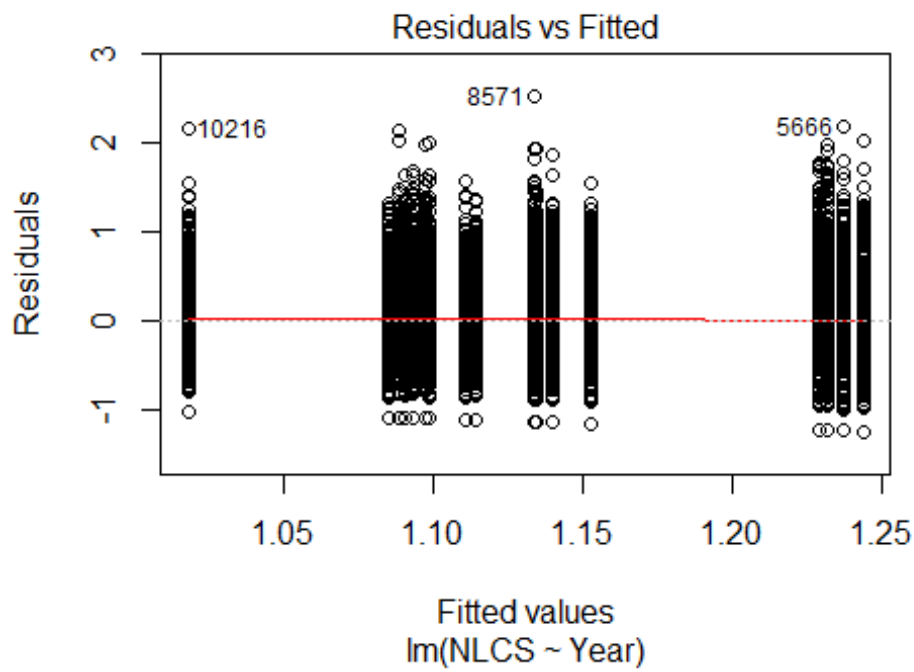
```
## Year2007           -0.05789    0.01674   -3.46  0.00055 ***
## Year2008           -0.02472    0.01698   -1.46  0.14541
## Year2009           -0.04999    0.01681   -2.97  0.00295 **
## Year2010           -0.07186    0.01715   -4.19  2.8e-05 ***
## Year2011           -0.07802    0.01773   -4.40  1.1e-05 ***
## Year2012           -0.04695    0.01784   -2.63  0.00849 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.342
## Multiple R-squared:  0.00695,    Adjusted R-squared:  0.00599
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
##  4 observations c(5003,7803,8168,15824)
##   are outliers with |weight| = 0 ( < 5.7e-06);
##  1490 weights are ~= 1. The remaining 16101 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0072  0.8650  0.9500  0.8920  0.9860  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          5.68e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##                   GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.004  1           1.002
## Year             1.004 16           1.000
```



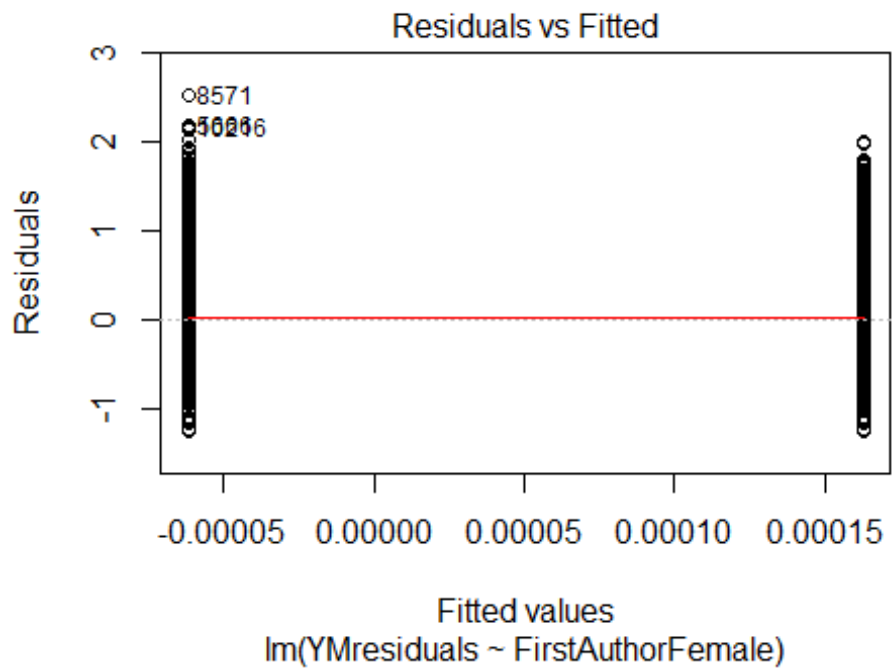

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.16170 -0.22988  0.00533  0.22981  1.89081
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.16170    0.01315   88.31  < 2e-16 ***
## LastAuthorFemale1 -0.01175    0.00773   -1.52  0.12847
## Year1997          -0.00787    0.01900   -0.41  0.67858
## Year1998          -0.01003    0.01753   -0.57  0.56712
## Year1999          -0.04212    0.01762   -2.39  0.01685 *
## Year2000          -0.06733    0.01842   -3.66  0.00026 ***
## Year2001          -0.09414    0.01739   -5.41  6.3e-08 ***
## Year2002          -0.08068    0.01662   -4.85  1.2e-06 ***
## Year2003          -0.07823    0.01642   -4.76  1.9e-06 ***
## Year2004          -0.07254    0.01643   -4.41  1.0e-05 ***
## Year2005          -0.09813    0.01721   -5.70  1.2e-08 ***
## Year2006          -0.08000    0.01673   -4.78  1.7e-06 ***
```

```
## Year2007          -0.05651    0.01672   -3.38  0.00073 ***
## Year2008          -0.02282    0.01695   -1.35  0.17802
## Year2009          -0.04891    0.01680   -2.91  0.00360 **
## Year2010          -0.07007    0.01711   -4.10  4.2e-05 ***
## Year2011          -0.07651    0.01772   -4.32  1.6e-05 ***
## Year2012          -0.04513    0.01781   -2.53  0.01129 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.342
## Multiple R-squared:  0.00692,    Adjusted R-squared:  0.00596
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
##  4 observations c(5003,7803,8168,15824)
##   are outliers with |weight| = 0 ( < 5.7e-06);
##  1507 weights are ~= 1. The remaining 16084 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0078  0.8650  0.9500  0.8920  0.9860  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          5.68e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                  psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Sample size for the above analysis:  17595"
## [1] ""
## [1] ""
## [1] "##################################"
## [1] "Analysis of AJSC 3003"
## [1] "##################################"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 1426 1294 1244 1353 1288 1335 1563 1453 1640 1620 1706 1817 2042 2014 2085
## 2011 2012
## 2189 2174
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
```

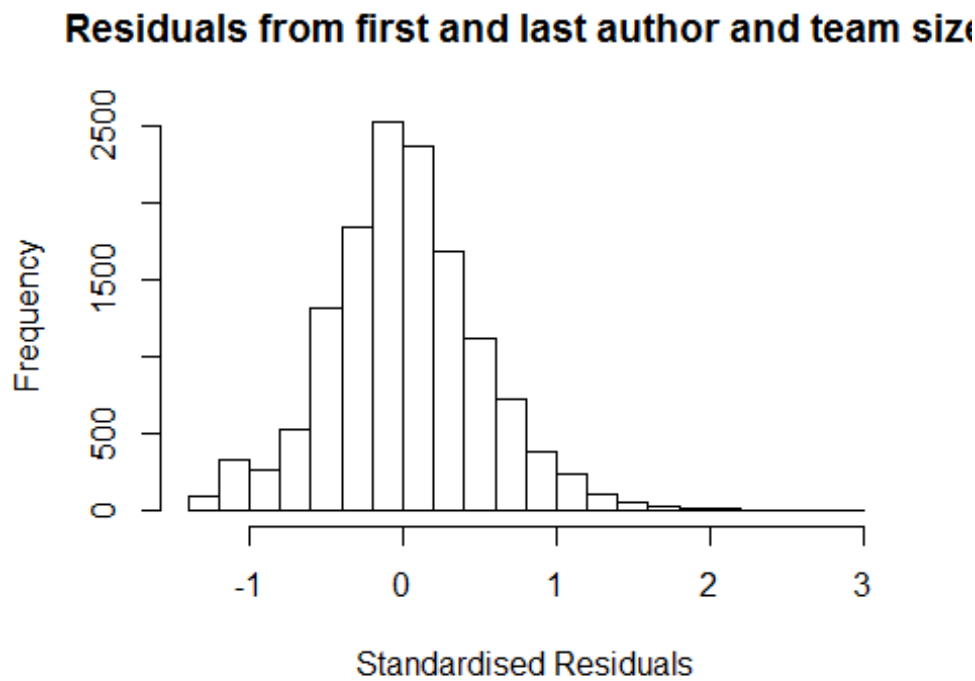
```
##  761  672  641  780  690  573  982  896  990  998 1041 1132 1295 1198 1281
## 2011 2012
## 1373 1377
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  657  552  537  639  582  491  839  745  814  785  850  904 1052  954 1028
## 2011 2012
## 1101 1108
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
##  Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 300, df = 16, p-value <2e-16
```




```
##
##  Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 69, df = 1, p-value <2e-16
```



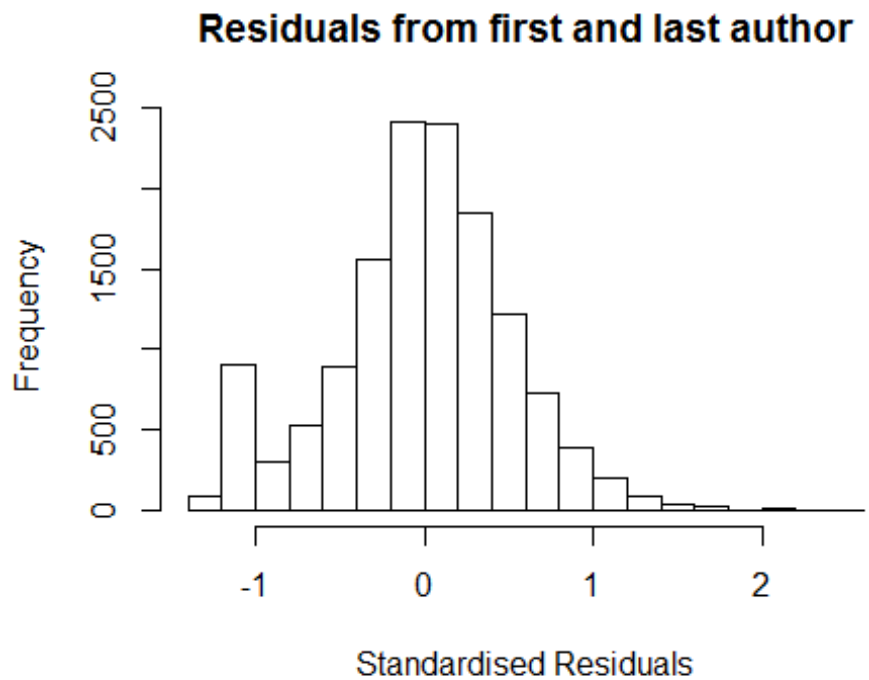

```
## [1] "Female first author team size 2018 geometric mean: 4.4203318949986"
## [1] "Male first author team size 2018 geometric mean: 4.51266225780054"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 110000, p-value = 0.4
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 4.41593786361254"
## [1] "Male last author team size 2018 geometric mean: 4.49578798077966"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 86000, p-value = 0.8
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.041  1           1.020
## LastAuthorFemale  1.032  1           1.016
## UniqueAuthors     1.102  4           1.012
## Year              1.094 16           1.003
```




```
## [1] "List of  2 outliers with residuals above  2.5"
##         ScopusId  NLCS Year OneField Fields residuals
## 5666  0032885450 3.413 1999     3003      1     2.800
## 10216 0037122675 3.175 2002     3003      1     2.773
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
UniqueAuthors +
##     Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##     k.max = 1000))
##  \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.35719 -0.29843 -0.00839  0.30510  2.80040
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         0.60097    0.03031   19.83  < 2e-16 ***
## FirstAuthorFemale1  0.00858    0.00971    0.88  0.37692
## LastAuthorFemale1  -0.06272    0.01103   -5.69  1.3e-08 ***
## UniqueAuthors2      0.60346    0.02422   24.91  < 2e-16 ***
## UniqueAuthors3      0.69277    0.02239   30.94  < 2e-16 ***
## UniqueAuthors4      0.72782    0.02194   33.17  < 2e-16 ***
## UniqueAuthors5      0.73601    0.01940   37.94  < 2e-16 ***
## Year1997           -0.05159    0.03600   -1.43  0.15191
## Year1998           -0.00136    0.03457   -0.04  0.96857
```

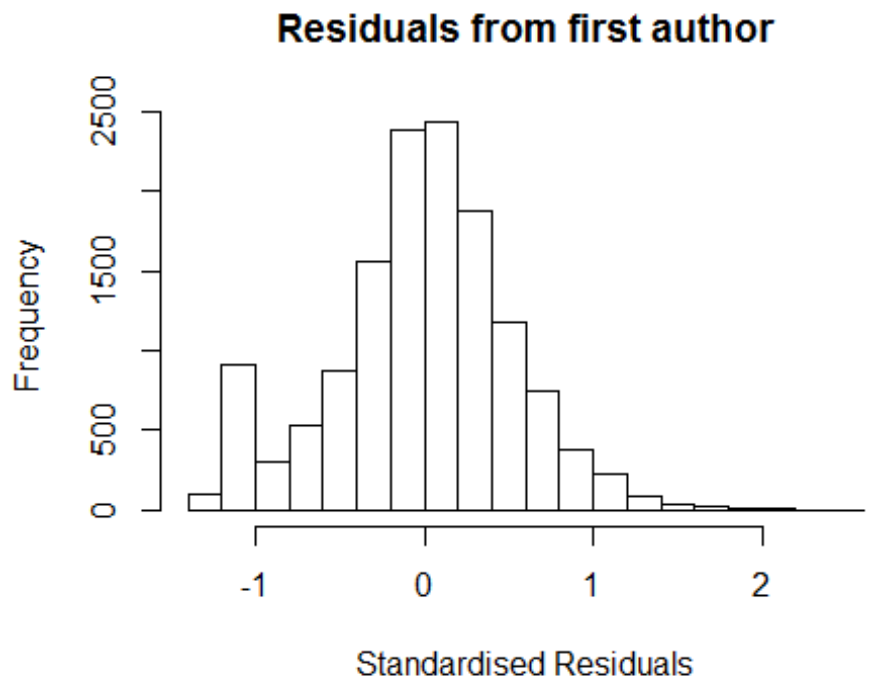
```
## Year1999            0.01162    0.03347    0.35  0.72839
## Year2000           -0.08262    0.03327   -2.48  0.01303 *
## Year2001           -0.12963    0.03270   -3.96  7.4e-05 ***
## Year2002           -0.19884    0.02897   -6.86  7.0e-12 ***
## Year2003           -0.15562    0.02863   -5.44  5.5e-08 ***
## Year2004           -0.13940    0.02807   -4.97  6.9e-07 ***
## Year2005           -0.10794    0.02857   -3.78  0.00016 ***
## Year2006           -0.15242    0.02799   -5.44  5.3e-08 ***
## Year2007           -0.09568    0.02824   -3.39  0.00070 ***
## Year2008           -0.15783    0.02741   -5.76  8.6e-09 ***
## Year2009           -0.17212    0.02742   -6.28  3.6e-10 ***
## Year2010           -0.15631    0.02720   -5.75  9.3e-09 ***
## Year2011           -0.14829    0.02784   -5.33  1.0e-07 ***
## Year2012           -0.14545    0.02820   -5.16  2.5e-07 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.424
## Multiple R-squared:  0.195,  Adjusted R-squared:  0.194
## Convergence in 15 IRWLS iterations
##
## Robustness weights:
##  15 observations
c(1059,1613,1618,1996,2271,2441,3349,3377,4012,4073,5666,6012,9055,10363,1255
0)
##   are outliers with |weight| = 0 ( < 7.3e-06);
##  1063 weights are ~= 1. The remaining 12560 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0001  0.8530  0.9440  0.8810  0.9860  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          7.33e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.022  1           1.011
```

```
## LastAuthorFemale  1.016  1           1.008
## Year              1.019 16           1.001
```



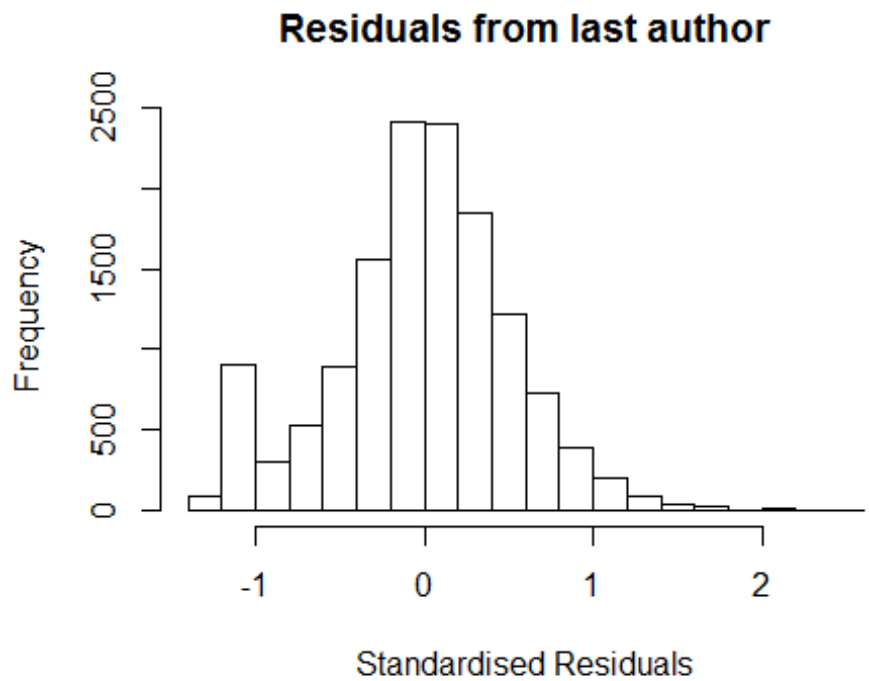

```
## [1] "List of  1 outliers with residuals above  2.5"
##        ScopusId  NLCS Year OneField Fields residuals
## 8571 0035289779 3.659 2001     3003      1     2.533
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.2881 -0.2968  0.0106  0.3139  2.5332
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)           1.2451     0.0255   48.83  < 2e-16 ***
## FirstAuthorFemale1    0.0120     0.0106    1.13  0.25985
## LastAuthorFemale1    -0.0888     0.0128   -6.94  4.2e-12 ***
## Year1997             -0.0471     0.0384   -1.23  0.21999
## Year1998              0.0311     0.0353    0.88  0.37954
## Year1999              0.0169     0.0343    0.49  0.62160
## Year2000             -0.0692     0.0349   -1.98  0.04747 *
## Year2001             -0.1193     0.0355   -3.36  0.00078 ***
## Year2002             -0.2302     0.0319   -7.21  6.0e-13 ***
```

```
## Year2003            -0.1433     0.0308   -4.66  3.2e-06 ***
## Year2004            -0.1154     0.0300   -3.84  0.00012 ***
## Year2005            -0.0962     0.0302   -3.19  0.00144 **
## Year2006            -0.1175     0.0297   -3.96  7.7e-05 ***
## Year2007            -0.0662     0.0302   -2.19  0.02827 *
## Year2008            -0.1333     0.0299   -4.46  8.4e-06 ***
## Year2009            -0.1374     0.0293   -4.69  2.7e-06 ***
## Year2010            -0.1167     0.0291   -4.01  6.0e-05 ***
## Year2011            -0.1419     0.0306   -4.63  3.6e-06 ***
## Year2012            -0.1359     0.0308   -4.42  1.0e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.461
## Multiple R-squared:  0.0189, Adjusted R-squared:  0.0176
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
##  2 observations c(3349,4073) are outliers with |weight| = 0 ( < 7.3e-06);
##  1279 weights are ~= 1. The remaining 12357 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0001  0.8370  0.9500  0.8800  0.9860  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          7.33e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##                   GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.01  1           1.005
## Year              1.01 16           1.000
```




```
## [1] "List of  1 outliers with residuals above  2.5"
##        ScopusId  NLCS Year OneField Fields residuals
## 8571 0035289779 3.659 2001     3003      1     2.533
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.2684 -0.2973  0.0124  0.3119  2.5457
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.22996    0.02555   48.14  < 2e-16 ***
## FirstAuthorFemale1  0.00507    0.01088    0.47  0.64088
## Year1997           -0.04436    0.03863   -1.15  0.25089
## Year1998            0.03339    0.03543    0.94  0.34596
## Year1999            0.02028    0.03456    0.59  0.55725
## Year2000           -0.06718    0.03512   -1.91  0.05581 .
## Year2001           -0.11665    0.03584   -3.25  0.00114 **
## Year2002           -0.22874    0.03214   -7.12  1.2e-12 ***
## Year2003           -0.14017    0.03088   -4.54  5.7e-06 ***
## Year2004           -0.11456    0.03017   -3.80  0.00015 ***
## Year2005           -0.09541    0.03032   -3.15  0.00165 **
## Year2006           -0.11714    0.02983   -3.93  8.7e-05 ***
```

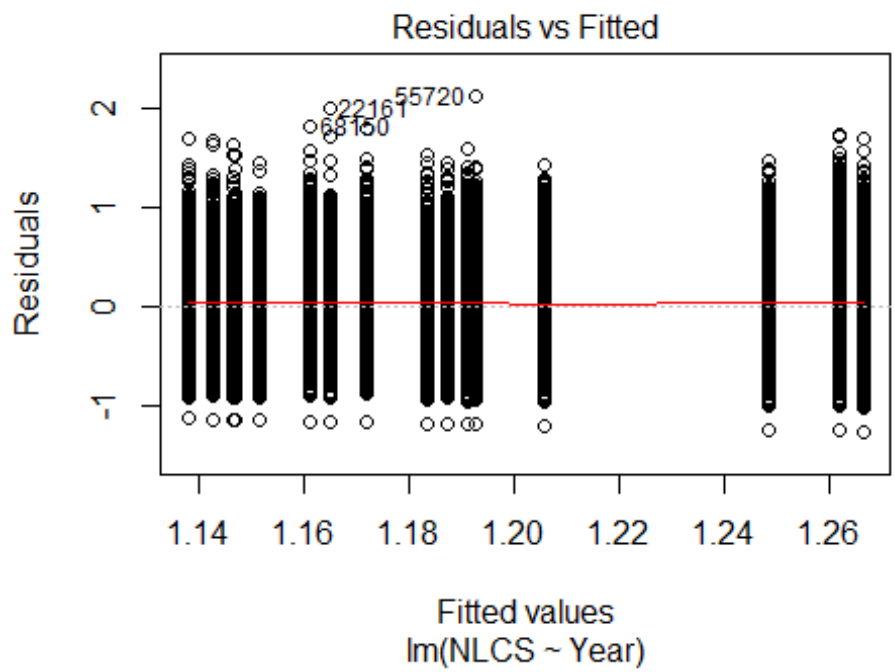
```
## Year2007           -0.06608    0.03033   -2.18  0.02937 *
## Year2008           -0.13322    0.03011   -4.43  9.7e-06 ***
## Year2009           -0.13685    0.02951   -4.64  3.6e-06 ***
## Year2010           -0.11499    0.02927   -3.93  8.6e-05 ***
## Year2011           -0.14186    0.03089   -4.59  4.4e-06 ***
## Year2012           -0.13610    0.03099   -4.39  1.1e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.461
## Multiple R-squared:  0.0141, Adjusted R-squared:  0.0128
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
##  3 observations c(2271,3349,4073)
##   are outliers with |weight| = 0 ( < 7.3e-06);
##  1257 weights are ~= 1. The remaining 12378 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0004  0.8360  0.9500  0.8800  0.9870  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          7.33e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##                   GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.006  1           1.003
## Year             1.006 16           1.000
```



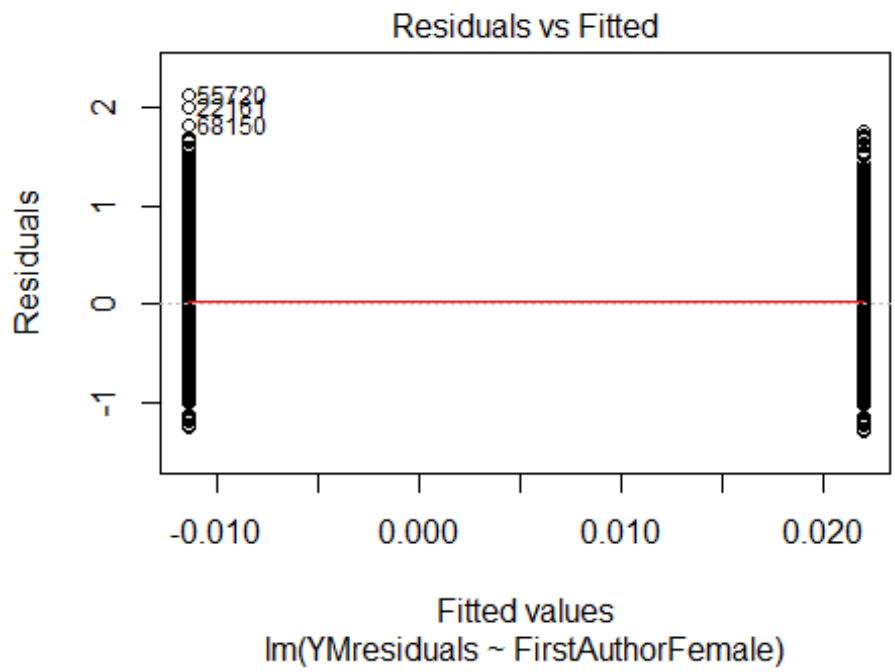

```
## [1] "List of  1 outliers with residuals above  2.5"
##        ScopusId  NLCS Year OneField Fields residuals
## 8571 0035289779 3.659 2001     3003      1     2.533
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.2788 -0.2958  0.0115  0.3123  2.5303
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.2475     0.0254   49.12  < 2e-16 ***
## LastAuthorFemale1  -0.0875     0.0129   -6.77  1.4e-11 ***
## Year1997           -0.0471     0.0384   -1.23  0.21962
## Year1998            0.0313     0.0354    0.88  0.37673
## Year1999            0.0167     0.0343    0.49  0.62592
## Year2000           -0.0689     0.0349   -1.97  0.04856 *
## Year2001           -0.1188     0.0355   -3.34  0.00083 ***
## Year2002           -0.2297     0.0320   -7.19  6.8e-13 ***
## Year2003           -0.1425     0.0308   -4.63  3.6e-06 ***
## Year2004           -0.1148     0.0300   -3.82  0.00013 ***
## Year2005           -0.0956     0.0302   -3.16  0.00156 **
## Year2006           -0.1170     0.0297   -3.94  8.2e-05 ***
```

```
## Year2007           -0.0660     0.0302   -2.19  0.02871 *
## Year2008           -0.1324     0.0299   -4.42  9.7e-06 ***
## Year2009           -0.1370     0.0293   -4.67  3.0e-06 ***
## Year2010           -0.1157     0.0291   -3.98  6.9e-05 ***
## Year2011           -0.1409     0.0306   -4.60  4.3e-06 ***
## Year2012           -0.1350     0.0308   -4.39  1.2e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.461
## Multiple R-squared:  0.0188, Adjusted R-squared:  0.0176
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
##  2 observations c(3349,4073)
##   are outliers with |weight| <= 3.4e-06 ( < 7.3e-06);
##  1297 weights are ~= 1. The remaining 12339 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0001  0.8360  0.9500  0.8800  0.9860  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          7.33e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                  psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Sample size for the above analysis:  13638"
## [1] ""
## [1] ""
## [1] "##################################"
## [1] "Analysis of AJSC 3004"
## [1] "##################################"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 4048 3826 3661 3850 3666 3953 3891 3384 3532 3581 3681 3992 4104 4217 3952
## 2011 2012
## 4582 4402
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
```

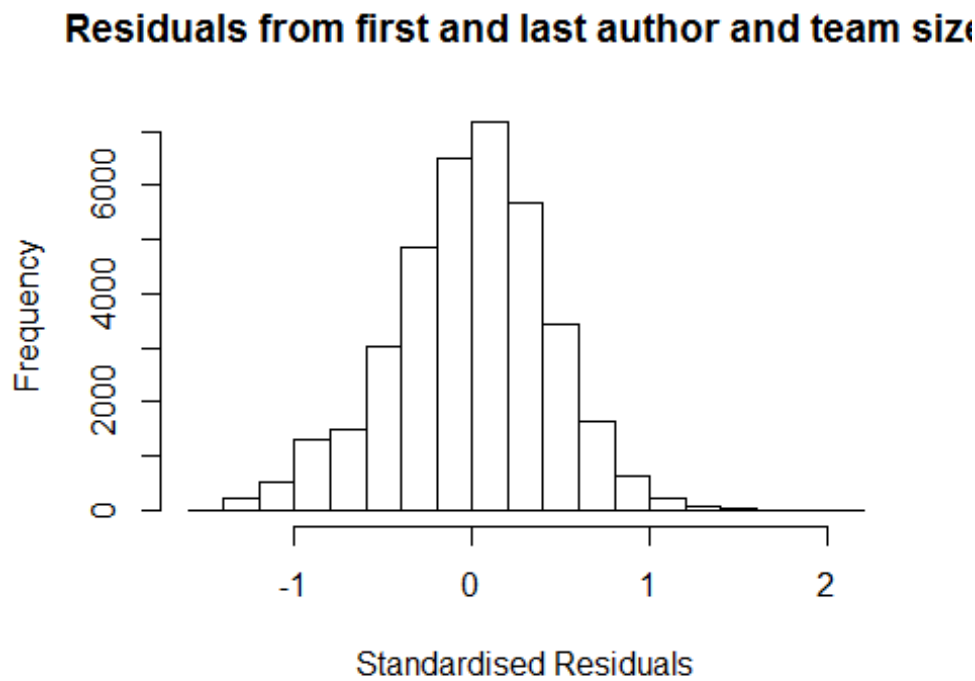
```
## 2538 2327 2315 2414 2067 1910 2595 2306 2434 2449 2456 2659 2766 2869 2664
## 2011 2012
## 3152 2988
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 2217 2032 2015 2090 1805 1681 2238 1963 2091 2080 2126 2275 2365 2431 2266
## 2011 2012
## 2681 2546
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
##  Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 250, df = 16, p-value <2e-16
```




```
##
##  Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 16, df = 1, p-value = 7e-05
```

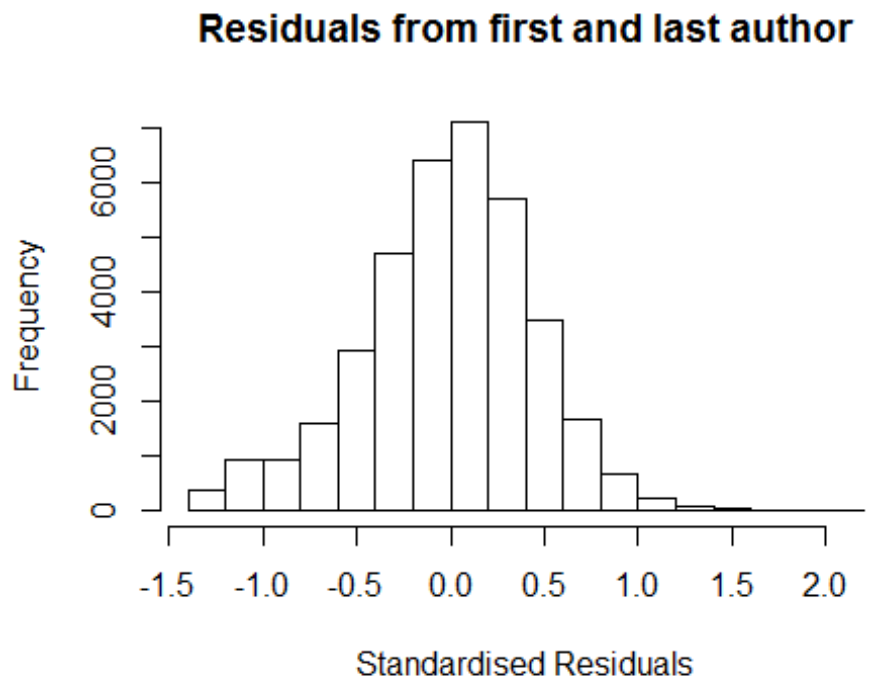



```
## [1] "Female first author team size 2018 geometric mean: 4.8070400487434"
## [1] "Male first author team size 2018 geometric mean: 4.63817087184691"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 680000, p-value = 0.5
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 4.48443161362917"
## [1] "Male last author team size 2018 geometric mean: 4.81281210211466"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 530000, p-value = 0.01
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##                     GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.015  1           1.007
## LastAuthorFemale  1.010  1           1.005
## UniqueAuthors     1.047  4           1.006
## Year              1.055 16           1.002
```



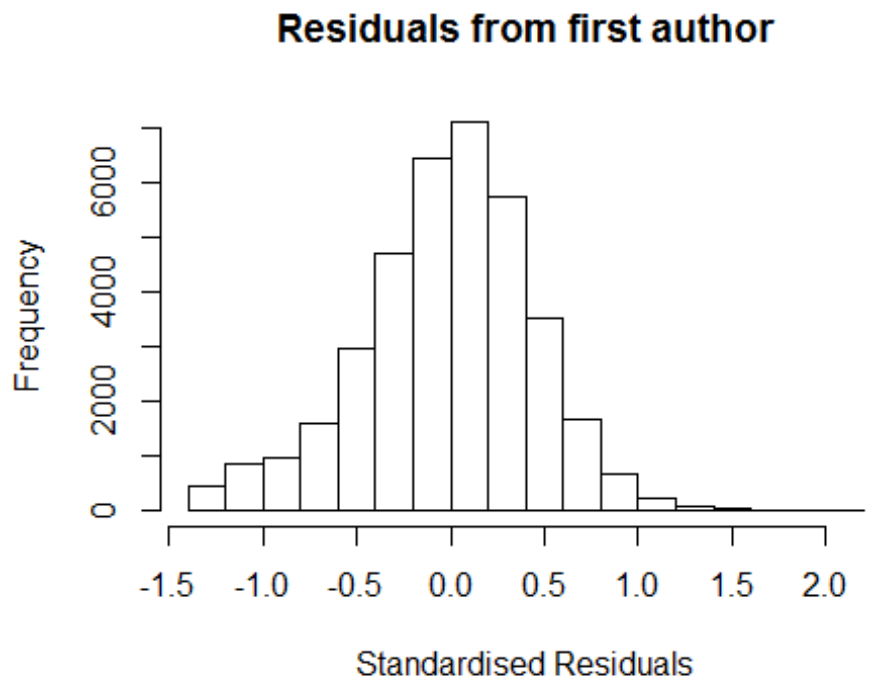

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
UniqueAuthors +
##     Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##     k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.4019 -0.2803  0.0126  0.2819  2.1459
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)         0.946574   0.016079   58.87  < 2e-16 ***
## FirstAuthorFemale1  0.025568   0.004831    5.29  1.2e-07 ***
## LastAuthorFemale1  -0.000275   0.005574   -0.05  0.96059
## UniqueAuthors2      0.261631   0.013608   19.23  < 2e-16 ***
## UniqueAuthors3      0.318153   0.013315   23.89  < 2e-16 ***
## UniqueAuthors4      0.351906   0.013316   26.43  < 2e-16 ***
## UniqueAuthors5      0.429754   0.012618   34.06  < 2e-16 ***
## Year1997            0.011041   0.015767    0.70  0.48377
## Year1998           -0.050747   0.015084   -3.36  0.00077 ***
## Year1999           -0.020800   0.015113   -1.38  0.16876
```

```
## Year2000           -0.073359   0.015176   -4.83  1.3e-06 ***
## Year2001           -0.093418   0.015309   -6.10  1.1e-09 ***
## Year2002           -0.130371   0.014501   -8.99  < 2e-16 ***
## Year2003           -0.139711   0.014540   -9.61  < 2e-16 ***
## Year2004           -0.134174   0.014160   -9.48  < 2e-16 ***
## Year2005           -0.136990   0.013972   -9.80  < 2e-16 ***
## Year2006           -0.156661   0.014192  -11.04  < 2e-16 ***
## Year2007           -0.136523   0.014057   -9.71  < 2e-16 ***
## Year2008           -0.105369   0.014049   -7.50  6.5e-14 ***
## Year2009           -0.105811   0.014094   -7.51  6.2e-14 ***
## Year2010           -0.109140   0.014119   -7.73  1.1e-14 ***
## Year2011           -0.124734   0.014306   -8.72  < 2e-16 ***
## Year2012           -0.124632   0.014381   -8.67  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.416
## Multiple R-squared:  0.079,  Adjusted R-squared:  0.0785
## Convergence in 15 IRWLS iterations
##
## Robustness weights:
##  observation 26818 is an outlier with |weight| = 0 ( < 2.7e-06);
##  3189 weights are ~= 1. The remaining 33712 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0018  0.8630  0.9500  0.8960  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          2.71e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.011  1           1.006
## LastAuthorFemale  1.006  1           1.003
## Year              1.012 16           1.000
```



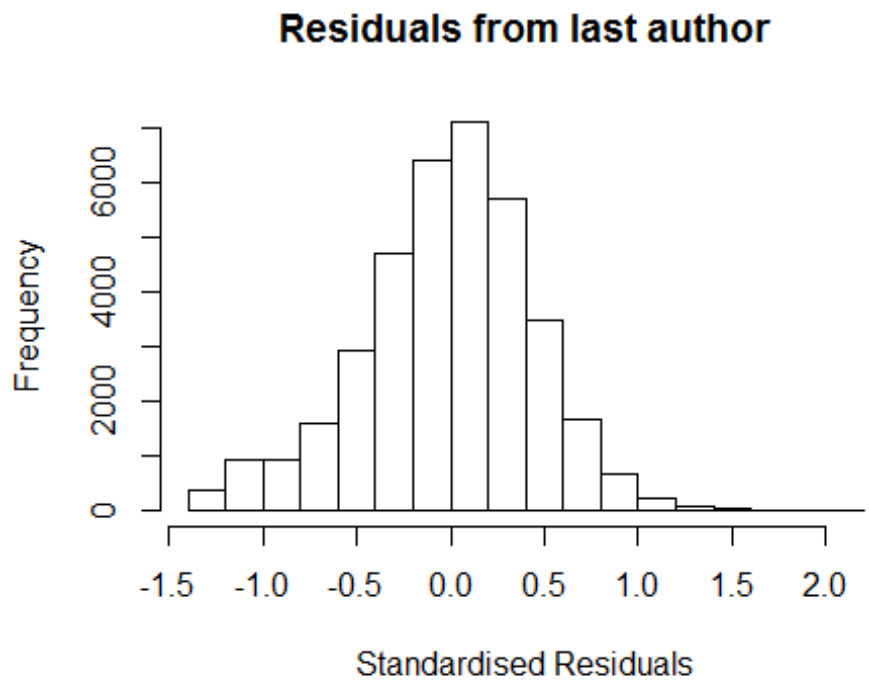

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.3159 -0.2861  0.0149  0.2824  2.1287
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)          1.27259    0.01123  113.37  < 2e-16 ***
## FirstAuthorFemale1   0.03448    0.00492    7.01  2.4e-12 ***
## LastAuthorFemale1   -0.01504    0.00570   -2.64  0.00832 **
## Year1997             0.00883    0.01596    0.55  0.57988
## Year1998            -0.05791    0.01530   -3.79  0.00015 ***
## Year1999            -0.02203    0.01526   -1.44  0.14890
## Year2000            -0.07385    0.01535   -4.81  1.5e-06 ***
## Year2001            -0.09507    0.01554   -6.12  9.6e-10 ***
## Year2002            -0.11929    0.01467   -8.13  4.4e-16 ***
## Year2003            -0.11845    0.01481   -8.00  1.3e-15 ***
## Year2004            -0.11178    0.01449   -7.71  1.2e-14 ***
## Year2005            -0.11167    0.01426   -7.83  5.0e-15 ***
```

```
## Year2006           -0.13248    0.01450   -9.14  < 2e-16 ***
## Year2007           -0.10828    0.01422   -7.61  2.7e-14 ***
## Year2008           -0.08120    0.01422   -5.71  1.1e-08 ***
## Year2009           -0.07482    0.01428   -5.24  1.6e-07 ***
## Year2010           -0.07916    0.01440   -5.50  3.9e-08 ***
## Year2011           -0.10194    0.01456   -7.00  2.6e-12 ***
## Year2012           -0.09285    0.01453   -6.39  1.7e-10 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.422
## Multiple R-squared:  0.00927,    Adjusted R-squared:  0.00878
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  observation 26818 is an outlier with |weight| = 0 ( < 2.7e-06);
##  3216 weights are ~= 1. The remaining 33685 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0274  0.8640  0.9500  0.8930  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          2.71e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.009  1           1.005
## Year              1.009 16           1.000
```




```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.3120 -0.2856  0.0147  0.2823  2.1167
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)          1.27021    0.01118  113.58  < 2e-16 ***
## FirstAuthorFemale1   0.03304    0.00492    6.71  2.0e-11 ***
## Year1997             0.00873    0.01596    0.55  0.58458
## Year1998            -0.05815    0.01530   -3.80  0.00014 ***
## Year1999            -0.02208    0.01527   -1.45  0.14812
## Year2000            -0.07415    0.01535   -4.83  1.4e-06 ***
## Year2001            -0.09550    0.01554   -6.14  8.1e-10 ***
## Year2002            -0.11956    0.01467   -8.15  3.8e-16 ***
## Year2003            -0.11912    0.01481   -8.05  8.8e-16 ***
## Year2004            -0.11238    0.01449   -7.76  8.9e-15 ***
## Year2005            -0.11180    0.01427   -7.84  4.8e-15 ***
## Year2006            -0.13296    0.01450   -9.17  < 2e-16 ***
```

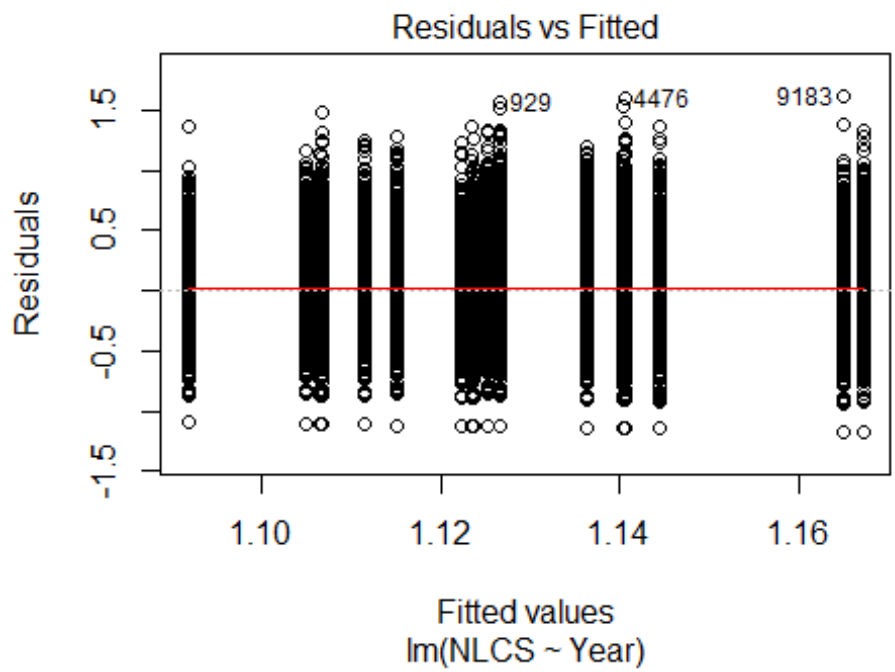
```
## Year2007           -0.10886    0.01422   -7.66  2.0e-14 ***
## Year2008           -0.08189    0.01422   -5.76  8.6e-09 ***
## Year2009           -0.07560    0.01429   -5.29  1.2e-07 ***
## Year2010           -0.08001    0.01439   -5.56  2.7e-08 ***
## Year2011           -0.10295    0.01456   -7.07  1.6e-12 ***
## Year2012           -0.09383    0.01453   -6.46  1.1e-10 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.422
## Multiple R-squared:  0.00909,    Adjusted R-squared:  0.00863
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  observation 26818 is an outlier with |weight| = 0 ( < 2.7e-06);
##  3234 weights are ~= 1. The remaining 33667 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0263  0.8640  0.9500  0.8930  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          2.71e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##                   GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.003  1           1.002
## Year             1.003 16           1.000
```



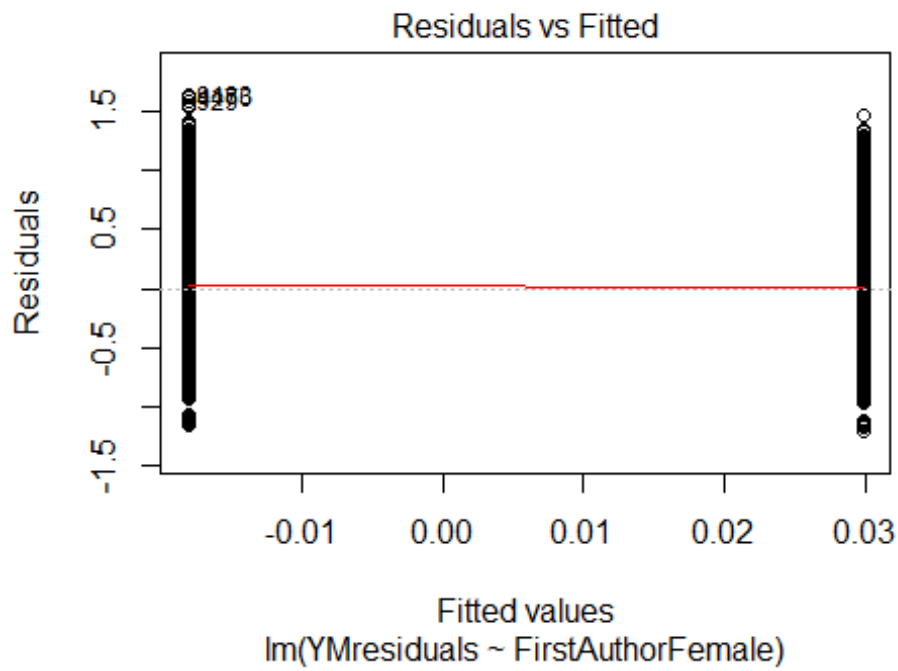

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.2913 -0.2861  0.0139  0.2837  2.1123
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)          1.2809     0.0112  114.74  < 2e-16 ***
## LastAuthorFemale1   -0.0107     0.0057   -1.89  0.05921 .
## Year1997             0.0103     0.0160    0.65  0.51728
## Year1998            -0.0567     0.0153   -3.71  0.00021 ***
## Year1999            -0.0206     0.0153   -1.35  0.17687
## Year2000            -0.0722     0.0153   -4.71  2.5e-06 ***
## Year2001            -0.0929     0.0155   -5.98  2.2e-09 ***
## Year2002            -0.1170     0.0147   -7.98  1.5e-15 ***
## Year2003            -0.1161     0.0148   -7.83  4.9e-15 ***
## Year2004            -0.1098     0.0145   -7.58  3.7e-14 ***
## Year2005            -0.1092     0.0143   -7.65  2.1e-14 ***
## Year2006            -0.1286     0.0145   -8.88  < 2e-16 ***
```

```
## Year2007           -0.1052     0.0142   -7.39  1.4e-13 ***
## Year2008           -0.0774     0.0142   -5.44  5.3e-08 ***
## Year2009           -0.0708     0.0143   -4.96  7.2e-07 ***
## Year2010           -0.0752     0.0144   -5.23  1.7e-07 ***
## Year2011           -0.0980     0.0146   -6.73  1.7e-11 ***
## Year2012           -0.0879     0.0145   -6.05  1.5e-09 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.422
## Multiple R-squared:  0.00793,    Adjusted R-squared:  0.00747
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  observation 26818 is an outlier with |weight| = 0 ( < 2.7e-06);
##  3139 weights are ~= 1. The remaining 33762 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0315  0.8640  0.9500  0.8930  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          2.71e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                    psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Sample size for the above analysis:  36902"
## [1] ""
## [1] ""
## [1] "####################################"
## [1] "Analysis of AJSC 3005"
## [1] "####################################"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 2332 2030 1839 1966 1700 1994 1921 1665 1783 1773 1957 2288 2316 2134 2033
## 2011 2012
## 2105 2076
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 1406 1210 1050 1229  879 1017 1269 1122 1189 1172 1357 1665 1619 1569 1513
```

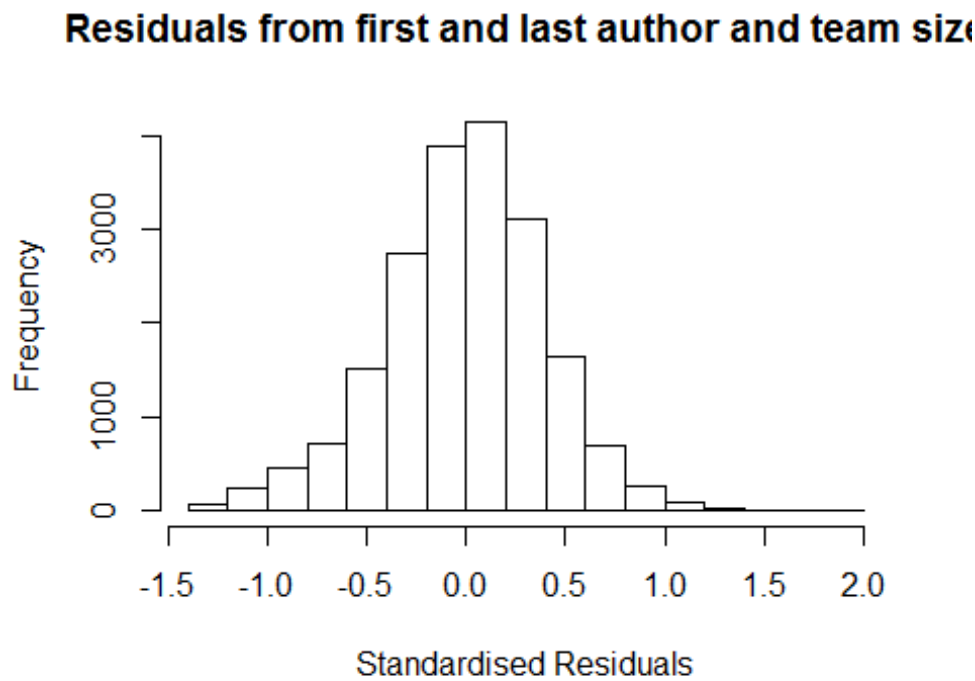
```
## 2011 2012
## 1563 1552
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 1251 1049  931 1061  761  896 1097  973 1017 1001 1199 1449 1428 1367 1310
## 2011 2012
## 1366 1368
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
##  Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 480, df = 16, p-value <2e-16
```




```
##
##  Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 52, df = 1, p-value = 4e-13
```

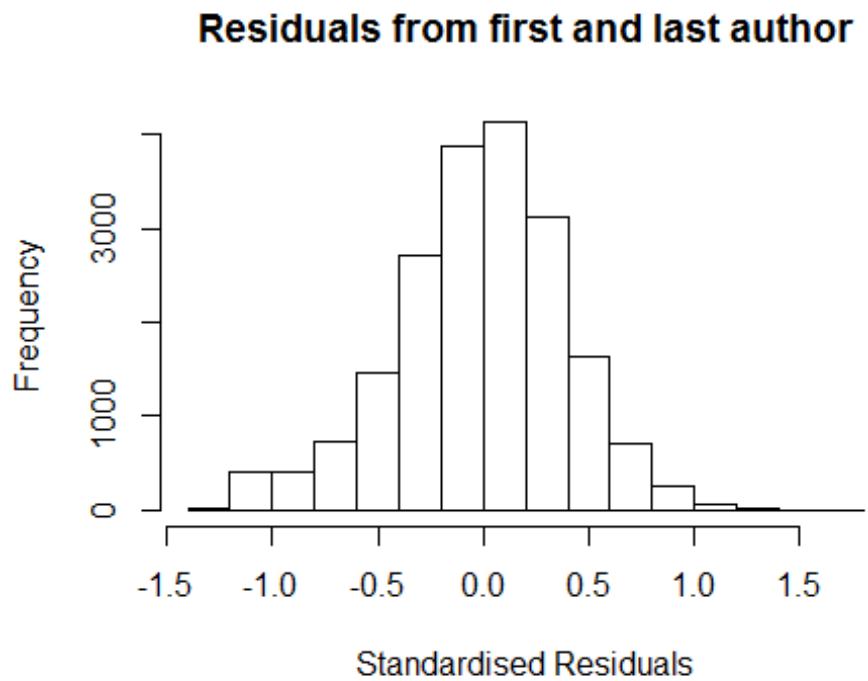



```
## [1] "Female first author team size 2018 geometric mean: 4.86435985214929"
## [1] "Male first author team size 2018 geometric mean: 4.53942498865901"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 270000, p-value = 0.04
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 4.72885493295085"
## [1] "Male last author team size 2018 geometric mean: 4.69300874331583"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 240000, p-value = 0.8
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##                     GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.036  1           1.018
## LastAuthorFemale  1.021  1           1.010
## UniqueAuthors     1.047  4           1.006
## Year              1.068 16           1.002
```



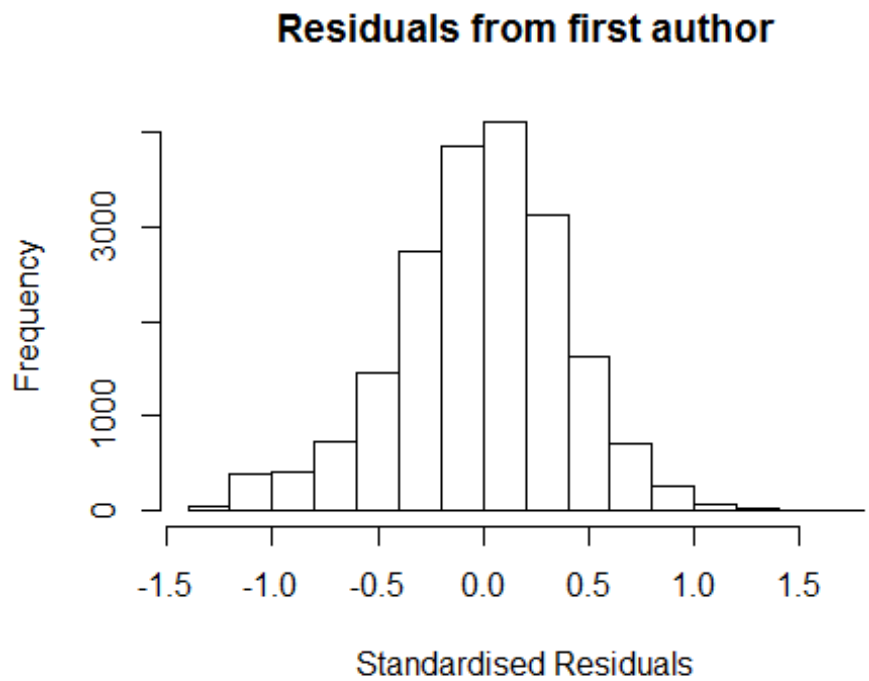

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
UniqueAuthors +
##     Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##     k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.2862 -0.2534  0.0083  0.2512  1.8028
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)          0.98331    0.02101   46.81  < 2e-16 ***
## FirstAuthorFemale1   0.03058    0.00585    5.23  1.7e-07 ***
## LastAuthorFemale1    0.03534    0.00641    5.51  3.6e-08 ***
## UniqueAuthors2       0.13446    0.01642    8.19  2.9e-16 ***
## UniqueAuthors3       0.16936    0.01606   10.55  < 2e-16 ***
## UniqueAuthors4       0.19041    0.01606   11.86  < 2e-16 ***
## UniqueAuthors5       0.23572    0.01533   15.38  < 2e-16 ***
## Year1997             0.00129    0.02118    0.06  0.95156
## Year1998             0.00547    0.02083    0.26  0.79277
## Year1999            -0.01594    0.01995   -0.80  0.42437
```

```
## Year2000           -0.00610    0.02035   -0.30  0.76454
## Year2001           -0.00414    0.02043   -0.20  0.83942
## Year2002           -0.06313    0.01934   -3.26  0.00110 **
## Year2003           -0.08746    0.01871   -4.67  3.0e-06 ***
## Year2004           -0.06632    0.01855   -3.58  0.00035 ***
## Year2005           -0.06001    0.01868   -3.21  0.00132 **
## Year2006           -0.06334    0.01848   -3.43  0.00061 ***
## Year2007           -0.05761    0.01808   -3.19  0.00144 **
## Year2008           -0.07492    0.01791   -4.18  2.9e-05 ***
## Year2009           -0.07108    0.01796   -3.96  7.6e-05 ***
## Year2010           -0.07733    0.01797   -4.30  1.7e-05 ***
## Year2011           -0.06933    0.01836   -3.78  0.00016 ***
## Year2012           -0.06868    0.01877   -3.66  0.00025 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.371
## Multiple R-squared:  0.0329, Adjusted R-squared:  0.0318
## Convergence in 15 IRWLS iterations
##
## Robustness weights:
##  observation 4346 is an outlier with |weight| = 0 ( < 5.1e-06);
##  1683 weights are ~= 1. The remaining 17840 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0016  0.8640  0.9500  0.8940  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          5.12e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.031  1           1.015
## LastAuthorFemale  1.020  1           1.010
## Year              1.028 16           1.001
```



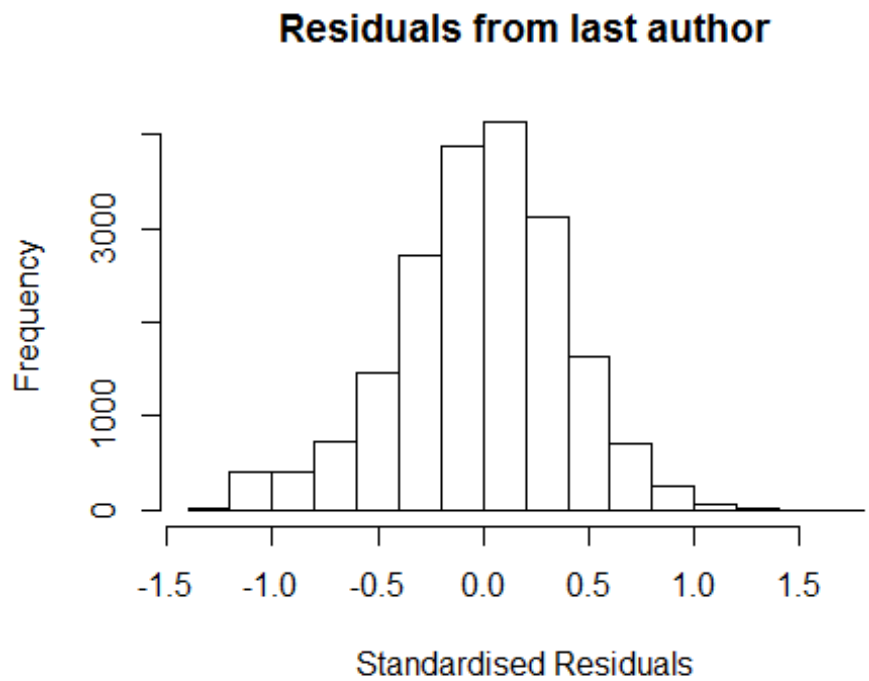

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.22859 -0.25449  0.00763  0.25005  1.63220
##
## Coefficients:
##                      Estimate Std. Error t value Pr(>|t|)
## (Intercept)          1.154154   0.015178   76.04  < 2e-16 ***
## FirstAuthorFemale1   0.037850   0.005880    6.44  1.2e-10 ***
## LastAuthorFemale1    0.032277   0.006447    5.01  5.6e-07 ***
## Year1997            -0.000701   0.021206   -0.03  0.97363
## Year1998             0.004313   0.020866    0.21  0.83624
## Year1999            -0.016322   0.020051   -0.81  0.41563
## Year2000            -0.006349   0.020458   -0.31  0.75629
## Year2001             0.003037   0.020512    0.15  0.88231
## Year2002            -0.061416   0.019423   -3.16  0.00157 **
## Year2003            -0.075921   0.018892   -4.02  5.9e-05 ***
## Year2004            -0.053791   0.018694   -2.88  0.00401 **
## Year2005            -0.048072   0.018781   -2.56  0.01049 *
```

```
## Year2006           -0.049283   0.018507   -2.66  0.00775 **
## Year2007           -0.040610   0.018125   -2.24  0.02506 *
## Year2008           -0.062844   0.017985   -3.49  0.00048 ***
## Year2009           -0.050771   0.017997   -2.82  0.00479 **
## Year2010           -0.057618   0.018070   -3.19  0.00143 **
## Year2011           -0.050995   0.018409   -2.77  0.00561 **
## Year2012           -0.044645   0.018800   -2.37  0.01757 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.373
## Multiple R-squared:  0.00752,    Adjusted R-squared:  0.0066
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
##  1690 weights are ~= 1. The remaining 17834 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0165  0.8630  0.9510  0.8930  0.9860  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          5.12e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##                   GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.02  1           1.010
## Year              1.02 16           1.001
```




```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.20627 -0.25035  0.00744  0.25021  1.62515
##
## Coefficients:
##                      Estimate Std. Error t value Pr(>|t|)
## (Intercept)          1.16e+00   1.52e-02   76.52  < 2e-16 ***
## FirstAuthorFemale1   4.16e-02   5.86e-03    7.10  1.3e-12 ***
## Year1997            -7.44e-06   2.12e-02    0.00  0.99972
## Year1998             4.89e-03   2.09e-02    0.23  0.81477
## Year1999            -1.59e-02   2.01e-02   -0.79  0.42875
## Year2000            -4.92e-03   2.05e-02   -0.24  0.80985
## Year2001             2.74e-03   2.05e-02    0.13  0.89368
## Year2002            -6.04e-02   1.94e-02   -3.11  0.00190 **
## Year2003            -7.50e-02   1.89e-02   -3.96  7.4e-05 ***
## Year2004            -5.21e-02   1.87e-02   -2.78  0.00539 **
## Year2005            -4.69e-02   1.88e-02   -2.49  0.01269 *
## Year2006            -4.78e-02   1.85e-02   -2.58  0.00994 **
```

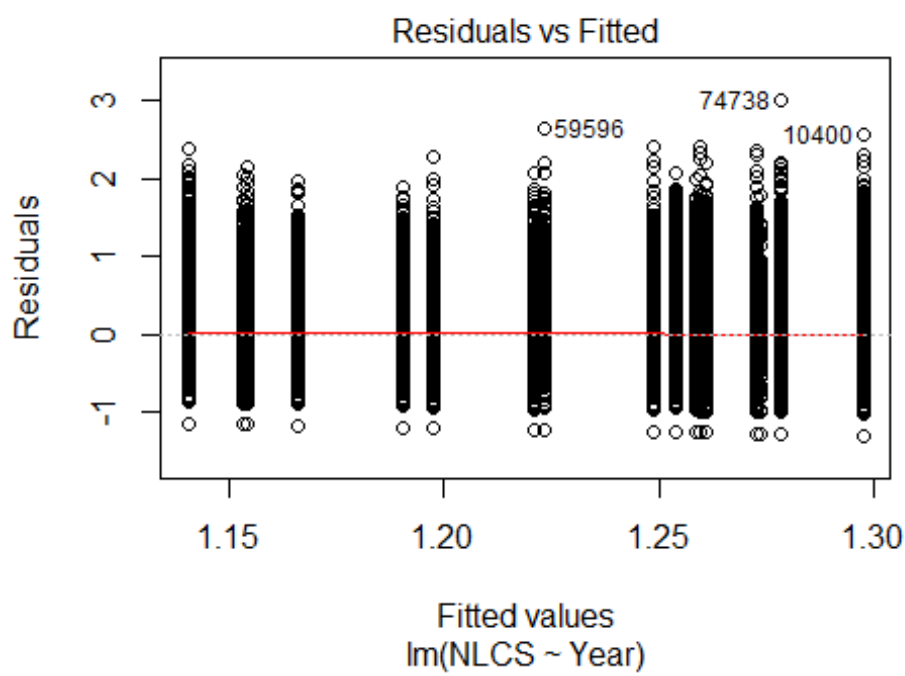
```
## Year2007           -3.88e-02   1.81e-02   -2.14  0.03276 *
## Year2008           -6.04e-02   1.80e-02   -3.36  0.00079 ***
## Year2009           -4.82e-02   1.80e-02   -2.68  0.00748 **
## Year2010           -5.54e-02   1.81e-02   -3.06  0.00222 **
## Year2011           -4.86e-02   1.84e-02   -2.64  0.00839 **
## Year2012           -4.12e-02   1.88e-02   -2.19  0.02866 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.374
## Multiple R-squared:  0.00615,    Adjusted R-squared:  0.00528
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
##  1659 weights are ~= 1. The remaining 17865 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0191  0.8630  0.9500  0.8930  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          5.12e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##                  GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.01  1           1.005
## Year             1.01 16           1.000
```



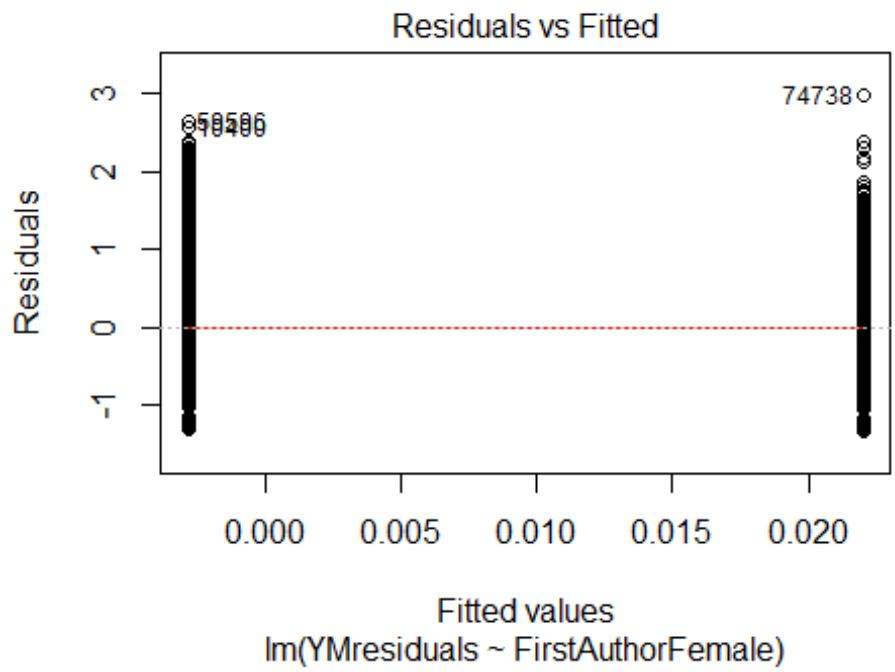

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.20722 -0.25204  0.00732  0.25030  1.62230
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)        1.164036   0.015049   77.35  < 2e-16 ***
## LastAuthorFemale1  0.037522   0.006416    5.85    5e-09 ***
## Year1997           0.000718   0.021217    0.03  0.97299
## Year1998           0.005390   0.020850    0.26  0.79603
## Year1999          -0.014355   0.020020   -0.72  0.47338
## Year2000          -0.006334   0.020426   -0.31  0.75647
## Year2001           0.005665   0.020492    0.28  0.78221
## Year2002          -0.059131   0.019360   -3.05  0.00226 **
## Year2003          -0.073031   0.018844   -3.88  0.00011 ***
## Year2004          -0.051216   0.018661   -2.74  0.00606 **
## Year2005          -0.045514   0.018762   -2.43  0.01528 *
## Year2006          -0.046025   0.018450   -2.49  0.01262 *
```

```
## Year2007          -0.036981   0.018108   -2.04  0.04114 *
## Year2008          -0.058689   0.017959   -3.27  0.00109 **
## Year2009          -0.045495   0.017941   -2.54  0.01123 *
## Year2010          -0.051770   0.018023   -2.87  0.00408 **
## Year2011          -0.045294   0.018378   -2.46  0.01373 *
## Year2012          -0.037402   0.018728   -2.00  0.04583 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.373
## Multiple R-squared:  0.0053, Adjusted R-squared:  0.00443
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
##  1738 weights are ~= 1. The remaining 17786 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0195  0.8630  0.9500  0.8930  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          5.12e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Sample size for the above analysis:  19524"
## [1] ""
## [1] ""
## [1] "####################################"
## [1] "Analysis of AJSC 3100"
## [1] "####################################"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  526 4559 4632 4843 3773 4836 5078 4999 4696 4687 4852 4373 4724 4420 4025
## 2011 2012
## 4374 4446
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  152 1758 1831 1953 1535 1671 1993 2034 2066 2163 2198 2026 2250 2117 2013
## 2011 2012
```

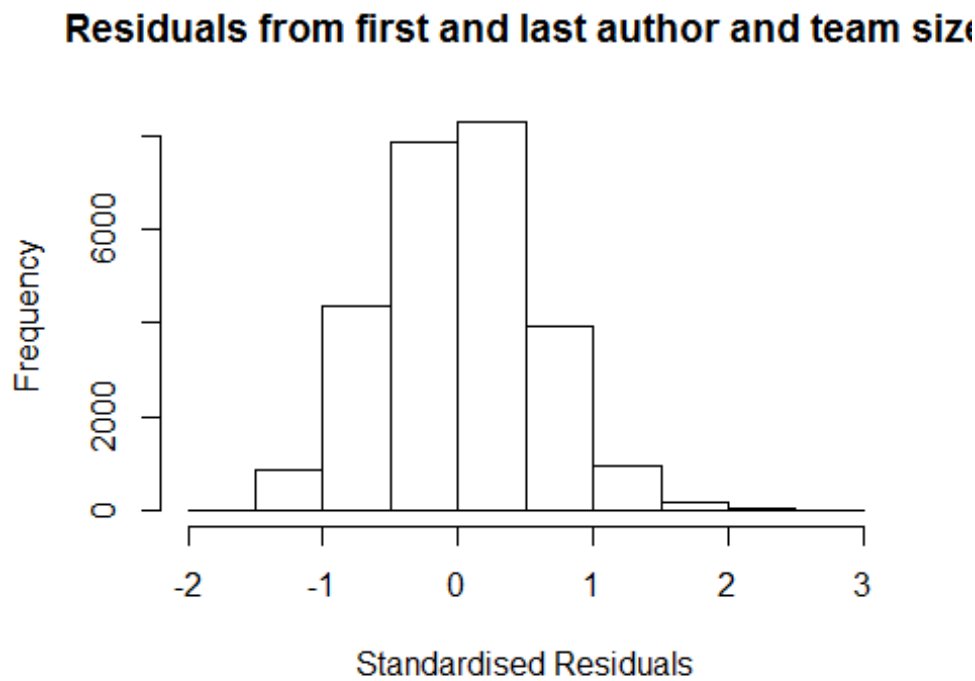
```
## 2201 2271
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  129 1437 1510 1594 1274 1351 1615 1668 1700 1792 1818 1696 1841 1766 1676
## 2011 2012
## 1796 1834
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
##  Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 180, df = 16, p-value <2e-16
```




```
##
##  Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 2.6, df = 1, p-value = 0.1
```



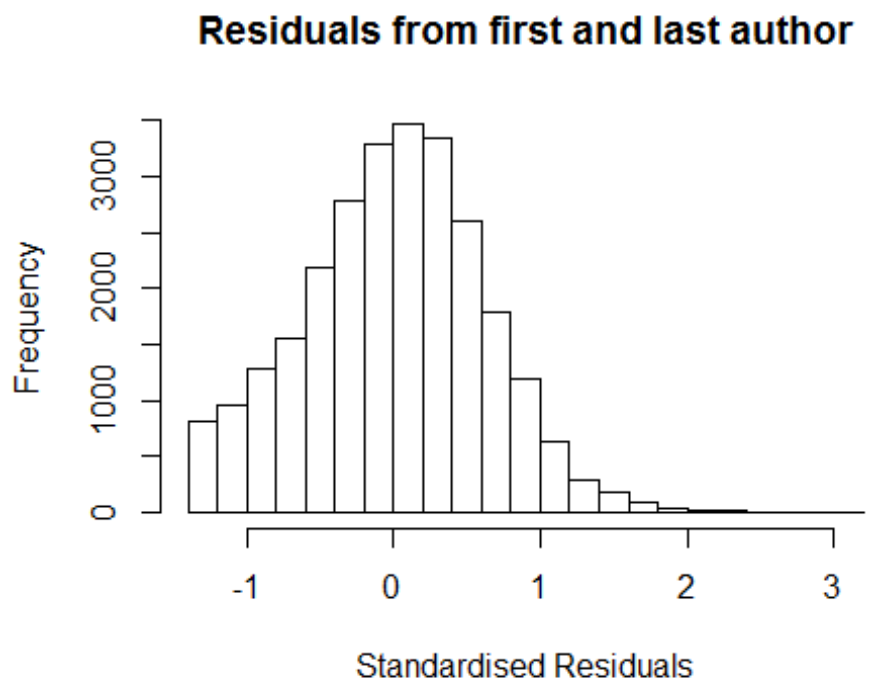

```
## [1] "Female first author team size 2018 geometric mean: 4.0574465577513"
## [1] "Male first author team size 2018 geometric mean: 3.19377552317283"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 480000, p-value = 6e-09
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 3.9111944995758"
## [1] "Male last author team size 2018 geometric mean: 3.25426095586965"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 380000, p-value = 7e-06
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##                     GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.044  1           1.022
## LastAuthorFemale  1.039  1           1.020
## UniqueAuthors     1.045  4           1.005
## Year              1.034 16           1.001
```




```
## [1] "List of  7 outliers with residuals above  2.5"
##          ScopusId  NLCS Year OneField Fields residuals
## 10018  4243684526 3.597 1998     3100      1     2.612
## 10400 78650234788 3.859 1998     3100      1     2.874
## 10908  0141465932 3.529 1998     2200      3     2.544
## 59596 34547309949 3.854 2007     2700      2     2.965
## 71138 65549134300 3.585 2009     2200      3     2.683
## 74738 78650092372 4.275 2010     2200      3     2.815
## 76404 77950857180 3.456 2010     2700      2     2.569
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
UniqueAuthors +
##     Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##     k.max = 1000))
##  \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.59822 -0.38414  0.00908  0.39019  2.96525
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)         0.945120   0.071735   13.18  < 2e-16 ***
## FirstAuthorFemale1 -0.039125   0.011605   -3.37  0.00075 ***
## LastAuthorFemale1   0.011233   0.012948    0.87  0.38567
## UniqueAuthors2      0.353959   0.010138   34.92  < 2e-16 ***
```

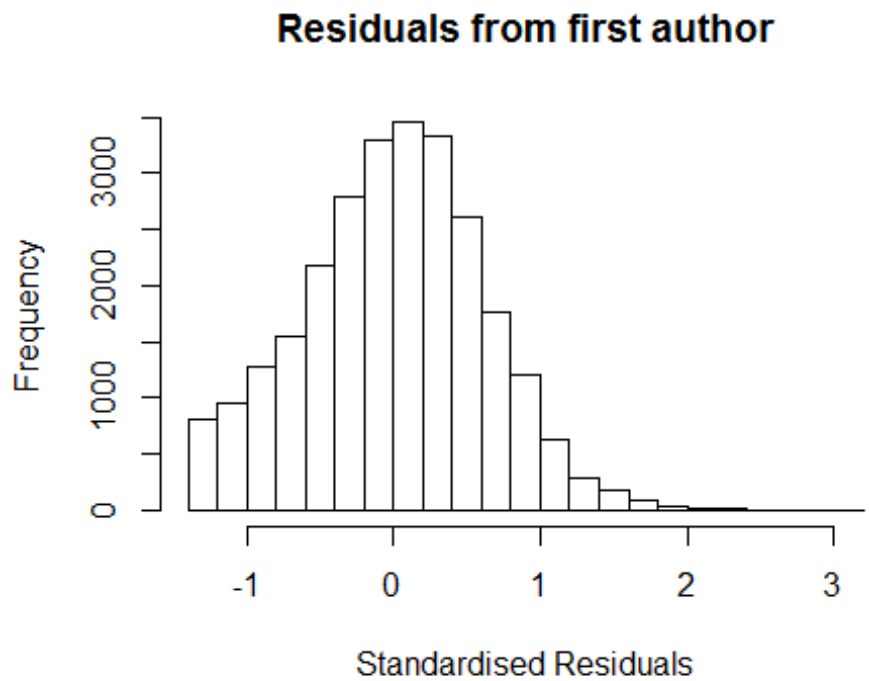
```
## UniqueAuthors3      0.446133   0.010961   40.70  < 2e-16 ***
## UniqueAuthors4      0.520916   0.013299   39.17  < 2e-16 ***
## UniqueAuthors5      0.612759   0.013498   45.40  < 2e-16 ***
## Year1997           -0.009094   0.073004   -0.12  0.90086
## Year1998            0.040343   0.072901    0.55  0.58000
## Year1999            0.000549   0.072805    0.01  0.99399
## Year2000           -0.062382   0.073004   -0.85  0.39284
## Year2001           -0.128705   0.073930   -1.74  0.08171 .
## Year2002           -0.095718   0.072984   -1.31  0.18970
## Year2003           -0.101897   0.073198   -1.39  0.16391
## Year2004           -0.084260   0.072803   -1.16  0.24713
## Year2005           -0.108197   0.072399   -1.49  0.13507
## Year2006           -0.082367   0.072425   -1.14  0.25544
## Year2007           -0.056365   0.072469   -0.78  0.43670
## Year2008           -0.036932   0.072451   -0.51  0.61023
## Year2009           -0.043586   0.072580   -0.60  0.54816
## Year2010           -0.058571   0.072704   -0.81  0.42048
## Year2011           -0.074944   0.072569   -1.03  0.30174
## Year2012           -0.092082   0.072715   -1.27  0.20540
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.578
## Multiple R-squared:  0.121,  Adjusted R-squared:  0.12
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
##  3 observations c(2826,16913,21772)
##   are outliers with |weight| = 0 ( < 3.8e-06);
##  2192 weights are ~= 1. The remaining 24302 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0003  0.8660  0.9510  0.9060  0.9860  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          3.77e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
```

```
##                     GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.014  1           1.007
## LastAuthorFemale  1.014  1           1.007
## Year              1.005 16           1.000
```



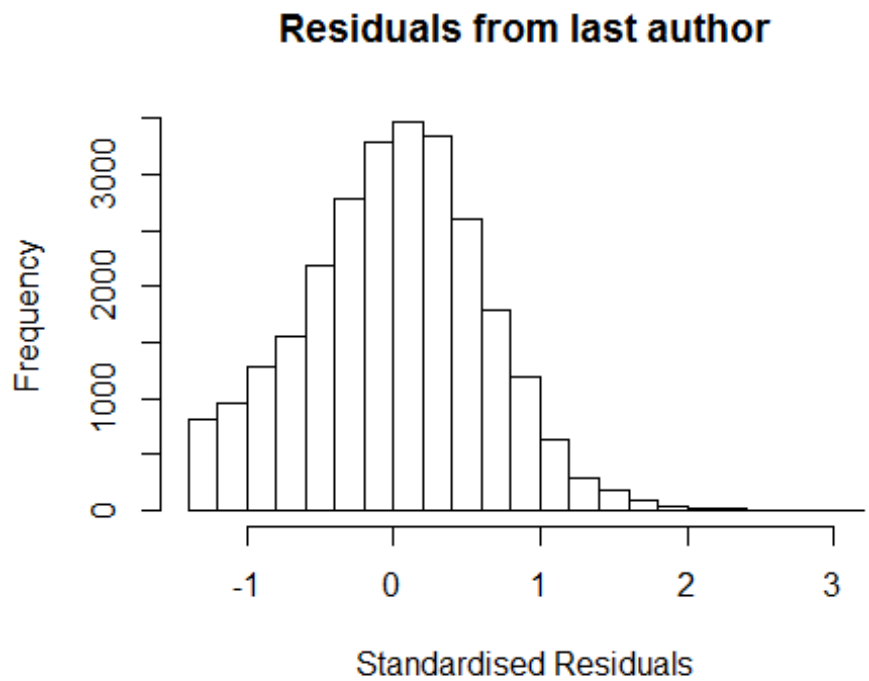

```
## [1] "List of  3 outliers with residuals above  2.5"
##          ScopusId  NLCS Year OneField Fields residuals
## 10400 78650234788 3.859 1998     3100      1     2.587
## 59596 34547309949 3.854 2007     2700      2     2.656
## 74738 78650092372 4.275 2010     2200      3     3.015
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.3280 -0.4101  0.0212  0.4127  3.0154
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)          1.2610     0.0699   18.04   <2e-16 ***
## FirstAuthorFemale1   0.0250     0.0122    2.04    0.041 *
## LastAuthorFemale1    0.0315     0.0140    2.25    0.024 *
## Year1997            -0.0299     0.0718   -0.42    0.677
## Year1998             0.0105     0.0717    0.15    0.883
```

```
## Year1999            -0.0250     0.0716   -0.35    0.727
## Year2000            -0.0722     0.0718   -1.01    0.315
## Year2001            -0.1616     0.0729   -2.22    0.027 *
## Year2002            -0.1204     0.0718   -1.68    0.094 .
## Year2003            -0.1148     0.0720   -1.60    0.111
## Year2004            -0.0936     0.0716   -1.31    0.191
## Year2005            -0.1310     0.0713   -1.84    0.066 .
## Year2006            -0.0939     0.0712   -1.32    0.187
## Year2007            -0.0635     0.0713   -0.89    0.373
## Year2008            -0.0255     0.0713   -0.36    0.721
## Year2009            -0.0167     0.0715   -0.23    0.816
## Year2010            -0.0264     0.0717   -0.37    0.713
## Year2011            -0.0288     0.0715   -0.40    0.687
## Year2012            -0.0394     0.0717   -0.55    0.583
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.612
## Multiple R-squared:  0.00624,    Adjusted R-squared:  0.00557
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
##  observation 21772 is an outlier with |weight| = 0 ( < 3.8e-06);
##  2184 weights are ~= 1. The remaining 24312 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0204  0.8650  0.9500  0.9060  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          3.77e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.002  1           1.001
## Year              1.002 16           1.000
```




```
## [1] "List of  3 outliers with residuals above  2.5"
##          ScopusId  NLCS Year OneField Fields residuals
## 10400 78650234788 3.859 1998     3100      1     2.587
## 59596 34547309949 3.854 2007     2700      2     2.656
## 74738 78650092372 4.275 2010     2200      3     3.015
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.3040 -0.4087  0.0208  0.4132  3.0075
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.26331    0.06992   18.07   <2e-16 ***
## FirstAuthorFemale1  0.03078    0.01215    2.53    0.011 *
## Year1997           -0.03054    0.07188   -0.42    0.671
## Year1998            0.00993    0.07172    0.14    0.890
## Year1999           -0.02508    0.07163   -0.35    0.726
## Year2000           -0.07264    0.07183   -1.01    0.312
## Year2001           -0.16209    0.07295   -2.22    0.026 *
## Year2002           -0.12099    0.07185   -1.68    0.092 .
## Year2003           -0.11488    0.07201   -1.60    0.111
## Year2004           -0.09379    0.07167   -1.31    0.191
```

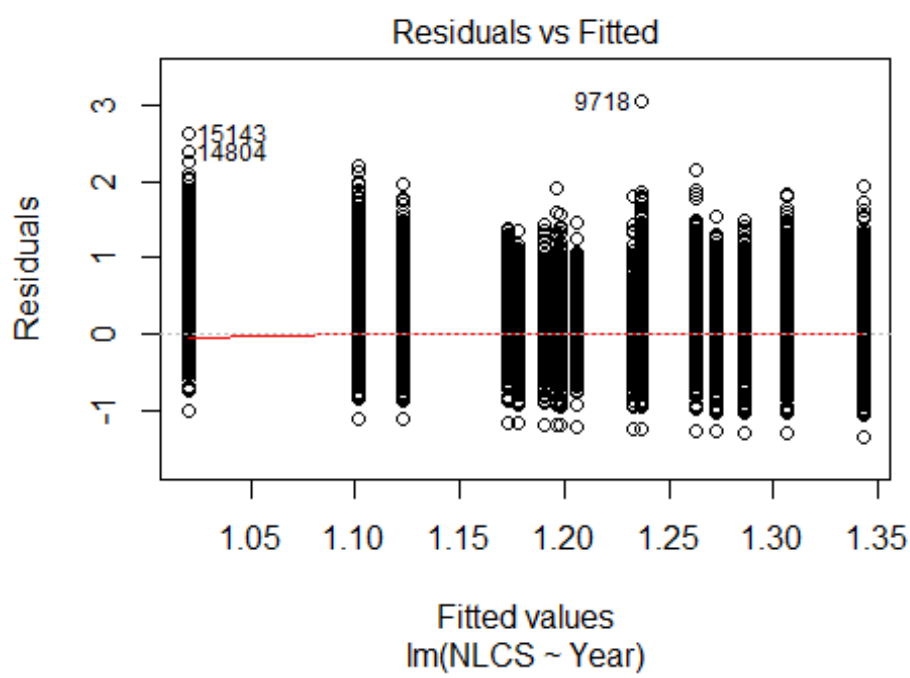
```
## Year2005           -0.13131    0.07133   -1.84    0.066 .
## Year2006           -0.09454    0.07126   -1.33    0.185
## Year2007           -0.06351    0.07131   -0.89    0.373
## Year2008           -0.02585    0.07129   -0.36    0.717
## Year2009           -0.01625    0.07154   -0.23    0.820
## Year2010           -0.02660    0.07170   -0.37    0.711
## Year2011           -0.02838    0.07151   -0.40    0.692
## Year2012           -0.03890    0.07175   -0.54    0.588
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.613
## Multiple R-squared:  0.00604,    Adjusted R-squared:  0.0054
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
##  observation 21772 is an outlier with |weight| = 0 ( < 3.8e-06);
##  2203 weights are ~= 1. The remaining 24293 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.021   0.865   0.950   0.906   0.985   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          3.77e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##                   GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.003  1           1.001
## Year             1.003 16           1.000
```



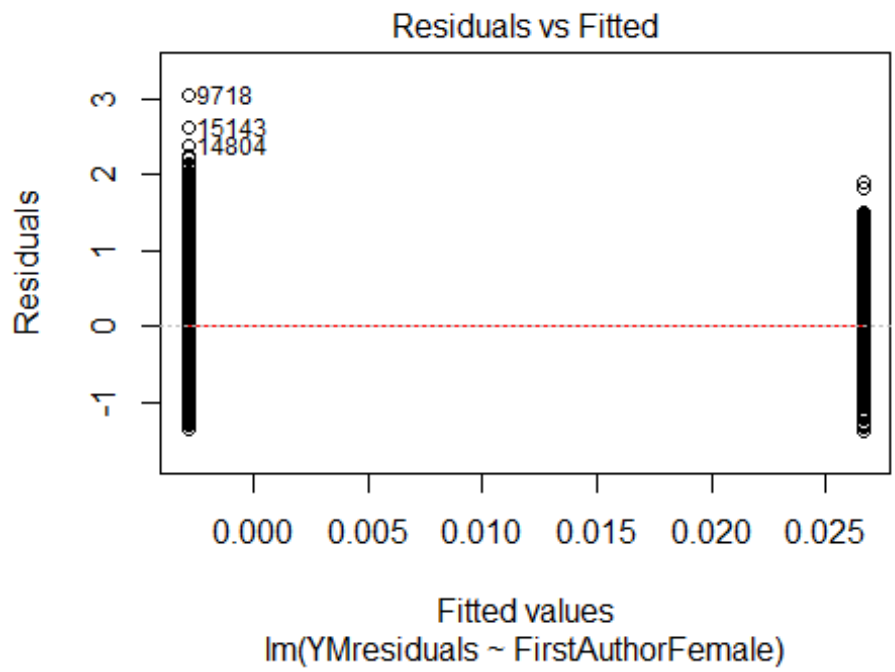

```
## [1] "List of  3 outliers with residuals above  2.5"
##          ScopusId  NLCS Year OneField Fields residuals
## 10400 78650234788 3.859 1998     3100      1     2.587
## 59596 34547309949 3.854 2007     2700      2     2.656
## 74738 78650092372 4.275 2010     2200      3     3.015
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.3104 -0.4086  0.0208  0.4129  3.0375
##
## Coefficients:
##                   Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.2630     0.0698   18.08   <2e-16 ***
## LastAuthorFemale1   0.0373     0.0139    2.68   0.0074 **
## Year1997           -0.0298     0.0718   -0.41   0.6786
## Year1998            0.0102     0.0716    0.14   0.8872
## Year1999           -0.0249     0.0715   -0.35   0.7276
## Year2000           -0.0715     0.0717   -1.00   0.3189
## Year2001           -0.1615     0.0729   -2.22   0.0266 *
## Year2002           -0.1203     0.0718   -1.68   0.0938 .
## Year2003           -0.1151     0.0719   -1.60   0.1096
## Year2004           -0.0936     0.0716   -1.31   0.1912
```

```
## Year2005           -0.1306     0.0712   -1.83   0.0667 .
## Year2006           -0.0936     0.0712   -1.32   0.1883
## Year2007           -0.0631     0.0712   -0.89   0.3758
## Year2008           -0.0250     0.0712   -0.35   0.7258
## Year2009           -0.0157     0.0714   -0.22   0.8257
## Year2010           -0.0254     0.0716   -0.36   0.7225
## Year2011           -0.0278     0.0714   -0.39   0.6973
## Year2012           -0.0387     0.0717   -0.54   0.5889
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.612
## Multiple R-squared:  0.00608,    Adjusted R-squared:  0.00544
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
##  observation 21772 is an outlier with |weight| = 0 ( < 3.8e-06);
##  2193 weights are ~= 1. The remaining 24303 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0208  0.8640  0.9500  0.9060  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          3.77e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Sample size for the above analysis:  26497"
## [1] ""
## [1] ""
## [1] "####################################"
## [1] "Analysis of AJSC 3101"
## [1] "####################################"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 2222 2110 2266 2529 2542 2585 2554 2289 2604 2338 2675 2335 2255 1793 1769
## 2011 2012
## 1651 1797
##
```

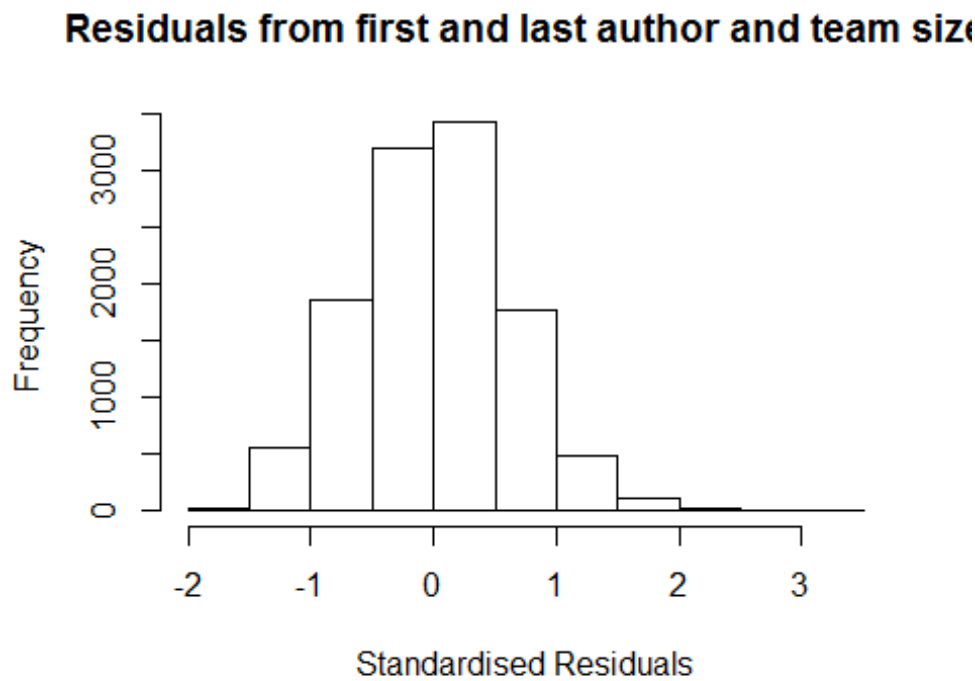
```
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  702  697  778  817 1015  966  990  884  990  834  938  870  806  775  723
## 2011 2012
##  751  770
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  566  564  633  678  851  812  804  712  794  608  709  692  623  618  572
## 2011 2012
##  601  602
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
##  Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 1200, df = 16, p-value <2e-16
```




```
##
##  Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 8, df = 1, p-value = 0.005
```

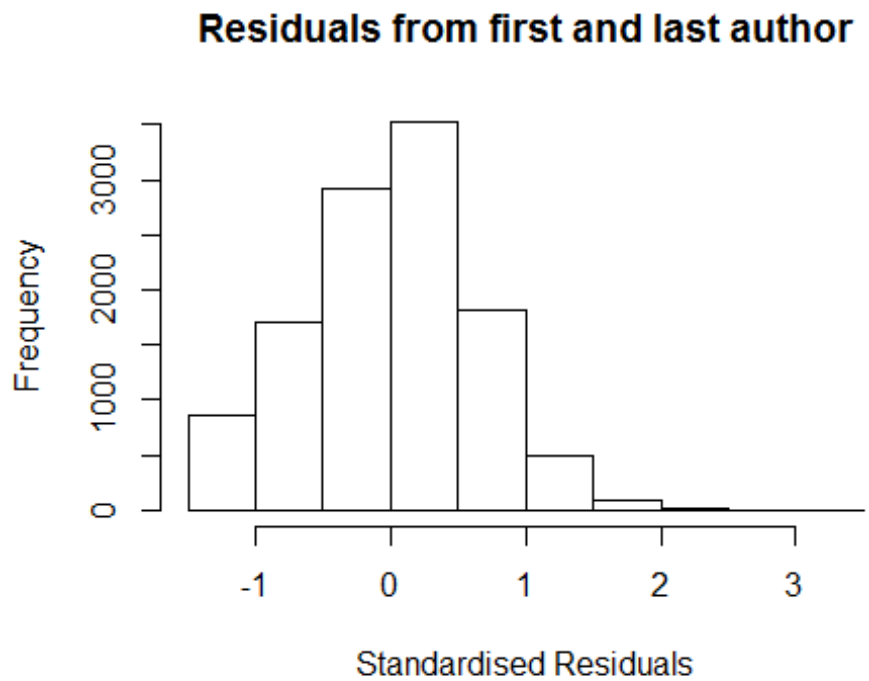



```
## [1] "Female first author team size 2018 geometric mean: 3.11935230973174"
## [1] "Male first author team size 2018 geometric mean: 2.62731156540493"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 27000, p-value = 0.02
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 3.20395944502861"
## [1] "Male last author team size 2018 geometric mean: 2.62139785018015"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 25000, p-value = 0.03
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.063  1           1.031
## LastAuthorFemale  1.059  1           1.029
## UniqueAuthors     1.118  4           1.014
## Year              1.128 16           1.004
```



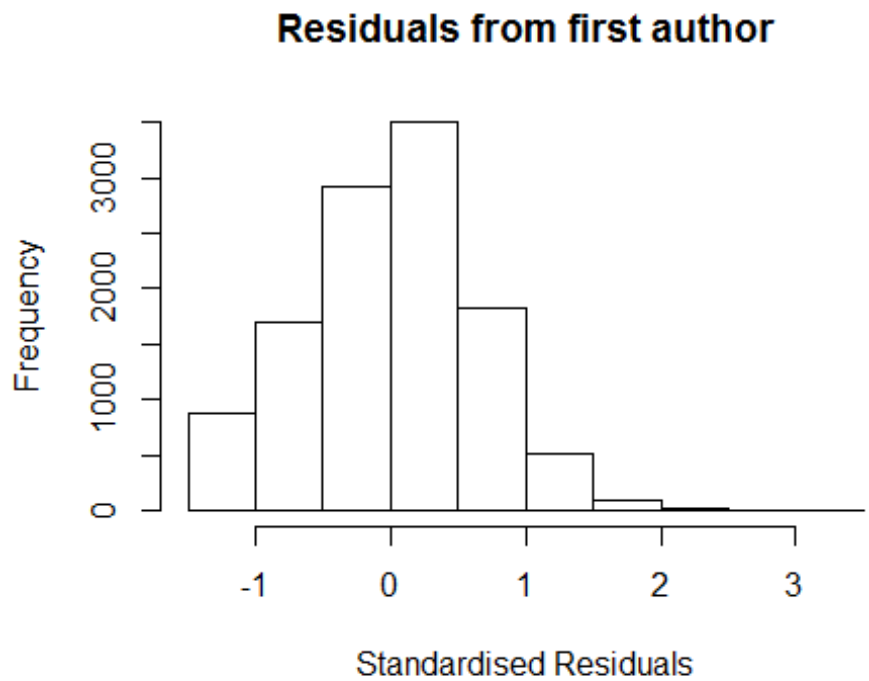

```
## [1] "List of  3 outliers with residuals above  2.5"
##         ScopusId  NLCS Year OneField Fields residuals
## 9718  0033239662 4.280 1999     2600      2     3.253
## 13512 0035423652 3.262 2001     3101      2     2.590
## 15143 0035288601 3.646 2001     3101      4     2.974
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
UniqueAuthors +
##     Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##     k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.5893 -0.4019  0.0136  0.4108  3.2525
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)          1.0907     0.0292   37.39  < 2e-16 ***
## FirstAuthorFemale1   0.0226     0.0193    1.17  0.24195
## LastAuthorFemale1   -0.0534     0.0204   -2.62  0.00890 **
## UniqueAuthors2       0.3228     0.0175   18.42  < 2e-16 ***
## UniqueAuthors3       0.4133     0.0184   22.49  < 2e-16 ***
## UniqueAuthors4       0.4571     0.0206   22.21  < 2e-16 ***
## UniqueAuthors5       0.4760     0.0187   25.52  < 2e-16 ***
## Year1997            -0.0391     0.0396   -0.99  0.32384
```

```
## Year1998            -0.0555     0.0397   -1.40  0.16173
## Year1999            -0.0632     0.0408   -1.55  0.12141
## Year2000            -0.2323     0.0408   -5.70  1.2e-08 ***
## Year2001            -0.4188     0.0427   -9.80  < 2e-16 ***
## Year2002            -0.3028     0.0422   -7.18  7.4e-13 ***
## Year2003            -0.1337     0.0357   -3.74  0.00018 ***
## Year2004            -0.2613     0.0369   -7.08  1.5e-12 ***
## Year2005            -0.1538     0.0352   -4.37  1.3e-05 ***
## Year2006            -0.2311     0.0332   -6.96  3.7e-12 ***
## Year2007            -0.2675     0.0335   -7.98  1.6e-15 ***
## Year2008            -0.2369     0.0340   -6.97  3.4e-12 ***
## Year2009            -0.2124     0.0334   -6.35  2.2e-10 ***
## Year2010            -0.1932     0.0336   -5.74  9.5e-09 ***
## Year2011            -0.2332     0.0343   -6.79  1.1e-11 ***
## Year2012            -0.2622     0.0343   -7.65  2.1e-14 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.598
## Multiple R-squared:  0.107,  Adjusted R-squared:  0.106
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
##  2 observations c(2199,3715) are outliers with |weight| = 0 ( < 8.7e-06);
##  905 weights are ~= 1. The remaining 10532 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0208  0.8600  0.9490  0.9050  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          8.74e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.043  1           1.022
## LastAuthorFemale  1.039  1           1.019
## Year              1.014 16           1.000
```



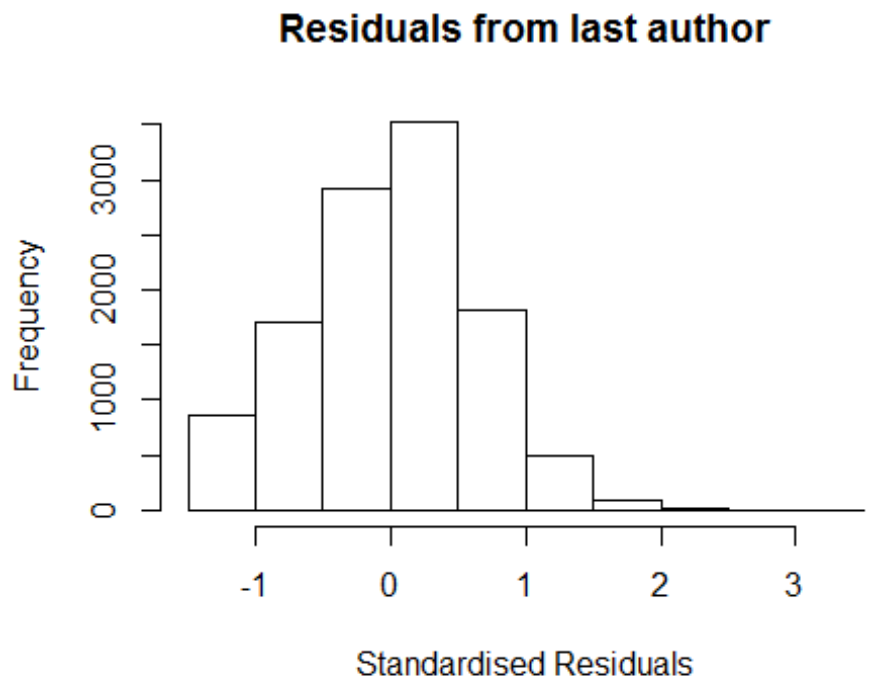

```
## [1] "List of  3 outliers with residuals above  2.5"
##         ScopusId  NLCS Year OneField Fields residuals
## 9718  0033239662 4.280 1999     2600      2     3.046
## 14804 0000653947 3.412 2001     1706      2     2.535
## 15143 0035288601 3.646 2001     3101      4     2.769
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.3836 -0.4249  0.0287  0.4254  3.0458
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)          1.3251     0.0286   46.31  < 2e-16 ***
## FirstAuthorFemale1   0.0585     0.0201    2.91  0.00362 **
## LastAuthorFemale1   -0.0430     0.0214   -2.01  0.04432 *
## Year1997            -0.0284     0.0403   -0.70  0.48201
## Year1998            -0.0637     0.0403   -1.58  0.11374
## Year1999            -0.0909     0.0431   -2.11  0.03486 *
## Year2000            -0.2584     0.0431   -6.00  2.0e-09 ***
## Year2001            -0.4477     0.0479   -9.36  < 2e-16 ***
## Year2002            -0.2873     0.0452   -6.36  2.1e-10 ***
## Year2003            -0.0760     0.0367   -2.07  0.03808 *
```

```
## Year2004            -0.2202     0.0389   -5.66  1.6e-08 ***
## Year2005            -0.0523     0.0359   -1.46  0.14470
## Year2006            -0.1358     0.0338   -4.02  5.8e-05 ***
## Year2007            -0.1575     0.0344   -4.57  4.9e-06 ***
## Year2008            -0.1354     0.0351   -3.86  0.00011 ***
## Year2009            -0.1220     0.0344   -3.55  0.00039 ***
## Year2010            -0.1028     0.0344   -2.99  0.00280 **
## Year2011            -0.1363     0.0354   -3.85  0.00012 ***
## Year2012            -0.1630     0.0351   -4.65  3.3e-06 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.612
## Multiple R-squared:  0.0293, Adjusted R-squared:  0.0277
## Convergence in 15 IRWLS iterations
##
## Robustness weights:
##  observation 2199 is an outlier with |weight| = 0 ( < 8.7e-06);
##  927 weights are ~= 1. The remaining 10511 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0044  0.8520  0.9470  0.9020  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          8.74e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.009  1           1.005
## Year              1.009 16           1.000
```




```
## [1] "List of  3 outliers with residuals above  2.5"
##         ScopusId  NLCS Year OneField Fields residuals
## 9718  0033239662 4.280 1999     2600      2     3.046
## 14804 0000653947 3.412 2001     1706      2     2.535
## 15143 0035288601 3.646 2001     3101      4     2.769
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.3718 -0.4252  0.0294  0.4272  3.0477
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)          1.3233     0.0286   46.27  < 2e-16 ***
## FirstAuthorFemale1   0.0485     0.0200    2.43  0.01509 *
## Year1997            -0.0282     0.0404   -0.70  0.48485
## Year1998            -0.0634     0.0403   -1.57  0.11590
## Year1999            -0.0910     0.0431   -2.11  0.03462 *
## Year2000            -0.2598     0.0430   -6.04  1.6e-09 ***
## Year2001            -0.4494     0.0479   -9.39  < 2e-16 ***
## Year2002            -0.2880     0.0452   -6.37  2.0e-10 ***
## Year2003            -0.0772     0.0366   -2.11  0.03524 *
## Year2004            -0.2210     0.0389   -5.68  1.4e-08 ***
```

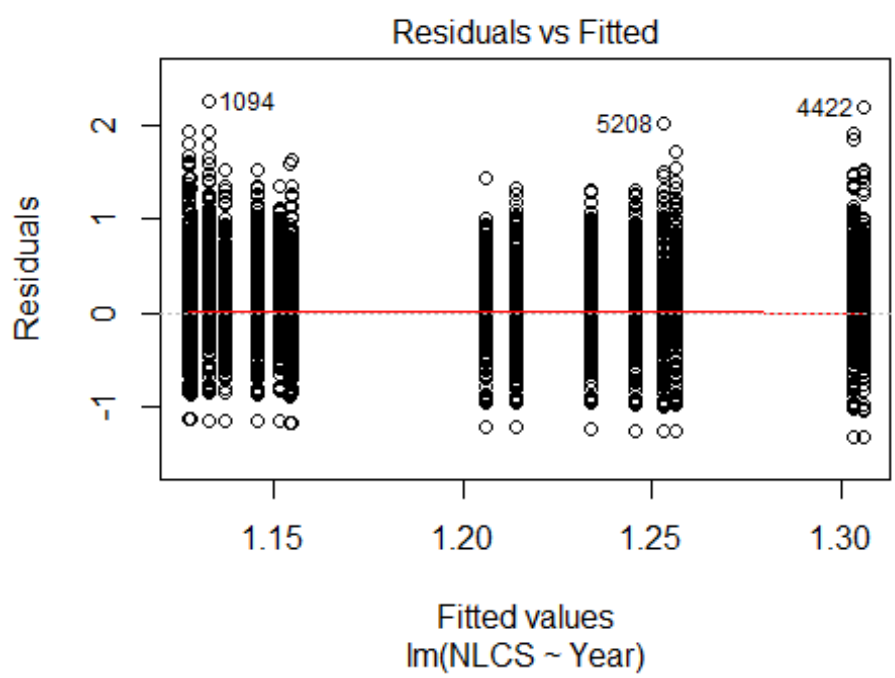
```
## Year2005            -0.0539     0.0359   -1.50  0.13276
## Year2006            -0.1368     0.0338   -4.05  5.1e-05 ***
## Year2007            -0.1587     0.0344   -4.61  4.1e-06 ***
## Year2008            -0.1368     0.0351   -3.90  9.5e-05 ***
## Year2009            -0.1236     0.0344   -3.60  0.00033 ***
## Year2010            -0.1038     0.0344   -3.02  0.00254 **
## Year2011            -0.1377     0.0354   -3.89  0.00010 ***
## Year2012            -0.1638     0.0351   -4.67  3.0e-06 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.612
## Multiple R-squared:  0.029,  Adjusted R-squared:  0.0275
## Convergence in 15 IRWLS iterations
##
## Robustness weights:
##  observation 2199 is an outlier with |weight| = 0 ( < 8.7e-06);
##  910 weights are ~= 1. The remaining 10528 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0041  0.8540  0.9470  0.9020  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          8.74e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##                   GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.005  1           1.003
## Year             1.005 16           1.000
```



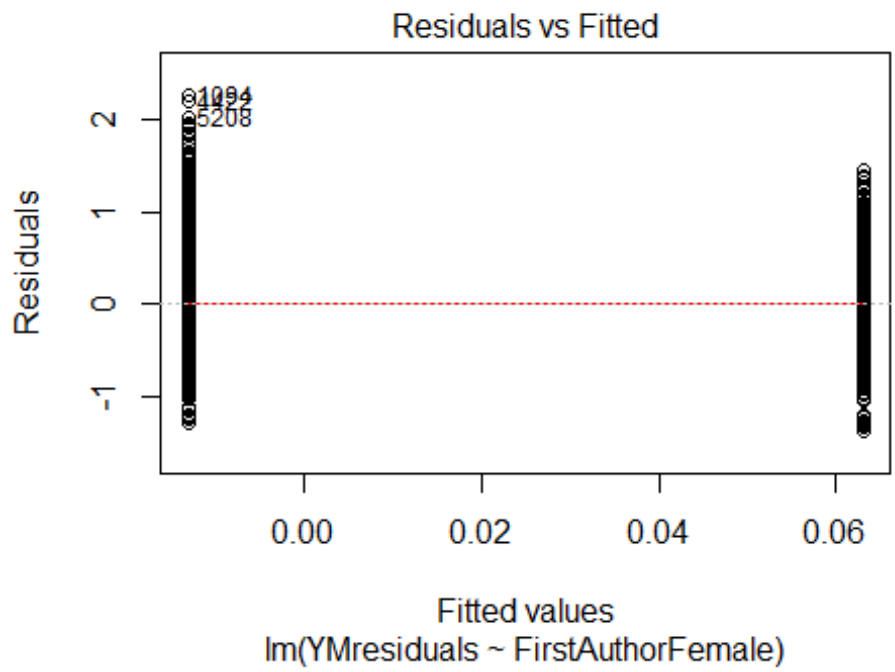

```
## [1] "List of  3 outliers with residuals above  2.5"
##         ScopusId  NLCS Year OneField Fields residuals
## 9718  0033239662 4.280 1999     2600      2     3.046
## 14804 0000653947 3.412 2001     1706      2     2.535
## 15143 0035288601 3.646 2001     3101      4     2.769
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.3290 -0.4221  0.0286  0.4258  3.0433
##
## Coefficients:
##                   Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.3290     0.0286   46.43  < 2e-16 ***
## LastAuthorFemale1  -0.0272     0.0209   -1.30  0.19250
## Year1997           -0.0284     0.0404   -0.70  0.48170
## Year1998           -0.0641     0.0404   -1.59  0.11214
## Year1999           -0.0923     0.0431   -2.14  0.03218 *
## Year2000           -0.2574     0.0431   -5.98  2.3e-09 ***
## Year2001           -0.4485     0.0479   -9.36  < 2e-16 ***
## Year2002           -0.2864     0.0453   -6.33  2.6e-10 ***
## Year2003           -0.0765     0.0367   -2.08  0.03731 *
## Year2004           -0.2200     0.0390   -5.65  1.7e-08 ***
```

```
## Year2005           -0.0523     0.0359   -1.46  0.14473
## Year2006           -0.1355     0.0338   -4.01  6.2e-05 ***
## Year2007           -0.1566     0.0345   -4.54  5.6e-06 ***
## Year2008           -0.1349     0.0351   -3.84  0.00012 ***
## Year2009           -0.1201     0.0344   -3.49  0.00048 ***
## Year2010           -0.1006     0.0344   -2.93  0.00344 **
## Year2011           -0.1351     0.0355   -3.81  0.00014 ***
## Year2012           -0.1606     0.0350   -4.58  4.6e-06 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.612
## Multiple R-squared:  0.0286, Adjusted R-squared:  0.0272
## Convergence in 15 IRWLS iterations
##
## Robustness weights:
##  observation 2199 is an outlier with |weight| = 0 ( < 8.7e-06);
##  936 weights are ~= 1. The remaining 10502 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0047  0.8530  0.9460  0.9020  0.9840  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          8.74e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Sample size for the above analysis:  11439"
## [1] ""
## [1] ""
## [1] "####################################"
## [1] "Analysis of AJSC 3102"
## [1] "####################################"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  901  866  745  632  704  660  555  602  540  637  759  715  678  704  623
## 2011 2012
##  669  692
##
```

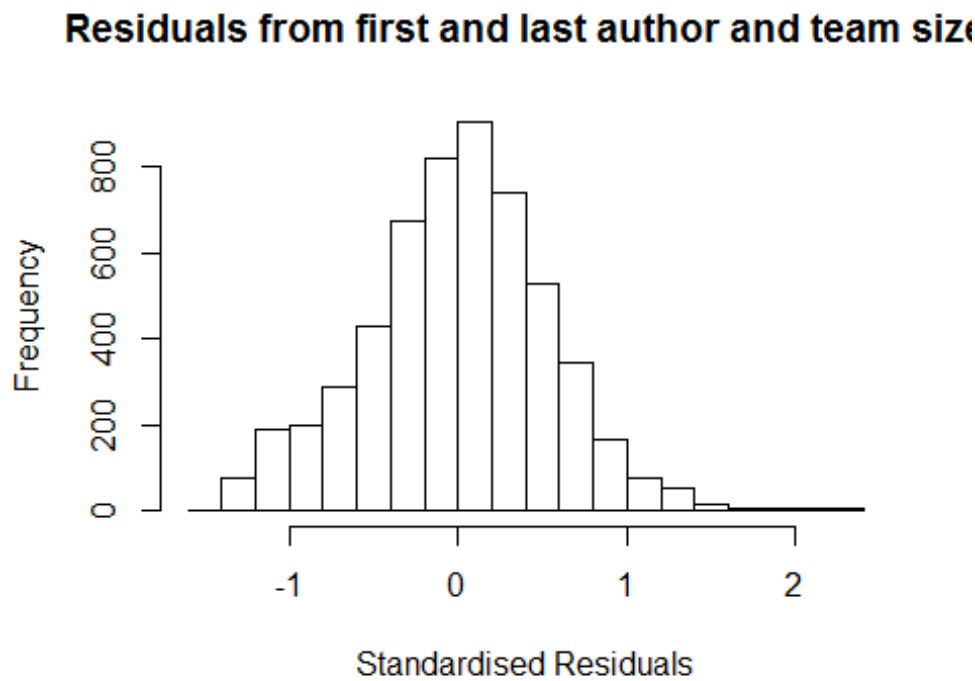
```
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  457  465  358  288  370  252  337  360  275  372  430  444  396  384  379
## 2011 2012
##  399  423
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  398  408  311  249  323  215  288  319  228  319  362  386  352  336  325
## 2011 2012
##  330  356
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
##  Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 230, df = 16, p-value <2e-16
```




```
##
##  Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 18, df = 1, p-value = 2e-05
```

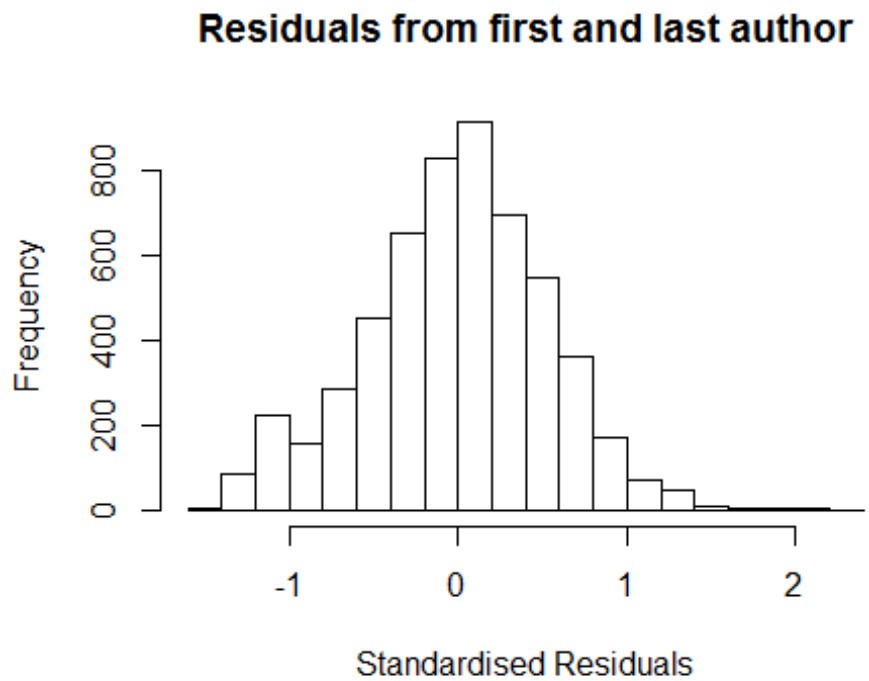



```
## [1] "Female first author team size 2018 geometric mean: 3.07996068542713"
## [1] "Male first author team size 2018 geometric mean: 2.93496958267232"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 9700, p-value = 0.5
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 2.90825976773522"
## [1] "Male last author team size 2018 geometric mean: 2.98270565709838"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 8200, p-value = 0.8
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.043  1           1.021
## LastAuthorFemale  1.040  1           1.020
## UniqueAuthors     1.082  4           1.010
## Year              1.088 16           1.003
```



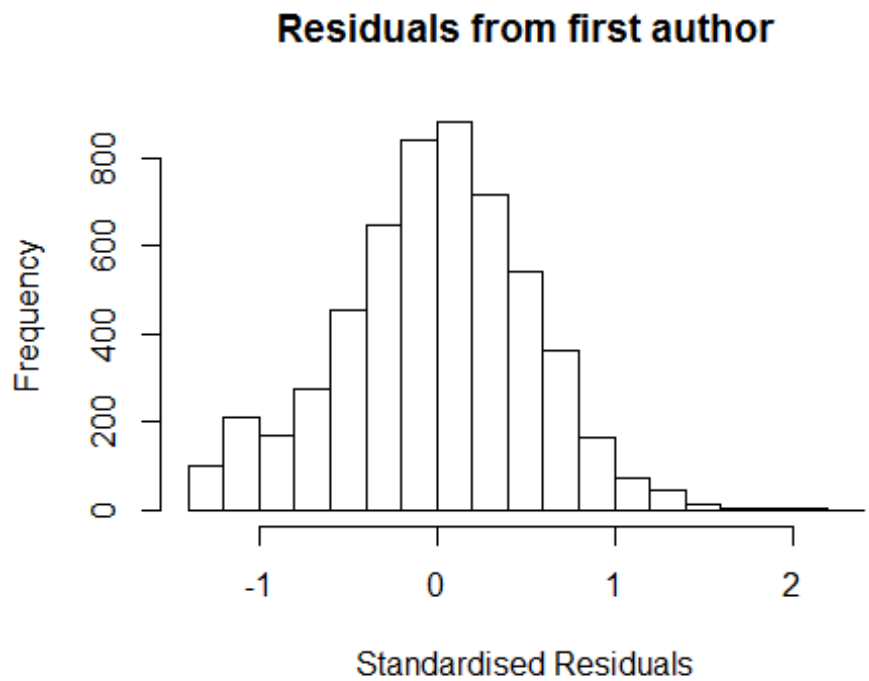

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
UniqueAuthors +
##     Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##     k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.4374 -0.3323  0.0207  0.3410  2.3491
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)          0.95946    0.04375   21.93  < 2e-16 ***
## FirstAuthorFemale1   0.05062    0.01812    2.79   0.0052 **
## LastAuthorFemale1    0.09533    0.02065    4.62  4.0e-06 ***
## UniqueAuthors2       0.15628    0.02377    6.58  5.3e-11 ***
## UniqueAuthors3       0.18046    0.02486    7.26  4.5e-13 ***
## UniqueAuthors4       0.24929    0.02856    8.73  < 2e-16 ***
## UniqueAuthors5       0.27262    0.02868    9.51  < 2e-16 ***
## Year1997             0.04516    0.05522    0.82   0.4135
## Year1998             0.13341    0.05129    2.60   0.0093 **
## Year1999             0.03480    0.05226    0.67   0.5055
```

```
## Year2000            0.15472    0.05321    2.91   0.0037 **
## Year2001            0.18046    0.05536    3.26   0.0011 **
## Year2002            0.09580    0.05042    1.90   0.0575 .
## Year2003            0.11911    0.05116    2.33   0.0199 *
## Year2004            0.07063    0.05157    1.37   0.1709
## Year2005            0.07221    0.04948    1.46   0.1445
## Year2006            0.07502    0.04719    1.59   0.1120
## Year2007           -0.00393    0.04604   -0.09   0.9320
## Year2008           -0.00890    0.04658   -0.19   0.8484
## Year2009           -0.04448    0.04677   -0.95   0.3417
## Year2010           -0.01119    0.04697   -0.24   0.8116
## Year2011           -0.01210    0.04867   -0.25   0.8037
## Year2012           -0.00646    0.04925   -0.13   0.8956
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.504
## Multiple R-squared:  0.0459, Adjusted R-squared:  0.042
## Convergence in 15 IRWLS iterations
##
## Robustness weights:
##  478 weights are ~= 1. The remaining 5027 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0001  0.8590  0.9500  0.8970  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.82e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.038  1           1.019
## LastAuthorFemale  1.044  1           1.022
## Year              1.019 16           1.001
```



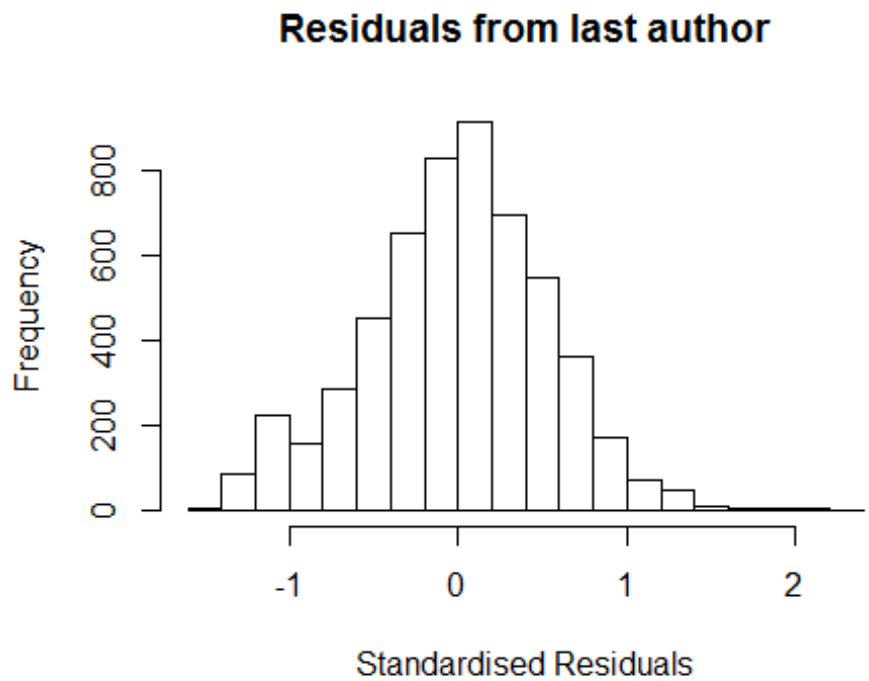

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##    Min     1Q Median     3Q    Max
## -1.402 -0.343  0.013  0.354  2.244
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)           1.0909     0.0404   26.98  < 2e-16 ***
## FirstAuthorFemale1    0.0680     0.0182    3.74  0.00018 ***
## LastAuthorFemale1     0.1029     0.0207    4.97  6.8e-07 ***
## Year1997              0.0522     0.0557    0.94  0.34881
## Year1998              0.1491     0.0521    2.86  0.00420 **
## Year1999              0.0293     0.0527    0.56  0.57787
## Year2000              0.1669     0.0532    3.14  0.00171 **
## Year2001              0.2102     0.0551    3.81  0.00014 ***
## Year2002              0.1086     0.0504    2.15  0.03132 *
## Year2003              0.1399     0.0513    2.73  0.00637 **
## Year2004              0.0848     0.0518    1.64  0.10175
## Year2005              0.0963     0.0500    1.93  0.05414 .
```

```
## Year2006             0.0999     0.0475    2.11  0.03522 *
## Year2007             0.0264     0.0462    0.57  0.56748
## Year2008             0.0149     0.0472    0.31  0.75288
## Year2009            -0.0119     0.0470   -0.25  0.79969
## Year2010             0.0162     0.0475    0.34  0.73325
## Year2011             0.0242     0.0488    0.50  0.62031
## Year2012             0.0315     0.0493    0.64  0.52331
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.513
## Multiple R-squared:  0.0205, Adjusted R-squared:  0.0172
## Convergence in 15 IRWLS iterations
##
## Robustness weights:
##  449 weights are ~= 1. The remaining 5056 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0162  0.8620  0.9500  0.8980  0.9870  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.82e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.007  1           1.003
## Year              1.007 16           1.000
```




```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.3515 -0.3405  0.0179  0.3529  2.2343
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)           1.1040     0.0404   27.35  < 2e-16 ***
## FirstAuthorFemale1    0.0893     0.0179    5.00    6e-07 ***
## Year1997              0.0487     0.0560    0.87  0.38446
## Year1998              0.1481     0.0520    2.85  0.00444 **
## Year1999              0.0256     0.0529    0.48  0.62893
## Year2000              0.1582     0.0531    2.98  0.00292 **
## Year2001              0.2071     0.0554    3.74  0.00019 ***
## Year2002              0.1085     0.0507    2.14  0.03241 *
## Year2003              0.1375     0.0514    2.67  0.00750 **
## Year2004              0.0812     0.0519    1.57  0.11741
## Year2005              0.0942     0.0500    1.88  0.05978 .
## Year2006              0.0982     0.0475    2.07  0.03889 *
```

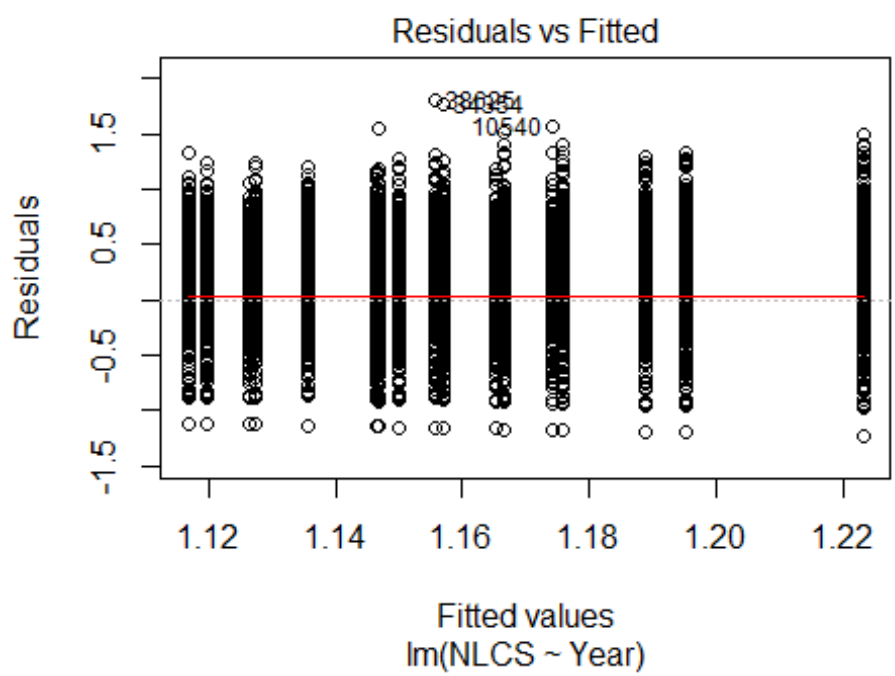
```
## Year2007             0.0235     0.0462    0.51  0.61044
## Year2008             0.0158     0.0473    0.33  0.73829
## Year2009            -0.0133     0.0470   -0.28  0.77786
## Year2010             0.0152     0.0476    0.32  0.74952
## Year2011             0.0246     0.0490    0.50  0.61507
## Year2012             0.0253     0.0494    0.51  0.60929
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.513
## Multiple R-squared:  0.0162, Adjusted R-squared:  0.0131
## Convergence in 15 IRWLS iterations
##
## Robustness weights:
##  466 weights are ~= 1. The remaining 5039 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0188  0.8620  0.9500  0.8980  0.9860  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.82e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##                   GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.012  1           1.006
## Year             1.012 16           1.000
```



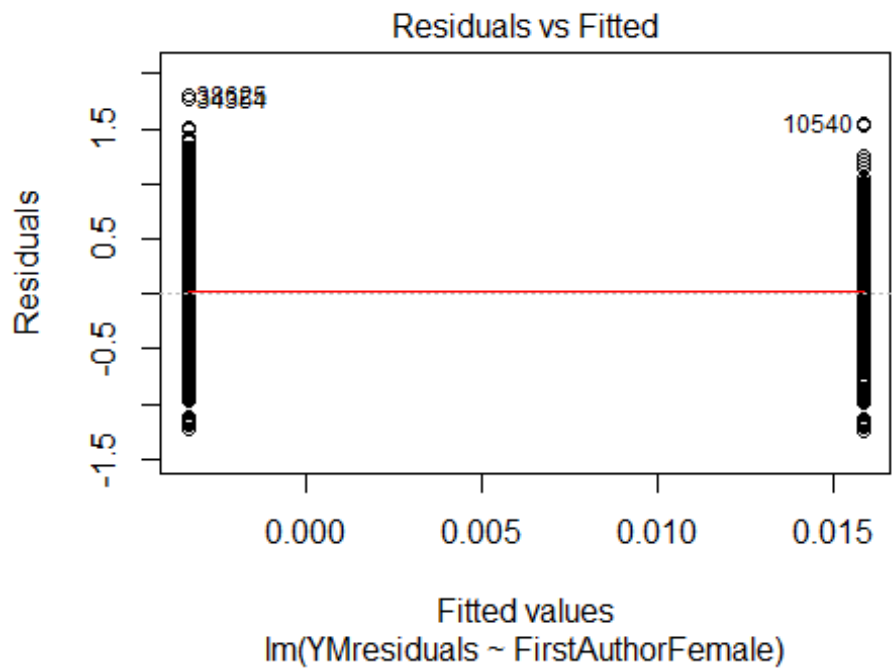

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##    Min     1Q Median     3Q    Max
## -1.358 -0.342  0.016  0.353  2.235
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)          1.1003     0.0406   27.13  < 2e-16 ***
## LastAuthorFemale1    0.1200     0.0203    5.91  3.6e-09 ***
## Year1997             0.0514     0.0560    0.92  0.35872
## Year1998             0.1518     0.0522    2.91  0.00367 **
## Year1999             0.0306     0.0531    0.58  0.56493
## Year2000             0.1670     0.0535    3.12  0.00181 **
## Year2001             0.2107     0.0554    3.80  0.00015 ***
## Year2002             0.1080     0.0507    2.13  0.03321 *
## Year2003             0.1374     0.0514    2.67  0.00754 **
## Year2004             0.0844     0.0520    1.62  0.10453
## Year2005             0.0959     0.0502    1.91  0.05630 .
## Year2006             0.1013     0.0477    2.12  0.03378 *
```

```
## Year2007            0.0281     0.0464    0.60  0.54536
## Year2008            0.0156     0.0474    0.33  0.74131
## Year2009           -0.0101     0.0473   -0.21  0.83029
## Year2010            0.0182     0.0477    0.38  0.70209
## Year2011            0.0236     0.0490    0.48  0.63000
## Year2012            0.0336     0.0496    0.68  0.49759
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.513
## Multiple R-squared:  0.0182, Adjusted R-squared:  0.0151
## Convergence in 15 IRWLS iterations
##
## Robustness weights:
##  473 weights are ~= 1. The remaining 5032 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0183  0.8620  0.9500  0.8980  0.9860  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.82e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Sample size for the above analysis:  5505"
## [1] ""
## [1] ""
## [1] "####################################"
## [1] "Analysis of AJSC 3103"
## [1] "####################################"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 2099 1715 1829 1688 1572 1583 1504 1484 1556 1471 1616 1664 1601 2113 2046
## 2011 2012
## 1922 1899
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 1196  964  979  951  931  875  827  853  838  828  953  956  886 1173 1194
## 2011 2012
```

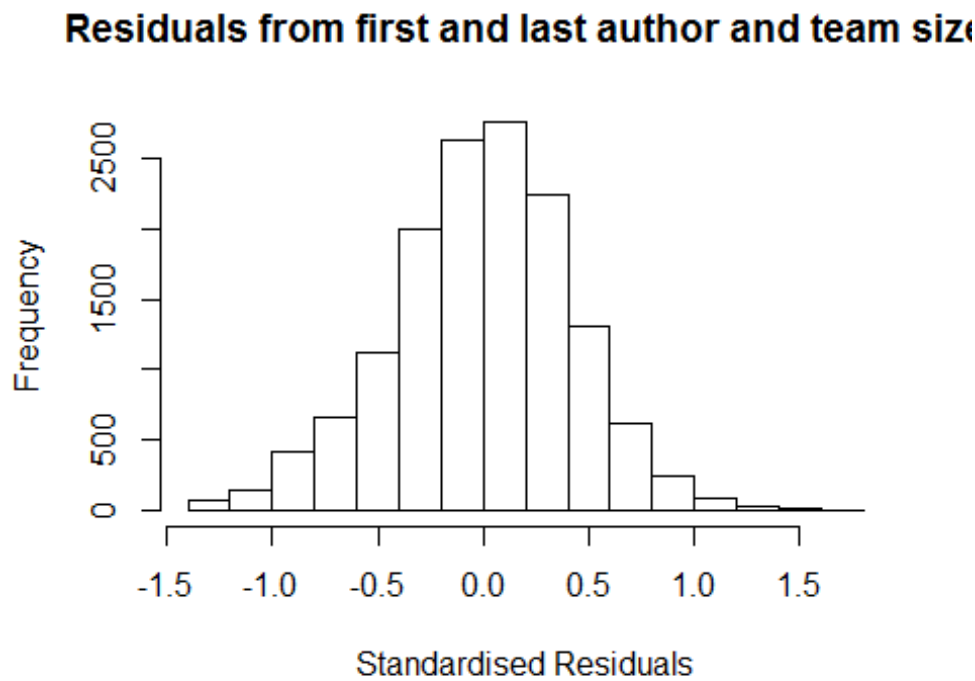
```
## 1135 1125
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 1032  829  855  829  802  759  704  713  713  710  812  837  767 1013 1022
## 2011 2012
##  996  961
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
##  Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 120, df = 16, p-value <2e-16
```




```
##
##  Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 39, df = 1, p-value = 4e-10
```

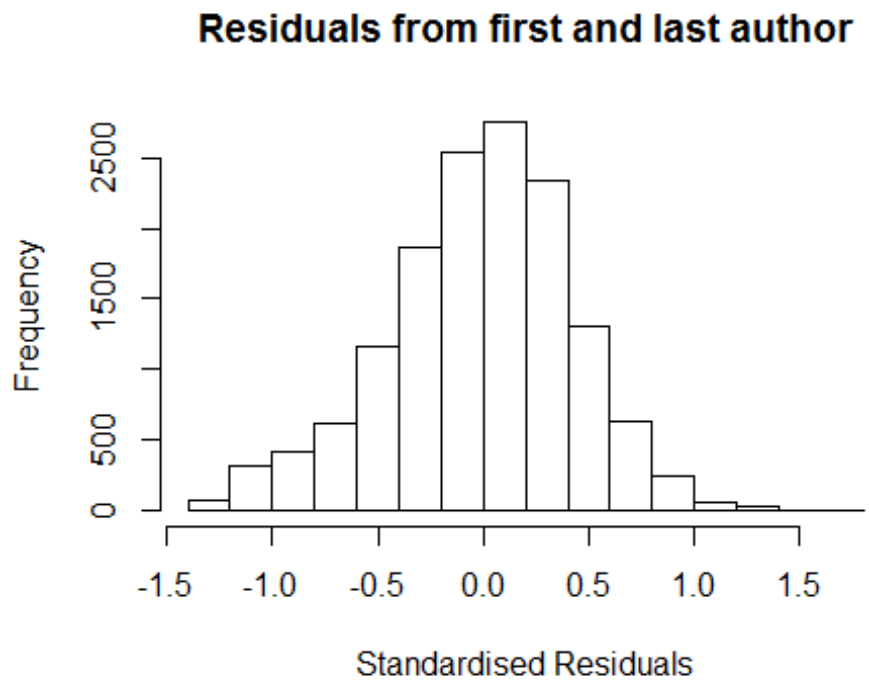



```
## [1] "Female first author team size 2018 geometric mean: 3.51125287518543"
## [1] "Male first author team size 2018 geometric mean: 2.80950460802796"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 99000, p-value = 8e-06
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 3.43491861099249"
## [1] "Male last author team size 2018 geometric mean: 2.84320573297875"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 88000, p-value = 4e-04
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.033  1           1.016
## LastAuthorFemale  1.027  1           1.013
## UniqueAuthors     1.054  4           1.007
## Year              1.043 16           1.001
```



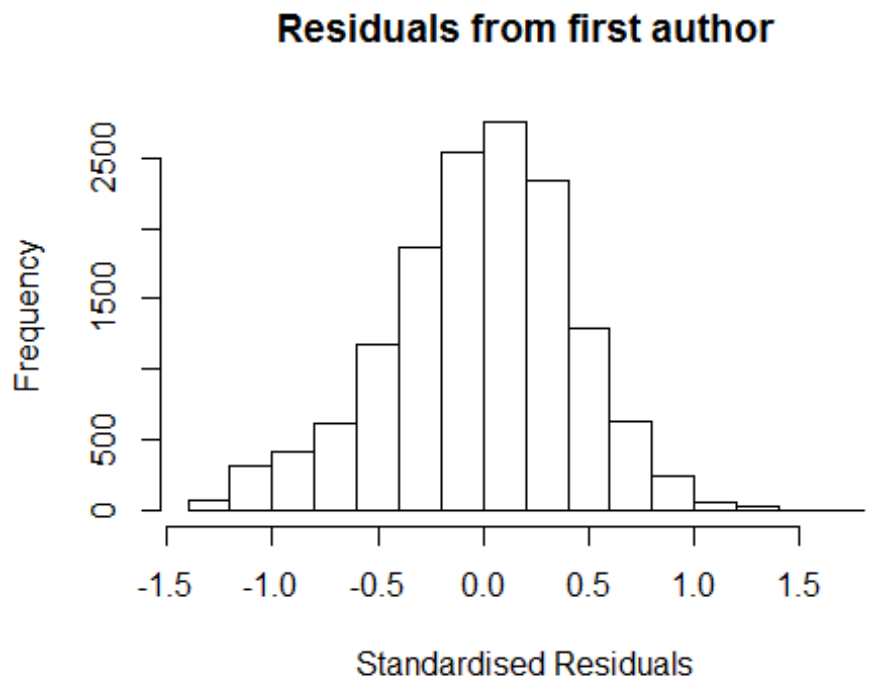

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
UniqueAuthors +
##     Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##     k.max = 1000))
##  \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.36690 -0.27419  0.00684  0.27318  1.77861
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.07217    0.01638   65.47  < 2e-16 ***
## FirstAuthorFemale1 -0.00937    0.00914   -1.03  0.30513
## LastAuthorFemale1  -0.01369    0.01043   -1.31  0.18924
## UniqueAuthors2      0.21550    0.01155   18.66  < 2e-16 ***
## UniqueAuthors3      0.25293    0.01215   20.82  < 2e-16 ***
## UniqueAuthors4      0.25049    0.01366   18.34  < 2e-16 ***
## UniqueAuthors5      0.29473    0.01261   23.37  < 2e-16 ***
## Year1997           -0.05746    0.02208   -2.60  0.00928 **
## Year1998           -0.08148    0.02144   -3.80  0.00015 ***
## Year1999           -0.04090    0.02176   -1.88  0.06017 .
```

```
## Year2000           -0.07148    0.02078   -3.44  0.00058 ***
## Year2001           -0.07690    0.02105   -3.65  0.00026 ***
## Year2002           -0.09224    0.02134   -4.32  1.6e-05 ***
## Year2003           -0.11435    0.02109   -5.42  6.0e-08 ***
## Year2004           -0.13528    0.02034   -6.65  3.1e-11 ***
## Year2005           -0.13363    0.02208   -6.05  1.5e-09 ***
## Year2006           -0.13905    0.01997   -6.96  3.5e-12 ***
## Year2007           -0.14060    0.02118   -6.64  3.3e-11 ***
## Year2008           -0.11075    0.02055   -5.39  7.2e-08 ***
## Year2009           -0.11445    0.01949   -5.87  4.4e-09 ***
## Year2010           -0.11310    0.02006   -5.64  1.8e-08 ***
## Year2011           -0.11129    0.02006   -5.55  3.0e-08 ***
## Year2012           -0.12639    0.01988   -6.36  2.1e-10 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.407
## Multiple R-squared:  0.0644, Adjusted R-squared:  0.063
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
##  1170 weights are ~= 1. The remaining 13184 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0171  0.8670  0.9510  0.8990  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          6.97e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.018  1           1.009
## LastAuthorFemale  1.021  1           1.010
## Year              1.012 16           1.000
```



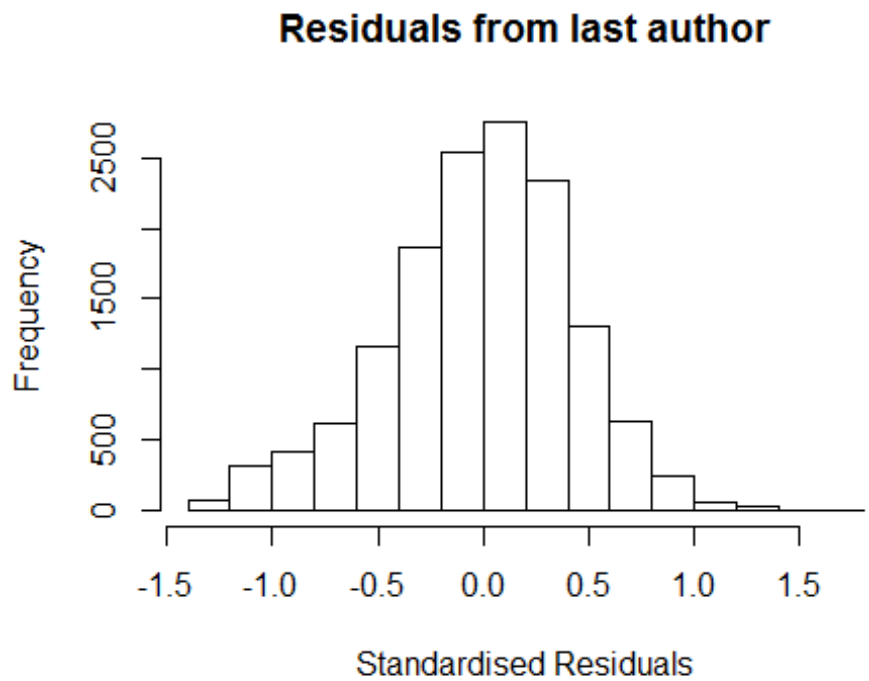

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.2478 -0.2762  0.0134  0.2766  1.7946
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)          1.22965    0.01448   84.89  < 2e-16 ***
## FirstAuthorFemale1   0.01927    0.00918    2.10  0.03593 *
## LastAuthorFemale1   -0.00111    0.01056   -0.10  0.91650
## Year1997            -0.03123    0.02233   -1.40  0.16185
## Year1998            -0.06822    0.02177   -3.13  0.00173 **
## Year1999            -0.02642    0.02218   -1.19  0.23364
## Year2000            -0.04286    0.02076   -2.06  0.03899 *
## Year2001            -0.04467    0.02114   -2.11  0.03463 *
## Year2002            -0.05638    0.02120   -2.66  0.00784 **
## Year2003            -0.07918    0.02122   -3.73  0.00019 ***
## Year2004            -0.09805    0.02034   -4.82  1.5e-06 ***
## Year2005            -0.09979    0.02215   -4.50  6.7e-06 ***
```

```
## Year2006           -0.10447    0.02026   -5.16  2.5e-07 ***
## Year2007           -0.09884    0.02150   -4.60  4.3e-06 ***
## Year2008           -0.07354    0.02110   -3.49  0.00049 ***
## Year2009           -0.08209    0.01989   -4.13  3.7e-05 ***
## Year2010           -0.06417    0.02046   -3.14  0.00171 **
## Year2011           -0.06929    0.02023   -3.43  0.00061 ***
## Year2012           -0.08648    0.01993   -4.34  1.4e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.413
## Multiple R-squared:  0.00473,    Adjusted R-squared:  0.00348
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  1208 weights are ~= 1. The remaining 13146 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0192  0.8650  0.9510  0.8950  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          6.97e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.005  1           1.002
## Year              1.005 16           1.000
```




```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.2486 -0.2763  0.0134  0.2765  1.7948
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)          1.22955    0.01444   85.13  < 2e-16 ***
## FirstAuthorFemale1  0.01907    0.00916    2.08  0.03732 *
## Year1997           -0.03126    0.02232   -1.40  0.16149
## Year1998           -0.06822    0.02177   -3.13  0.00173 **
## Year1999           -0.02644    0.02218   -1.19  0.23323
## Year2000           -0.04286    0.02076   -2.06  0.03899 *
## Year2001           -0.04466    0.02114   -2.11  0.03467 *
## Year2002           -0.05638    0.02120   -2.66  0.00785 **
## Year2003           -0.07920    0.02122   -3.73  0.00019 ***
## Year2004           -0.09808    0.02035   -4.82  1.4e-06 ***
## Year2005           -0.09981    0.02216   -4.50  6.7e-06 ***
## Year2006           -0.10449    0.02026   -5.16  2.5e-07 ***
```

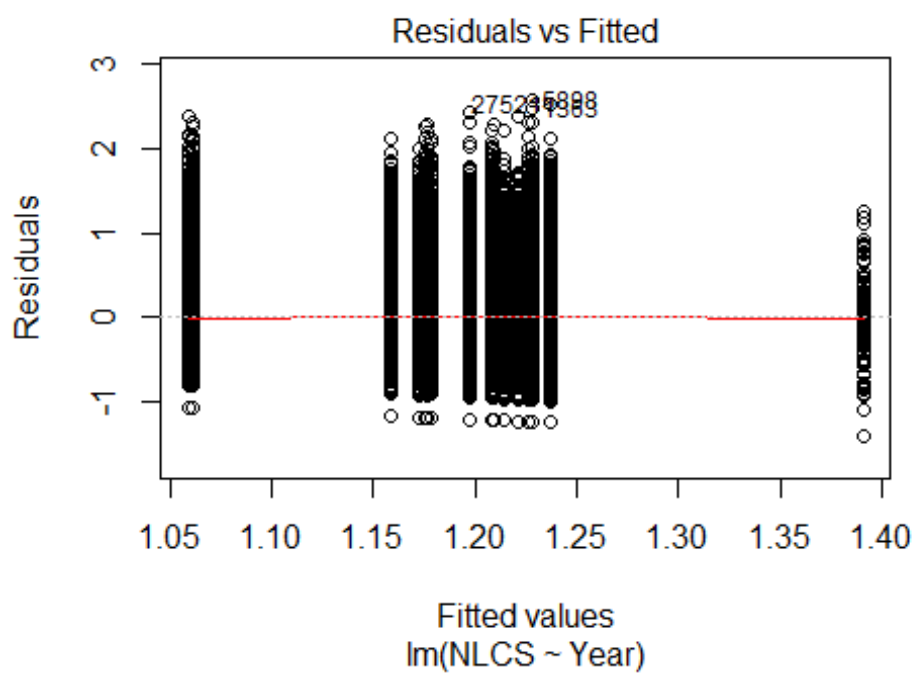
```
## Year2007           -0.09886    0.02150   -4.60  4.3e-06 ***
## Year2008           -0.07356    0.02110   -3.49  0.00049 ***
## Year2009           -0.08212    0.01988   -4.13  3.6e-05 ***
## Year2010           -0.06423    0.02044   -3.14  0.00168 **
## Year2011           -0.06932    0.02023   -3.43  0.00061 ***
## Year2012           -0.08652    0.01993   -4.34  1.4e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.413
## Multiple R-squared:  0.00473,    Adjusted R-squared:  0.00355
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  1206 weights are ~= 1. The remaining 13148 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0192  0.8650  0.9510  0.8950  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          6.97e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##                   GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.008  1           1.004
## Year             1.008 16           1.000
```



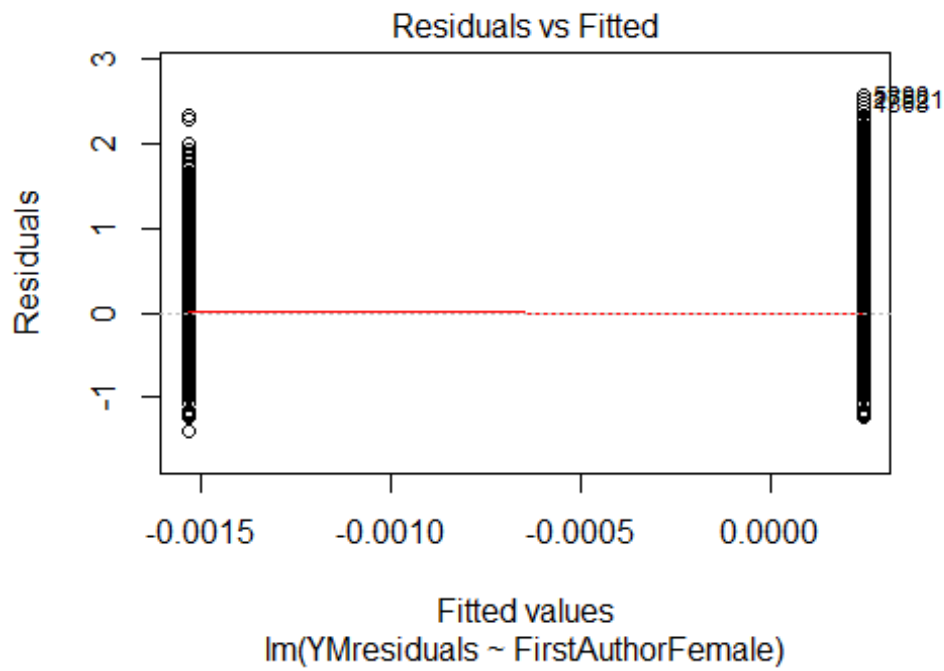

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.2349 -0.2777  0.0141  0.2784  1.7913
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.23171    0.01444   85.28  < 2e-16 ***
## LastAuthorFemale1   0.00315    0.01050    0.30  0.76434
## Year1997           -0.03134    0.02234   -1.40  0.16063
## Year1998           -0.06780    0.02180   -3.11  0.00188 **
## Year1999           -0.02560    0.02220   -1.15  0.24890
## Year2000           -0.04227    0.02077   -2.04  0.04185 *
## Year2001           -0.04408    0.02115   -2.08  0.03714 *
## Year2002           -0.05531    0.02124   -2.60  0.00921 **
## Year2003           -0.07868    0.02122   -3.71  0.00021 ***
## Year2004           -0.09701    0.02036   -4.76  1.9e-06 ***
## Year2005           -0.09922    0.02215   -4.48  7.6e-06 ***
## Year2006           -0.10305    0.02028   -5.08  3.8e-07 ***
```

```
## Year2007          -0.09794    0.02152   -4.55  5.4e-06 ***
## Year2008          -0.07252    0.02110   -3.44  0.00059 ***
## Year2009          -0.08115    0.01990   -4.08  4.6e-05 ***
## Year2010          -0.06255    0.02045   -3.06  0.00223 **
## Year2011          -0.06802    0.02023   -3.36  0.00077 ***
## Year2012          -0.08539    0.01993   -4.28  1.8e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.413
## Multiple R-squared:  0.00447,    Adjusted R-squared:  0.00329
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  1217 weights are ~= 1. The remaining 13137 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0201  0.8630  0.9510  0.8950  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          6.97e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Sample size for the above analysis:  14354"
## [1] ""
## [1] ""
## [1] "####################################"
## [1] "Analysis of AJSC 3104"
## [1] "####################################"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
##  1996  1997  1998  1999  2000  2001  2002  2003  2004  2005  2006  2007
##  9063  9133  8430   550  8841 10270  9548  7234  7046  7280  7808  7686
##  2008  2009  2010  2011  2012
##  7874  8249  8161  8456  8576
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 2931 3043 2810  147 3068 3485 3909 2804 2826 2975 3221 3304 3366 3762 3716
## 2011 2012
```

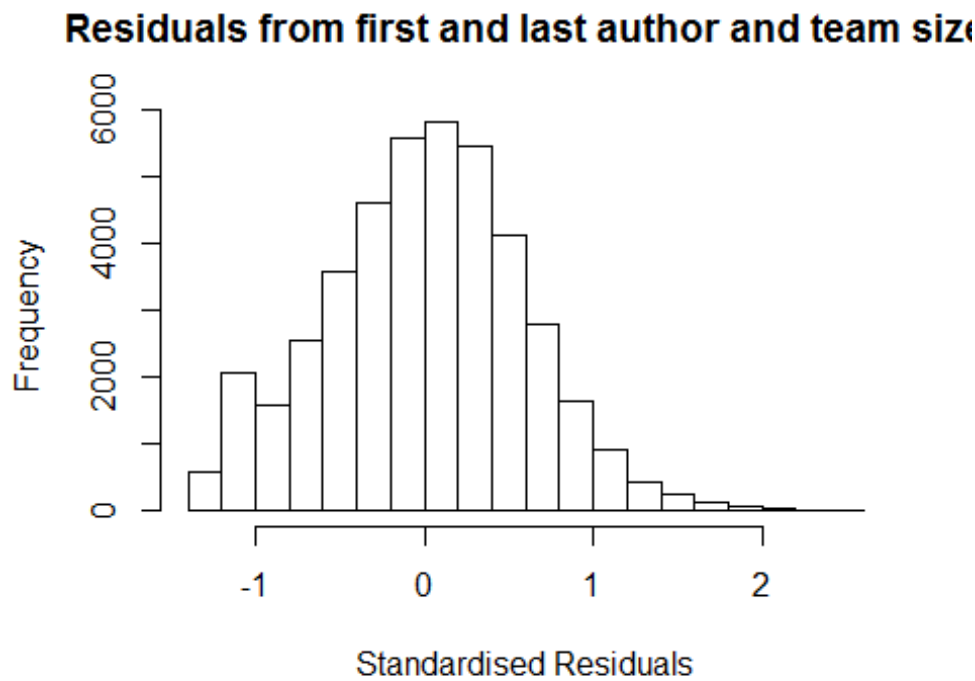
```
## 3926 4043
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 2365 2472 2244  120 2474 2742 3133 2174 2251 2306 2503 2554 2667 2964 2912
## 2011 2012
## 3041 3191
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
##  Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 610, df = 16, p-value <2e-16
```




```
##
##  Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 27, df = 1, p-value = 2e-07
```

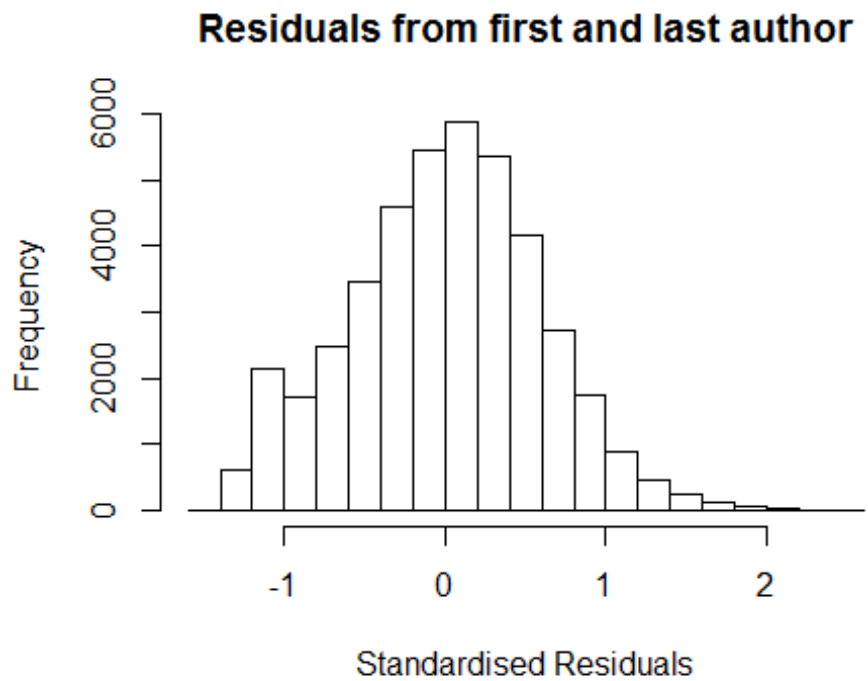



```
## [1] "Female first author team size 2018 geometric mean: 3.61567500530582"
## [1] "Male first author team size 2018 geometric mean: 3.12407500233381"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 540000, p-value = 2e-05
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 3.52309920107376"
## [1] "Male last author team size 2018 geometric mean: 3.1518387400107"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 450000, p-value = 0.002
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.039  1           1.019
## LastAuthorFemale  1.028  1           1.014
## UniqueAuthors     1.057  4           1.007
## Year              1.047 16           1.001
```



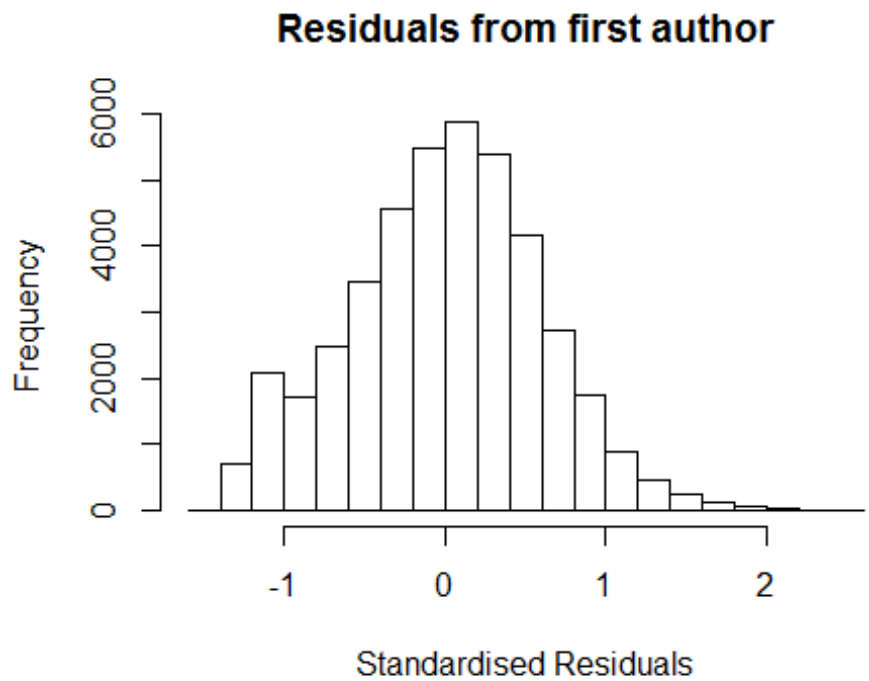

```
## [1] "List of  1 outliers with residuals above  2.5"
##        ScopusId  NLCS Year OneField Fields residuals
## 5898 2842565972 3.795 1996     2504      2     2.543
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
UniqueAuthors +
##     Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##     k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.3830 -0.3892  0.0186  0.3924  2.5428
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.09723    0.01508   72.74  < 2e-16 ***
## FirstAuthorFemale1 -0.01552    0.00844   -1.84  0.06606 .
## LastAuthorFemale1   0.00754    0.00964    0.78  0.43384
## UniqueAuthors2      0.15493    0.00964   16.06  < 2e-16 ***
## UniqueAuthors3      0.16221    0.01016   15.96  < 2e-16 ***
## UniqueAuthors4      0.18365    0.01130   16.25  < 2e-16 ***
## UniqueAuthors5      0.25715    0.01068   24.08  < 2e-16 ***
## Year1997            0.00343    0.01834    0.19  0.85178
## Year1998            0.00954    0.01899    0.50  0.61555
## Year1999            0.14636    0.04532    3.23  0.00124 **
```

```
## Year2000           -0.08016    0.01886   -4.25  2.1e-05 ***
## Year2001           -0.22770    0.02042  -11.15  < 2e-16 ***
## Year2002           -0.21330    0.01876  -11.37  < 2e-16 ***
## Year2003           -0.07180    0.01944   -3.69  0.00022 ***
## Year2004           -0.08021    0.01853   -4.33  1.5e-05 ***
## Year2005           -0.04778    0.01748   -2.73  0.00628 **
## Year2006           -0.05225    0.01756   -2.98  0.00293 **
## Year2007           -0.05884    0.01772   -3.32  0.00090 ***
## Year2008           -0.06818    0.01740   -3.92  9.0e-05 ***
## Year2009           -0.06194    0.01704   -3.63  0.00028 ***
## Year2010           -0.06159    0.01694   -3.64  0.00028 ***
## Year2011           -0.10009    0.01686   -5.94  3.0e-09 ***
## Year2012           -0.09093    0.01666   -5.46  4.9e-08 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.58
## Multiple R-squared:  0.0259, Adjusted R-squared:  0.0254
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  3577 weights are ~= 1. The remaining 38536 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0158  0.8660  0.9500  0.9050  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          2.37e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.025  1           1.013
## LastAuthorFemale  1.024  1           1.012
## Year              1.008 16           1.000
```



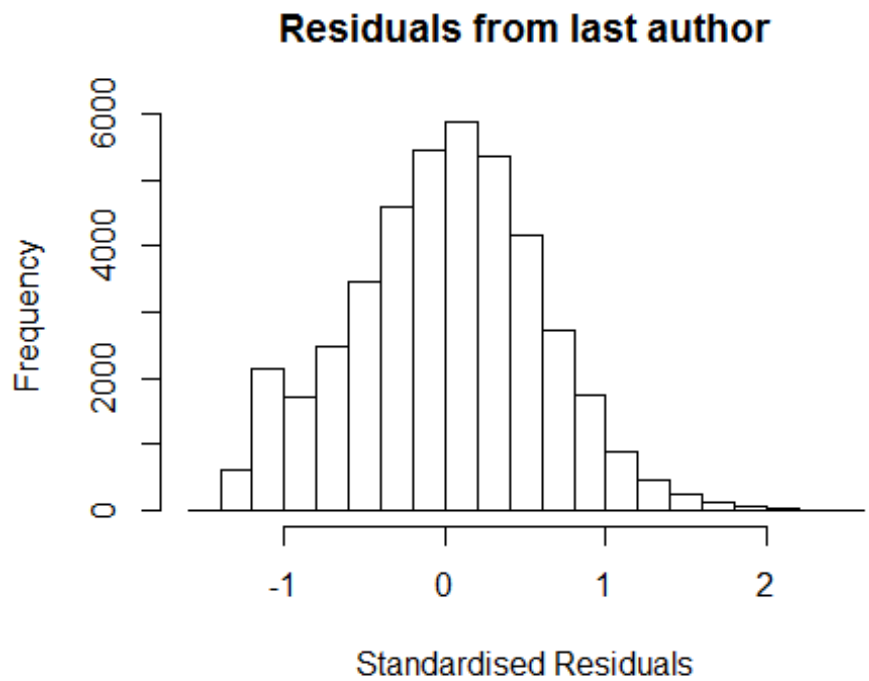

```
## [1] "List of  1 outliers with residuals above  2.5"
##        ScopusId  NLCS Year OneField Fields residuals
## 5898 2842565972 3.795 1996     2504      2     2.575
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.4135 -0.3933  0.0192  0.3940  2.5750
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.220038   0.013536   90.13  < 2e-16 ***
## FirstAuthorFemale1  0.007707   0.008439    0.91  0.36116
## LastAuthorFemale1   0.017774   0.009698    1.83  0.06686 .
## Year1997            0.000979   0.018448    0.05  0.95768
## Year1998            0.014075   0.019069    0.74  0.46047
## Year1999            0.185744   0.046269    4.01    6e-05 ***
## Year2000           -0.064402   0.018809   -3.42  0.00062 ***
## Year2001           -0.194409   0.020268   -9.59  < 2e-16 ***
## Year2002           -0.180588   0.018669   -9.67  < 2e-16 ***
## Year2003           -0.049809   0.019496   -2.55  0.01063 *
## Year2004           -0.053544   0.018666   -2.87  0.00413 **
## Year2005           -0.018975   0.017588   -1.08  0.28066
```

```
## Year2006           -0.020182   0.017640   -1.14  0.25259
## Year2007           -0.025248   0.017750   -1.42  0.15491
## Year2008           -0.037035   0.017464   -2.12  0.03395 *
## Year2009           -0.029055   0.017124   -1.70  0.08976 .
## Year2010           -0.023812   0.016993   -1.40  0.16115
## Year2011           -0.060761   0.016877   -3.60  0.00032 ***
## Year2012           -0.051557   0.016701   -3.09  0.00202 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.585
## Multiple R-squared:  0.00985,    Adjusted R-squared:  0.00943
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  3545 weights are ~= 1. The remaining 38568 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0134  0.8680  0.9500  0.9050  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          2.37e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.005  1           1.002
## Year              1.005 16           1.000
```




```
## [1] "List of  1 outliers with residuals above  2.5"
##        ScopusId  NLCS Year OneField Fields residuals
## 5898 2842565972 3.795 1996     2504      2     2.575
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.4174 -0.3938  0.0204  0.3946  2.5740
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)          1.22098    0.01352   90.29  < 2e-16 ***
## FirstAuthorFemale1   0.01004    0.00835    1.20  0.22963
## Year1997             0.00101    0.01845    0.05  0.95614
## Year1998             0.01427    0.01907    0.75  0.45423
## Year1999             0.18634    0.04622    4.03  5.6e-05 ***
## Year2000            -0.06413    0.01881   -3.41  0.00065 ***
## Year2001            -0.19396    0.02027   -9.57  < 2e-16 ***
## Year2002            -0.18011    0.01867   -9.65  < 2e-16 ***
## Year2003            -0.04935    0.01950   -2.53  0.01138 *
## Year2004            -0.05301    0.01866   -2.84  0.00451 **
## Year2005            -0.01841    0.01758   -1.05  0.29504
## Year2006            -0.01976    0.01764   -1.12  0.26275
```

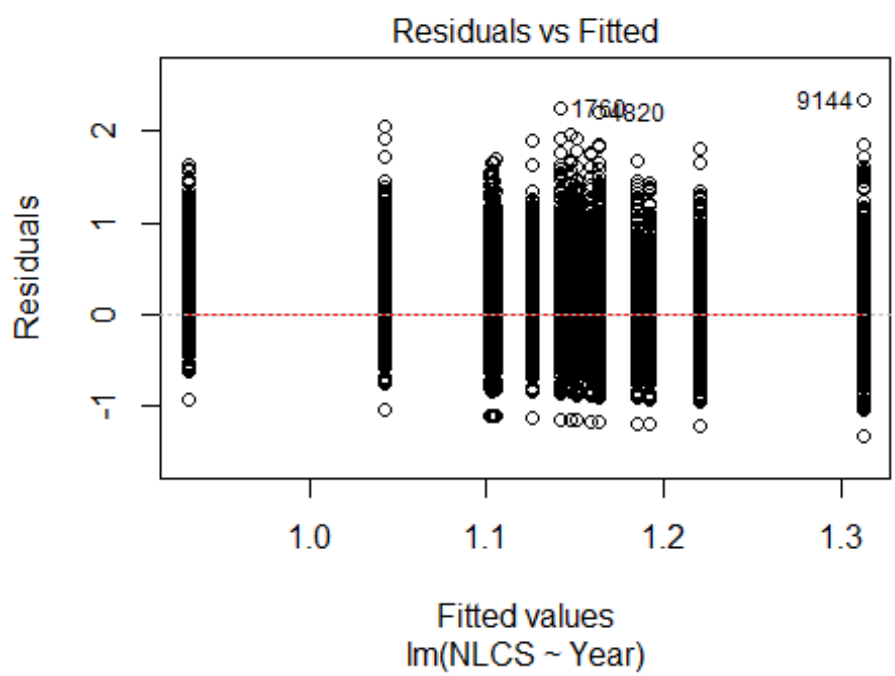
```
## Year2007           -0.02461    0.01775   -1.39  0.16557
## Year2008           -0.03643    0.01746   -2.09  0.03696 *
## Year2009           -0.02843    0.01712   -1.66  0.09683 .
## Year2010           -0.02301    0.01699   -1.35  0.17563
## Year2011           -0.06004    0.01688   -3.56  0.00037 ***
## Year2012           -0.05055    0.01669   -3.03  0.00246 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.585
## Multiple R-squared:  0.00977,    Adjusted R-squared:  0.00937
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  3553 weights are ~= 1. The remaining 38560 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0136  0.8670  0.9500  0.9050  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          2.37e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##                   GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.004  1           1.002
## Year             1.004 16           1.000
```



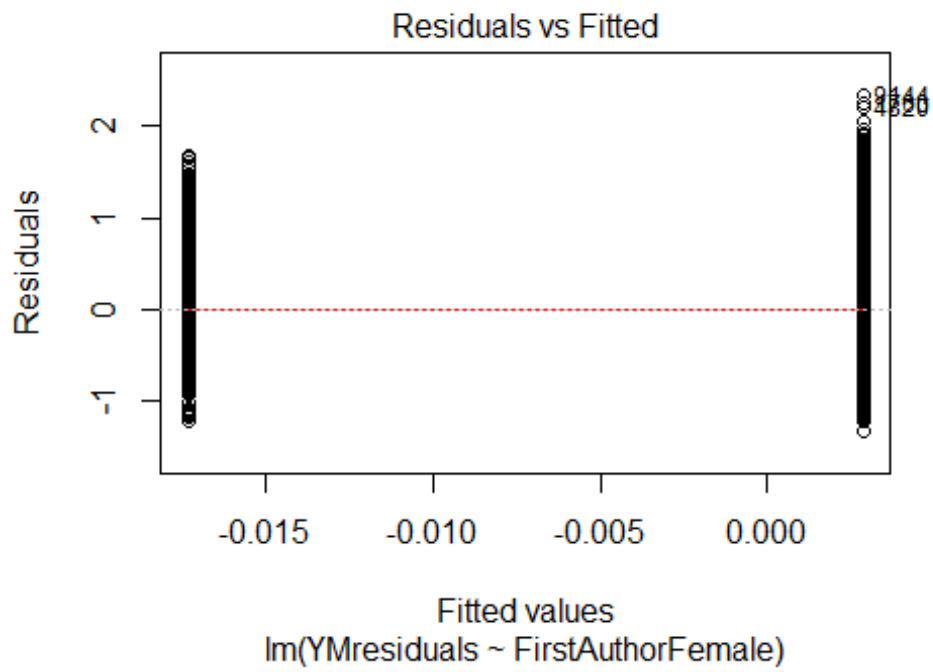

```
## [1] "List of  1 outliers with residuals above  2.5"
##        ScopusId  NLCS Year OneField Fields residuals
## 5898 2842565972 3.795 1996     2504      2     2.575
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.4071 -0.3934  0.0196  0.3939  2.5742
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.220818   0.013505   90.40  < 2e-16 ***
## LastAuthorFemale1   0.019124   0.009601    1.99  0.04639 *
## Year1997            0.000958   0.018449    0.05  0.95857
## Year1998            0.014106   0.019071    0.74  0.45951
## Year1999            0.186309   0.046253    4.03  5.6e-05 ***
## Year2000           -0.064444   0.018808   -3.43  0.00061 ***
## Year2001           -0.194341   0.020268   -9.59  < 2e-16 ***
## Year2002           -0.180478   0.018668   -9.67  < 2e-16 ***
## Year2003           -0.049566   0.019496   -2.54  0.01101 *
## Year2004           -0.053505   0.018667   -2.87  0.00416 **
## Year2005           -0.018924   0.017589   -1.08  0.28196
## Year2006           -0.019940   0.017638   -1.13  0.25827
```

```
## Year2007          -0.025097   0.017751   -1.41  0.15742
## Year2008          -0.036761   0.017463   -2.11  0.03529 *
## Year2009          -0.028787   0.017122   -1.68  0.09272 .
## Year2010          -0.023471   0.016990   -1.38  0.16715
## Year2011          -0.060445   0.016873   -3.58  0.00034 ***
## Year2012          -0.051186   0.016694   -3.07  0.00217 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.585
## Multiple R-squared:  0.00983,    Adjusted R-squared:  0.00943
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  3556 weights are ~= 1. The remaining 38557 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0136  0.8680  0.9500  0.9050  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          2.37e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Sample size for the above analysis:  42113"
## [1] ""
## [1] ""
## [1] "####################################"
## [1] "Analysis of AJSC 3105"
## [1] "####################################"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 1410 1443 1427 1480 1401 1417 1218  955  931 1076 1383 1635 1226 1711 1122
## 2011 2012
## 1379 1717
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  589  604  671  515  670  515  618  432  445  551  673  746  609  715  575
## 2011 2012
```

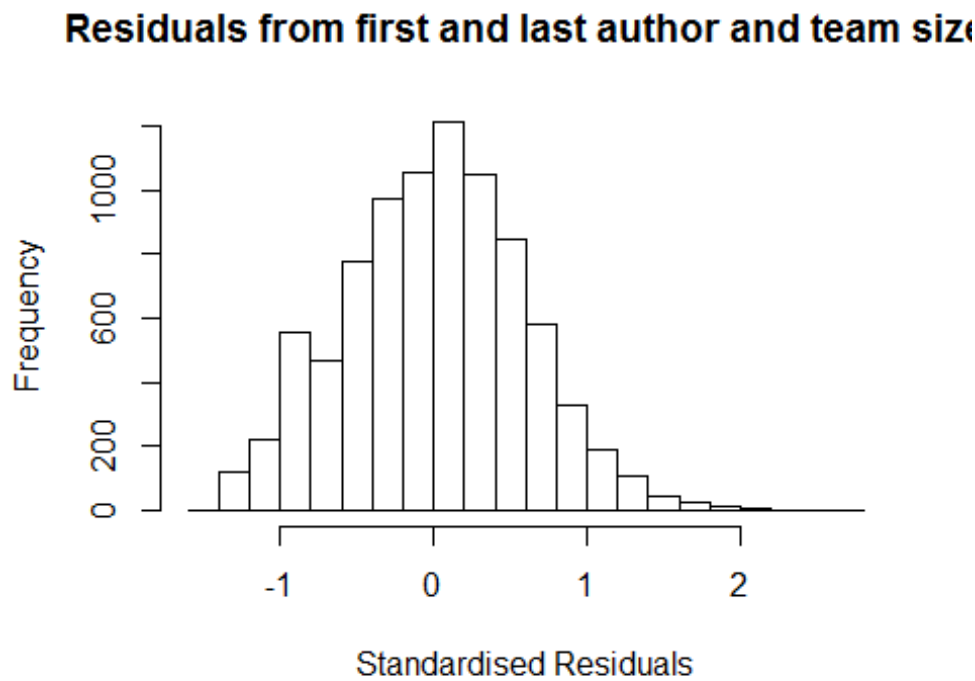
```
##  688  733
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  503  520  578  453  576  423  503  338  351  450  551  621  493  581  478
## 2011 2012
##  565  588
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
##  Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 120, df = 16, p-value <2e-16
```




```
##
##  Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 6.8, df = 1, p-value = 0.009
```

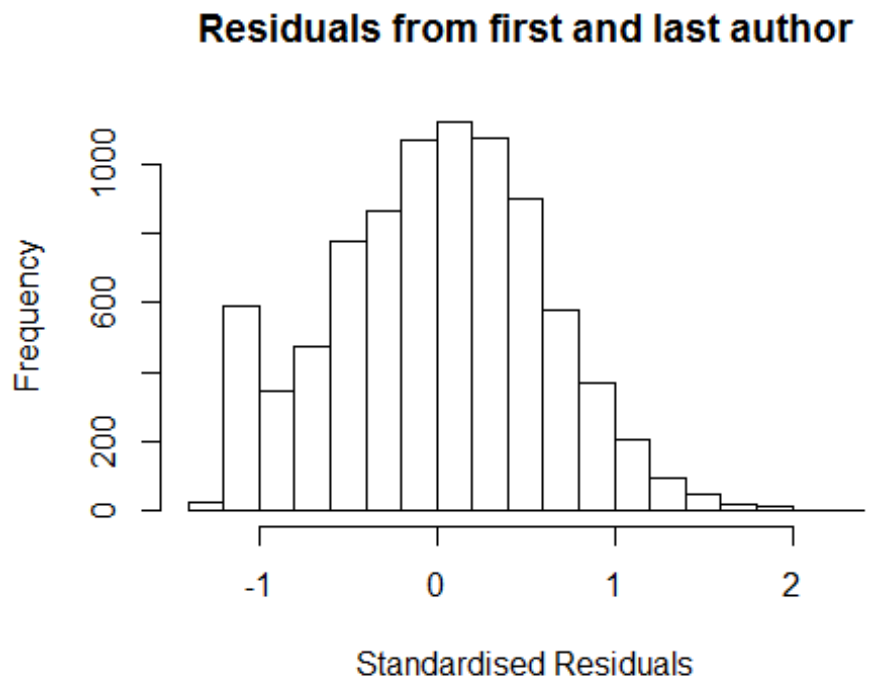



```
## [1] "Female first author team size 2018 geometric mean: 4.0774604662463"
## [1] "Male first author team size 2018 geometric mean: 3.40382493301218"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 53000, p-value = 5e-04
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 3.7795599263403"
## [1] "Male last author team size 2018 geometric mean: 3.48870380472555"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 37000, p-value = 0.2
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.037  1           1.018
## LastAuthorFemale  1.034  1           1.017
## UniqueAuthors     1.080  4           1.010
## Year              1.082 16           1.002
```



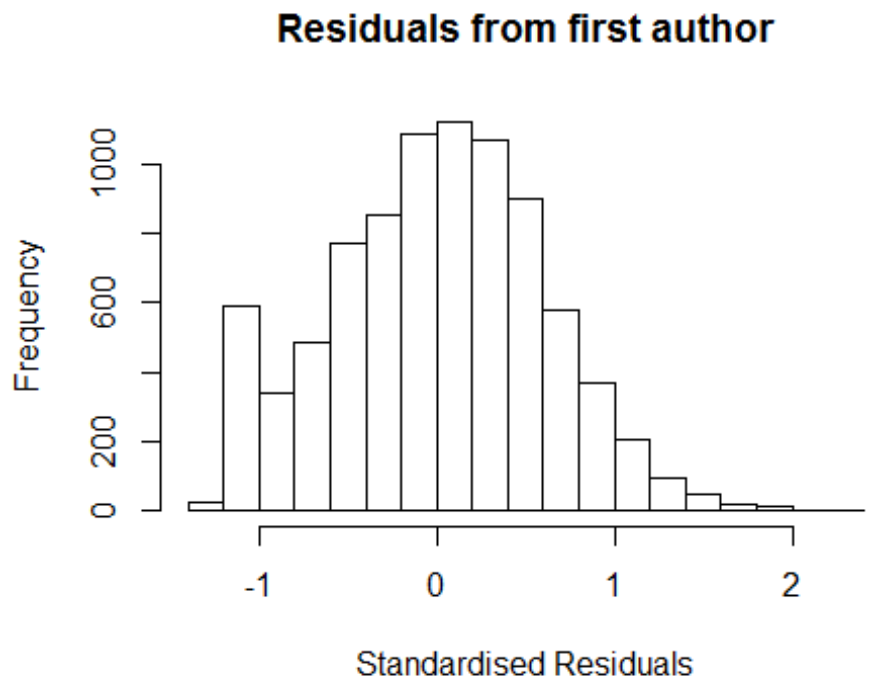

```
## [1] "List of  1 outliers with residuals above  2.5"
##        ScopusId  NLCS Year OneField Fields residuals
## 9144 0035288601 3.646 2001     3101      4     2.617
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
UniqueAuthors +
##     Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##     k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.4351 -0.4004  0.0202  0.3996  2.6166
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)          0.9158     0.0351   26.07  < 2e-16 ***
## FirstAuthorFemale1  -0.0379     0.0186   -2.04  0.04098 *
## LastAuthorFemale1   -0.0111     0.0227   -0.49  0.62306
## UniqueAuthors2       0.2736     0.0218   12.54  < 2e-16 ***
## UniqueAuthors3       0.3548     0.0223   15.93  < 2e-16 ***
## UniqueAuthors4       0.3756     0.0241   15.59  < 2e-16 ***
## UniqueAuthors5       0.4057     0.0228   17.81  < 2e-16 ***
## Year1997            -0.0143     0.0428   -0.34  0.73749
## Year1998             0.0158     0.0408    0.39  0.69811
## Year1999            -0.0698     0.0407   -1.72  0.08608 .
```

```
## Year2000            -0.0530     0.0427   -1.24  0.21490
## Year2001             0.1136     0.0449    2.53  0.01147 *
## Year2002            -0.0616     0.0430   -1.43  0.15247
## Year2003            -0.0240     0.0460   -0.52  0.60158
## Year2004            -0.0394     0.0446   -0.88  0.37652
## Year2005            -0.0125     0.0414   -0.30  0.76356
## Year2006            -0.0459     0.0395   -1.16  0.24525
## Year2007            -0.0781     0.0381   -2.05  0.04017 *
## Year2008            -0.0835     0.0405   -2.06  0.03919 *
## Year2009            -0.1589     0.0409   -3.89  0.00010 ***
## Year2010            -0.1014     0.0401   -2.53  0.01154 *
## Year2011            -0.1308     0.0382   -3.42  0.00062 ***
## Year2012            -0.2998     0.0423   -7.10  1.4e-12 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.593
## Multiple R-squared:  0.0679, Adjusted R-squared:  0.0655
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  687 weights are ~= 1. The remaining 7885 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0124  0.8680  0.9510  0.9090  0.9860  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.17e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.020  1           1.010
## LastAuthorFemale  1.019  1           1.009
## Year              1.010 16           1.000
```



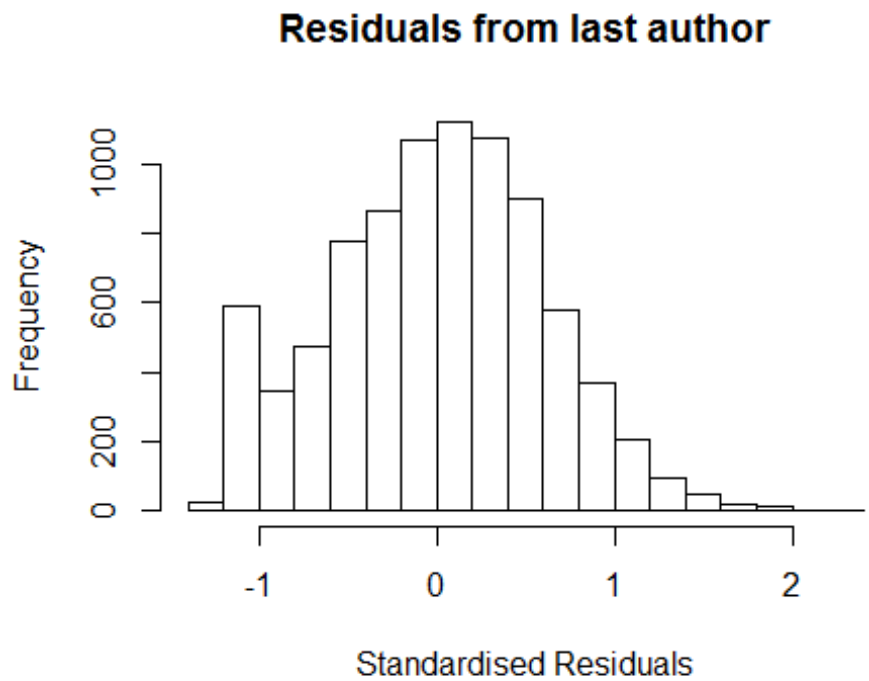

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.3009 -0.4177  0.0297  0.4146  2.3451
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)          1.15006    0.03209   35.84  < 2e-16 ***
## FirstAuthorFemale1 -0.00438    0.01893   -0.23  0.81724
## LastAuthorFemale1  -0.00832    0.02317   -0.36  0.71971
## Year1997           -0.00357    0.04395   -0.08  0.93525
## Year1998            0.01598    0.04228    0.38  0.70557
## Year1999           -0.05238    0.04153   -1.26  0.20726
## Year2000           -0.05695    0.04367   -1.30  0.19223
## Year2001            0.15083    0.04548    3.32  0.00092 ***
## Year2002           -0.02877    0.04444   -0.65  0.51742
## Year2003            0.02698    0.04678    0.58  0.56416
## Year2004            0.01940    0.04594    0.42  0.67289
## Year2005            0.04435    0.04233    1.05  0.29477
```

```
## Year2006           -0.00747    0.04064   -0.18  0.85415
## Year2007           -0.01938    0.03890   -0.50  0.61824
## Year2008           -0.01500    0.04154   -0.36  0.71798
## Year2009           -0.11130    0.04212   -2.64  0.00825 **
## Year2010           -0.02675    0.04094   -0.65  0.51344
## Year2011           -0.05025    0.03836   -1.31  0.19026
## Year2012           -0.22462    0.04297   -5.23  1.8e-07 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.608
## Multiple R-squared:  0.0147, Adjusted R-squared:  0.0126
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
##  695 weights are ~= 1. The remaining 7877 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.104   0.865   0.950   0.909   0.985   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.17e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.005  1           1.003
## Year              1.005 16           1.000
```




```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.3003 -0.4191  0.0298  0.4151  2.3457
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.14946    0.03203   35.89  < 2e-16 ***
## FirstAuthorFemale1 -0.00566    0.01885   -0.30  0.76414
## Year1997           -0.00352    0.04394   -0.08  0.93624
## Year1998            0.01608    0.04228    0.38  0.70367
## Year1999           -0.05239    0.04154   -1.26  0.20723
## Year2000           -0.05696    0.04367   -1.30  0.19214
## Year2001            0.15089    0.04548    3.32  0.00091 ***
## Year2002           -0.02862    0.04443   -0.64  0.51951
## Year2003            0.02719    0.04678    0.58  0.56112
## Year2004            0.01928    0.04595    0.42  0.67477
## Year2005            0.04427    0.04234    1.05  0.29580
## Year2006           -0.00746    0.04064   -0.18  0.85444
```

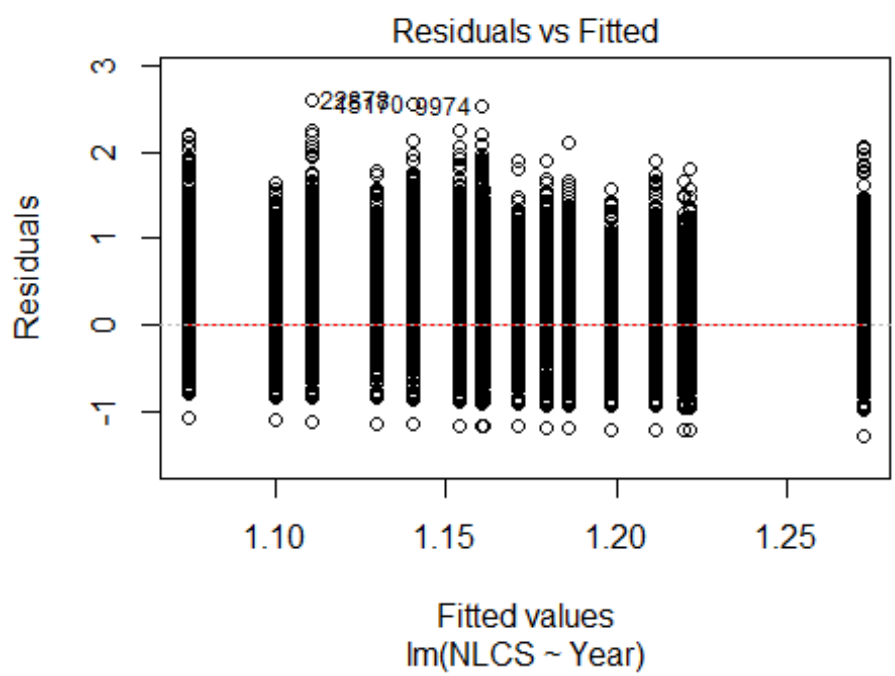
```
## Year2007           -0.01954    0.03889   -0.50  0.61539
## Year2008           -0.01512    0.04155   -0.36  0.71597
## Year2009           -0.11146    0.04213   -2.65  0.00817 **
## Year2010           -0.02689    0.04094   -0.66  0.51129
## Year2011           -0.05048    0.03836   -1.32  0.18818
## Year2012           -0.22484    0.04298   -5.23  1.7e-07 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.608
## Multiple R-squared:  0.0147, Adjusted R-squared:  0.0127
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
##  686 weights are ~= 1. The remaining 7886 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.104   0.865   0.950   0.910   0.985   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.17e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##                   GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.004  1           1.002
## Year             1.004 16           1.000
```



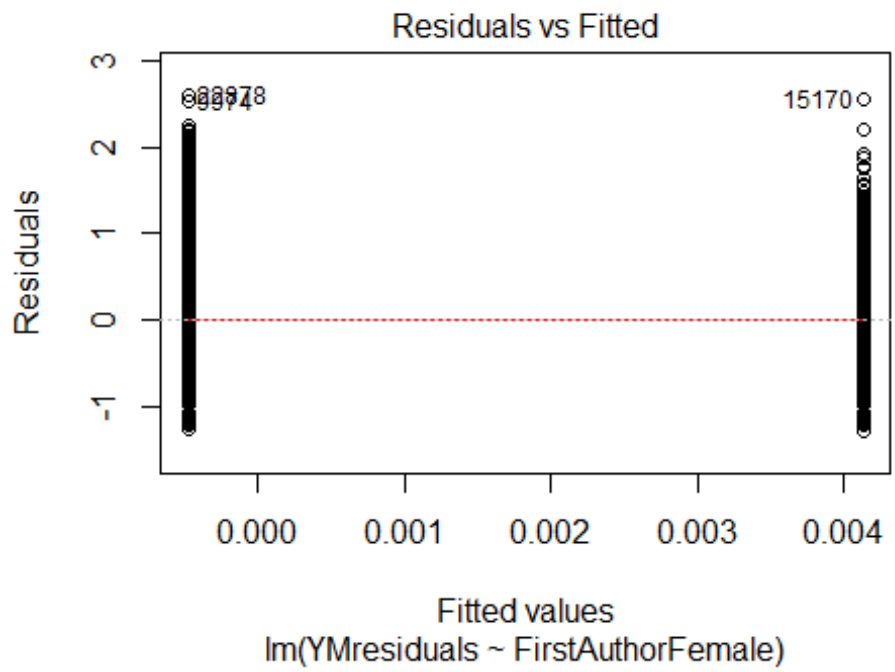

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.3004 -0.4180  0.0296  0.4150  2.3456
##
## Coefficients:
##                   Estimate Std. Error t value Pr(>|t|)
## (Intercept)        1.14960    0.03202   35.90  < 2e-16 ***
## LastAuthorFemale1 -0.00924    0.02304   -0.40  0.68839
## Year1997          -0.00345    0.04395   -0.08  0.93744
## Year1998           0.01592    0.04228    0.38  0.70661
## Year1999          -0.05249    0.04153   -1.26  0.20634
## Year2000          -0.05697    0.04367   -1.30  0.19211
## Year2001           0.15078    0.04548    3.32  0.00092 ***
## Year2002          -0.02895    0.04443   -0.65  0.51467
## Year2003           0.02678    0.04676    0.57  0.56691
## Year2004           0.01927    0.04593    0.42  0.67475
## Year2005           0.04428    0.04233    1.05  0.29563
## Year2006          -0.00752    0.04065   -0.18  0.85324
```

```
## Year2007          -0.01948    0.03890   -0.50  0.61649
## Year2008          -0.01519    0.04154   -0.37  0.71456
## Year2009          -0.11139    0.04213   -2.64  0.00821 **
## Year2010          -0.02681    0.04094   -0.65  0.51254
## Year2011          -0.05033    0.03837   -1.31  0.18960
## Year2012          -0.22472    0.04298   -5.23  1.8e-07 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.608
## Multiple R-squared:  0.0147, Adjusted R-squared:  0.0127
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
##  690 weights are ~= 1. The remaining 7882 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.104   0.865   0.950   0.909   0.985   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.17e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Sample size for the above analysis:  8572"
## [1] ""
## [1] ""
## [1] "####################################"
## [1] "Analysis of AJSC 3106"
## [1] "####################################"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 2379 2658 3036 3180 3284 2972 3043 2328 2309 2445 2263 2557 2306 2691 2229
## 2011 2012
## 2338 2276
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  894 1195 1342 1376 1589 1415 1497 1142 1281 1264 1205 1274 1261 1445 1220
## 2011 2012
```

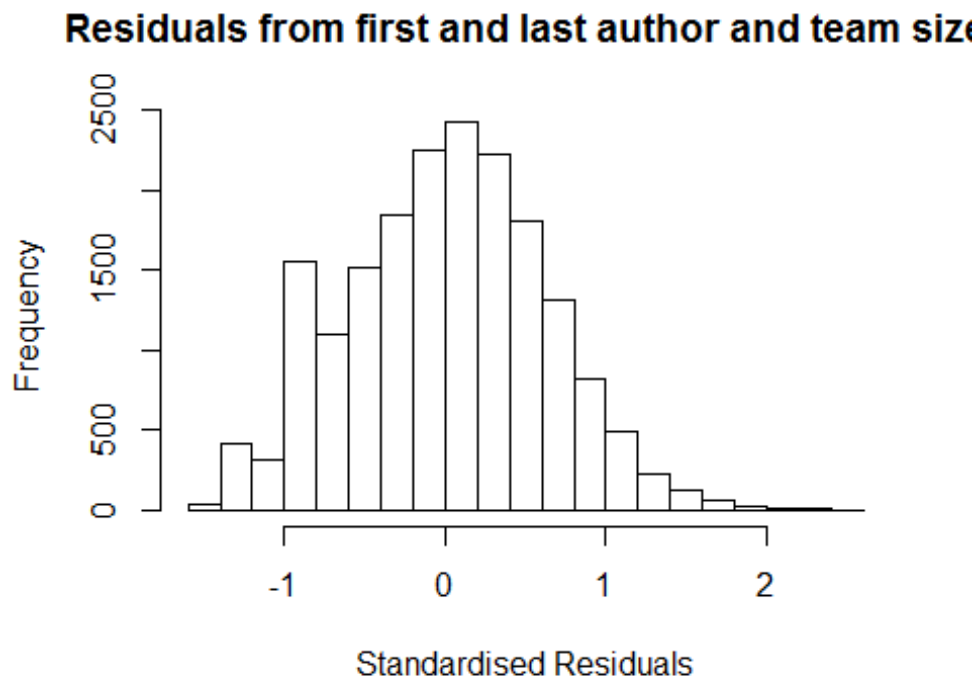
```
## 1286 1227
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  761 1000 1144 1172 1379 1234 1275  969 1094 1084 1024 1090 1070 1238 1018
## 2011 2012
## 1069 1007
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
##  Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 700, df = 16, p-value <2e-16
```




```
##
##  Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 7.9, df = 1, p-value = 0.005
```

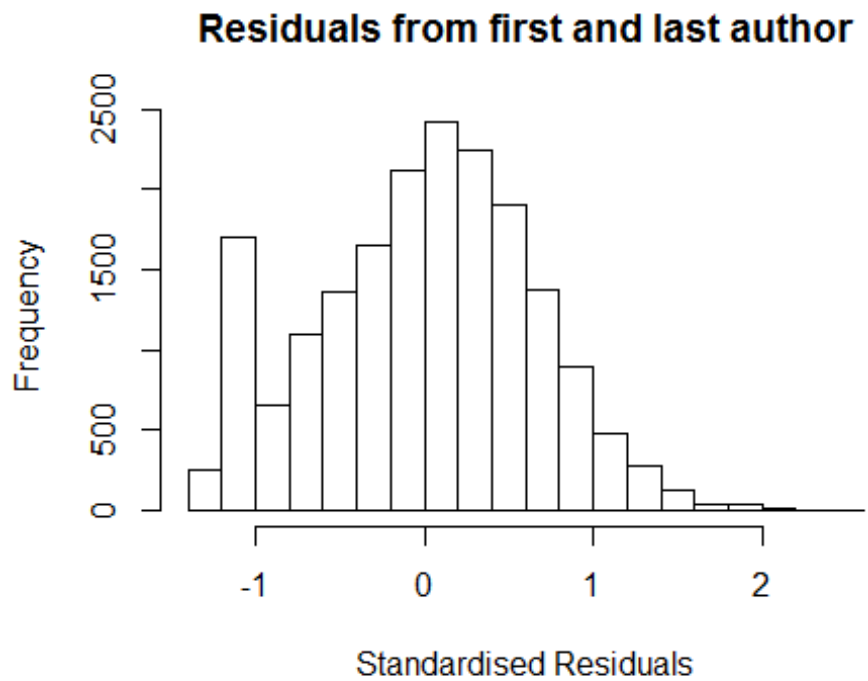



```
## [1] "Female first author team size 2018 geometric mean: 2.75701370818306"
## [1] "Male first author team size 2018 geometric mean: 2.65478454977709"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 20000, p-value = 0.4
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 2.32123230837087"
## [1] "Male last author team size 2018 geometric mean: 2.70225263033271"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 11000, p-value = 0.2
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.081  1           1.040
## LastAuthorFemale  1.076  1           1.037
## UniqueAuthors     1.056  4           1.007
## Year              1.057 16           1.002
```



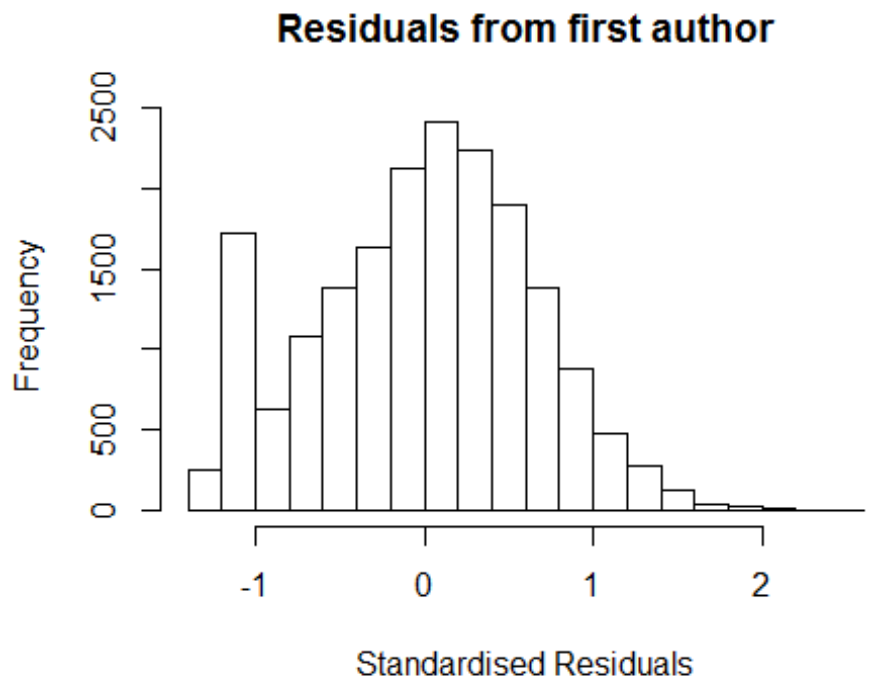

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
UniqueAuthors +
##     Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##     k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.4875 -0.4326  0.0242  0.4261  2.4709
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)          1.0336     0.0262   39.41  < 2e-16 ***
## FirstAuthorFemale1  -0.0168     0.0152   -1.10  0.26953
## LastAuthorFemale1   -0.0323     0.0168   -1.92  0.05496 .
## UniqueAuthors2       0.3664     0.0123   29.71  < 2e-16 ***
## UniqueAuthors3       0.4539     0.0137   33.22  < 2e-16 ***
## UniqueAuthors4       0.4385     0.0177   24.83  < 2e-16 ***
## UniqueAuthors5       0.3638     0.0165   22.09  < 2e-16 ***
## Year1997            -0.0546     0.0329   -1.66  0.09678 .
## Year1998            -0.0869     0.0323   -2.69  0.00722 **
## Year1999            -0.1013     0.0336   -3.01  0.00261 **
```

```
## Year2000            -0.1407     0.0328   -4.29  1.8e-05 ***
## Year2001            -0.2424     0.0349   -6.94  3.9e-12 ***
## Year2002            -0.2000     0.0340   -5.88  4.1e-09 ***
## Year2003            -0.1211     0.0338   -3.58  0.00034 ***
## Year2004            -0.1210     0.0323   -3.75  0.00018 ***
## Year2005            -0.1877     0.0313   -6.00  2.0e-09 ***
## Year2006            -0.1081     0.0307   -3.52  0.00043 ***
## Year2007            -0.1445     0.0303   -4.77  1.8e-06 ***
## Year2008            -0.0960     0.0300   -3.20  0.00136 **
## Year2009            -0.1573     0.0295   -5.34  9.7e-08 ***
## Year2010            -0.1178     0.0303   -3.89  1.0e-04 ***
## Year2011            -0.1870     0.0315   -5.94  2.9e-09 ***
## Year2012            -0.1628     0.0315   -5.18  2.3e-07 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.633
## Multiple R-squared:  0.0932, Adjusted R-squared:  0.0922
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  1623 weights are ~= 1. The remaining 17005 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0937  0.8630  0.9500  0.9100  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          5.37e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.055  1           1.027
## LastAuthorFemale  1.055  1           1.027
## Year              1.007 16           1.000
```



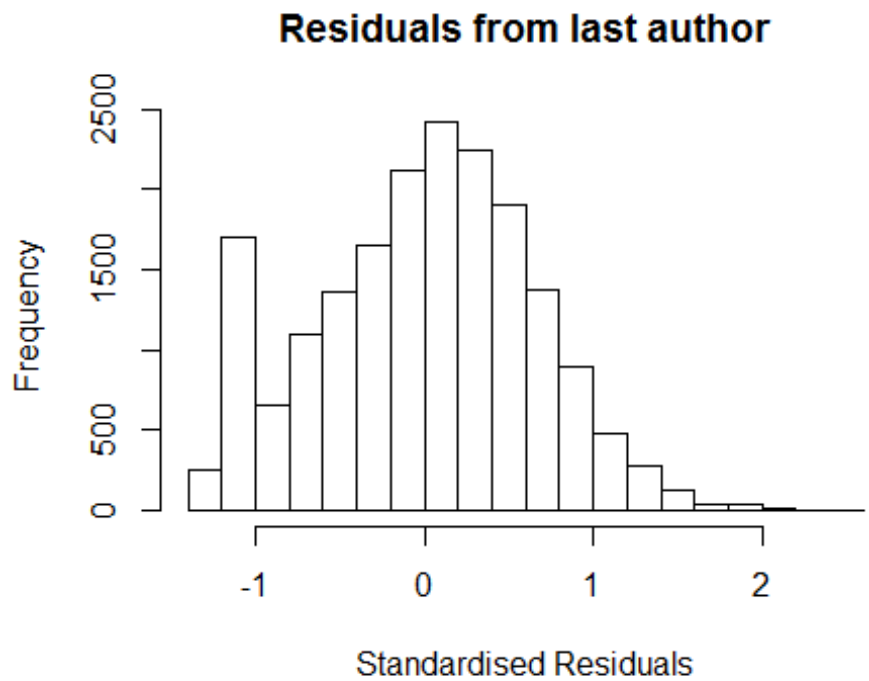

```
## [1] "List of  2 outliers with residuals above  2.5"
##         ScopusId  NLCS Year OneField Fields residuals
## 9974  0007084601 3.700 1999     3106      1     2.558
## 15170 0034644072 3.704 2000     3106      1     2.567
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.2837 -0.4647  0.0425  0.4473  2.5673
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)           1.2622     0.0266   47.37  < 2e-16 ***
## FirstAuthorFemale1    0.0215     0.0158    1.36  0.17368
## LastAuthorFemale1    -0.0378     0.0178   -2.13  0.03340 *
## Year1997             -0.0560     0.0343   -1.64  0.10204
## Year1998             -0.1261     0.0335   -3.76  0.00017 ***
## Year1999             -0.1198     0.0353   -3.40  0.00069 ***
## Year2000             -0.1470     0.0343   -4.29  1.8e-05 ***
## Year2001             -0.2522     0.0376   -6.71  2.0e-11 ***
## Year2002             -0.1837     0.0360   -5.10  3.5e-07 ***
## Year2003             -0.0841     0.0353   -2.38  0.01730 *
## Year2004             -0.0822     0.0335   -2.45  0.01420 *
```

```
## Year2005            -0.1812     0.0334   -5.43  5.6e-08 ***
## Year2006            -0.0680     0.0319   -2.13  0.03332 *
## Year2007            -0.0939     0.0319   -2.94  0.00329 **
## Year2008            -0.0378     0.0313   -1.21  0.22798
## Year2009            -0.1141     0.0311   -3.67  0.00024 ***
## Year2010            -0.0437     0.0316   -1.39  0.16599
## Year2011            -0.1355     0.0331   -4.09  4.3e-05 ***
## Year2012            -0.0924     0.0327   -2.83  0.00471 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.654
## Multiple R-squared:  0.00833,    Adjusted R-squared:  0.00737
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
##  1565 weights are ~= 1. The remaining 17063 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0889  0.8620  0.9480  0.9090  0.9860  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          5.37e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.003  1           1.002
## Year              1.003 16           1.000
```




```
## [1] "List of  2 outliers with residuals above  2.5"
##         ScopusId  NLCS Year OneField Fields residuals
## 9974  0007084601 3.700 1999     3106      1     2.558
## 15170 0034644072 3.704 2000     3106      1     2.567
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.2704 -0.4641  0.0426  0.4468  2.5802
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)           1.2599     0.0266   47.37  < 2e-16 ***
## FirstAuthorFemale1    0.0106     0.0156    0.68  0.49745
## Year1997             -0.0553     0.0342   -1.62  0.10621
## Year1998             -0.1254     0.0335   -3.74  0.00018 ***
## Year1999             -0.1195     0.0353   -3.39  0.00071 ***
## Year2000             -0.1467     0.0343   -4.28  1.9e-05 ***
## Year2001             -0.2527     0.0376   -6.72  1.8e-11 ***
## Year2002             -0.1838     0.0360   -5.10  3.5e-07 ***
## Year2003             -0.0843     0.0353   -2.39  0.01696 *
## Year2004             -0.0819     0.0335   -2.45  0.01445 *
## Year2005             -0.1810     0.0333   -5.43  5.7e-08 ***
```

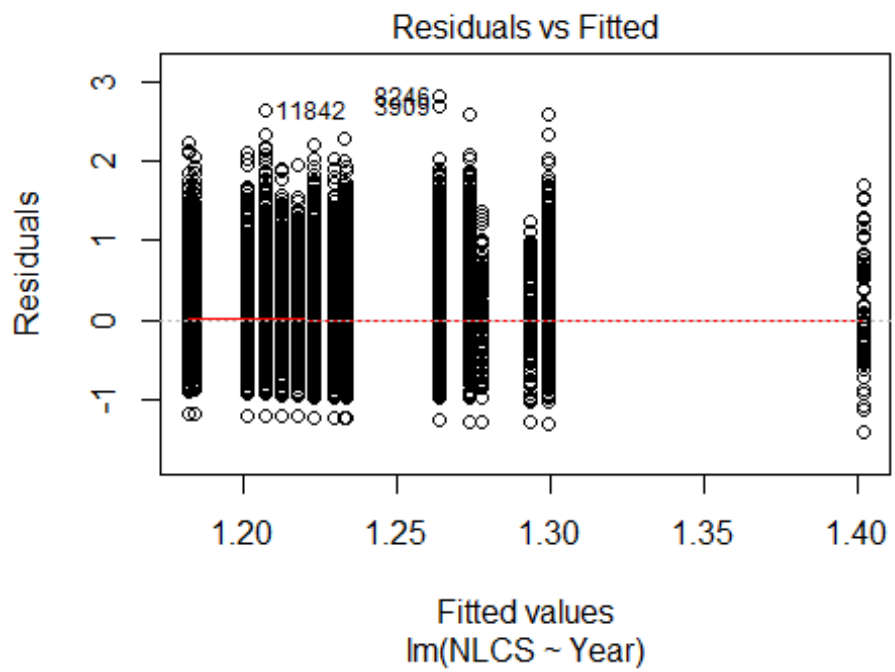
```
## Year2006            -0.0669     0.0319   -2.10  0.03610 *
## Year2007            -0.0939     0.0319   -2.94  0.00326 **
## Year2008            -0.0373     0.0313   -1.19  0.23417
## Year2009            -0.1145     0.0310   -3.69  0.00023 ***
## Year2010            -0.0434     0.0316   -1.38  0.16906
## Year2011            -0.1358     0.0331   -4.10  4.1e-05 ***
## Year2012            -0.0918     0.0327   -2.81  0.00500 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.654
## Multiple R-squared:  0.0081, Adjusted R-squared:  0.0072
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
##  1564 weights are ~= 1. The remaining 17064 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0847  0.8620  0.9480  0.9090  0.9860  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          5.37e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##                   GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.003  1           1.001
## Year             1.003 16           1.000
```



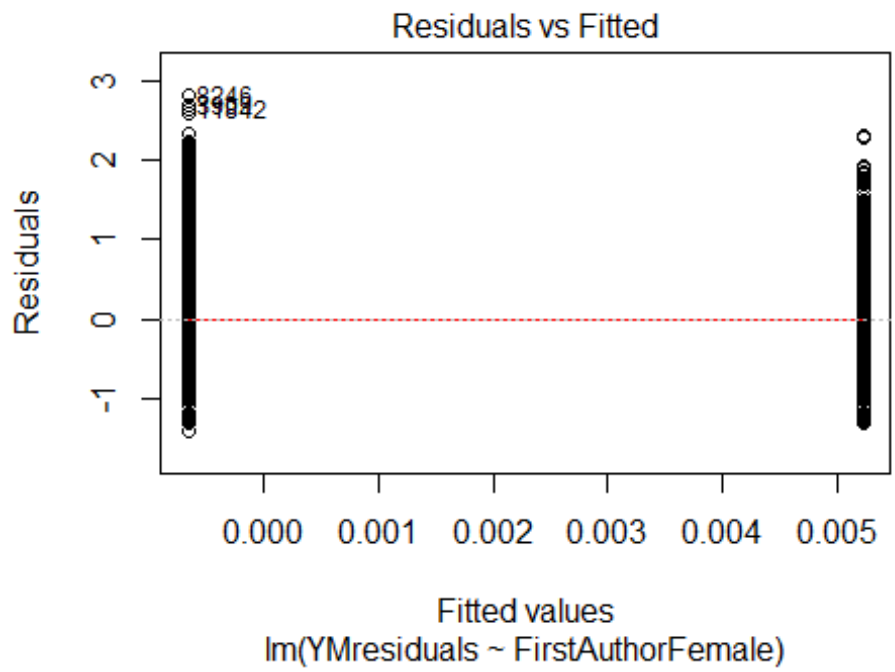

```
## [1] "List of  2 outliers with residuals above  2.5"
##         ScopusId  NLCS Year OneField Fields residuals
## 9974  0007084601 3.700 1999     3106      1     2.558
## 15170 0034644072 3.704 2000     3106      1     2.567
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.2631 -0.4658  0.0425  0.4465  2.5874
##
## Coefficients:
##                   Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.2631     0.0267   47.36  < 2e-16 ***
## LastAuthorFemale1  -0.0302     0.0173   -1.74  0.08160 .
## Year1997           -0.0553     0.0343   -1.61  0.10715
## Year1998           -0.1255     0.0336   -3.74  0.00018 ***
## Year1999           -0.1197     0.0353   -3.39  0.00071 ***
## Year2000           -0.1465     0.0343   -4.27  2.0e-05 ***
## Year2001           -0.2522     0.0376   -6.71  2.1e-11 ***
## Year2002           -0.1833     0.0361   -5.07  3.9e-07 ***
## Year2003           -0.0828     0.0354   -2.34  0.01926 *
## Year2004           -0.0813     0.0335   -2.42  0.01540 *
## Year2005           -0.1804     0.0334   -5.40  6.6e-08 ***
```

```
## Year2006           -0.0665     0.0320   -2.08  0.03744 *
## Year2007           -0.0930     0.0320   -2.91  0.00363 **
## Year2008           -0.0365     0.0314   -1.16  0.24425
## Year2009           -0.1130     0.0311   -3.63  0.00028 ***
## Year2010           -0.0422     0.0316   -1.33  0.18204
## Year2011           -0.1341     0.0331   -4.05  5.2e-05 ***
## Year2012           -0.0912     0.0327   -2.79  0.00532 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.649
## Multiple R-squared:  0.00828,    Adjusted R-squared:  0.00738
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
##  1552 weights are ~= 1. The remaining 17076 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0762  0.8600  0.9480  0.9080  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          5.37e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Sample size for the above analysis:  18628"
## [1] ""
## [1] ""
## [1] "####################################"
## [1] "Analysis of AJSC 3107"
## [1] "####################################"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 2738 2404 2506  324 2977 2991 3163 2980  210 3204  355 3428 3304 3128 3423
## 2011 2012
## 3377 3348
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 1151 1070 1177   88 1336 1097 1473 1375   96 1567  159 1686 1726 1610 1773
```

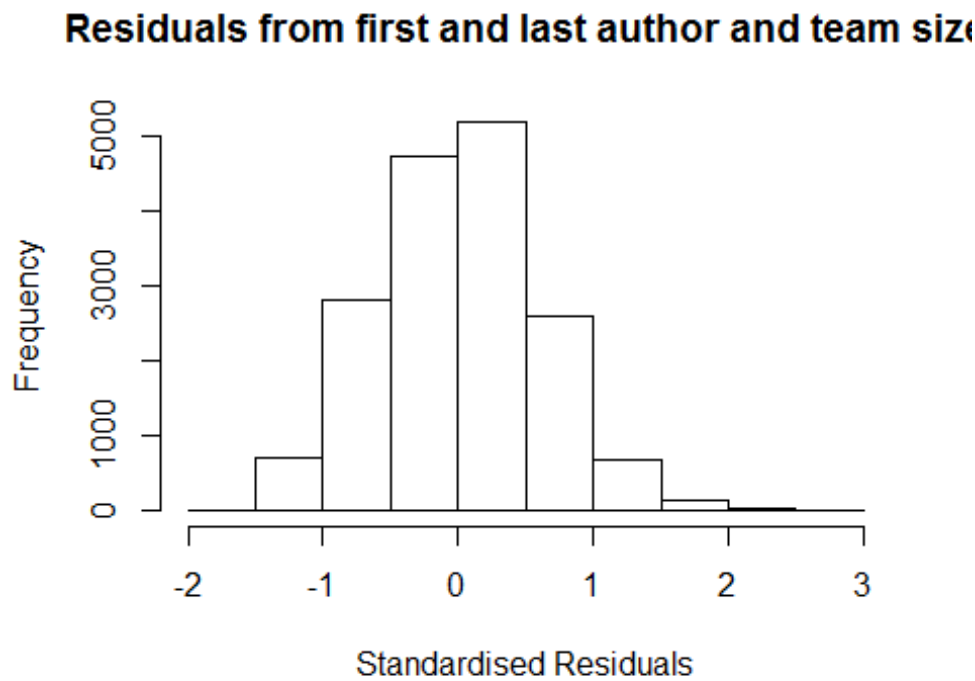
```
## 2011 2012
## 1733 1757
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  946  868  958   71 1109  885 1179 1105   83 1290  127 1356 1368 1285 1427
## 2011 2012
## 1383 1381
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
##  Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 350, df = 16, p-value <2e-16
```




```
##
##  Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 9.8, df = 1, p-value = 0.002
```

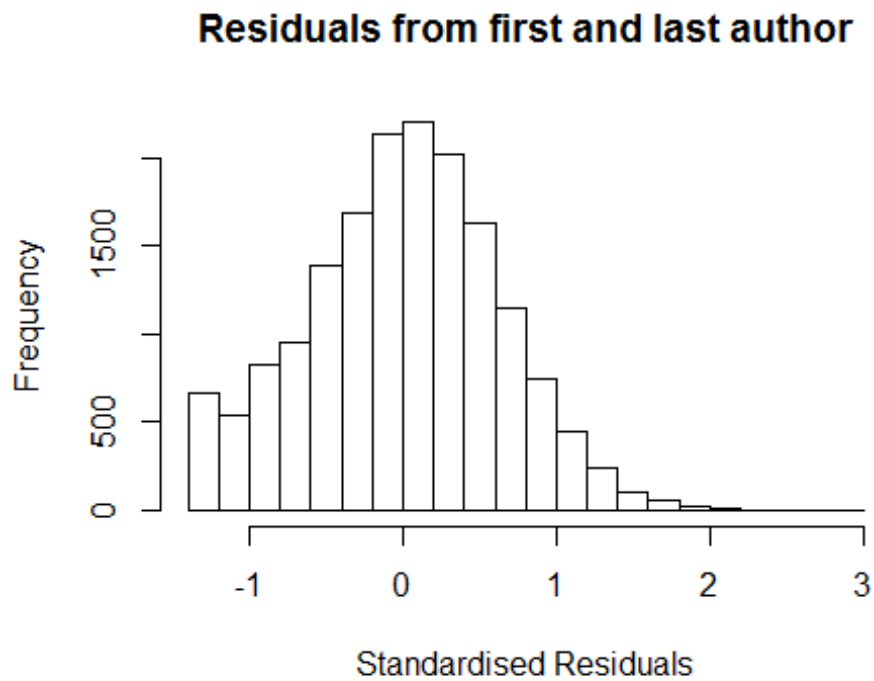



```
## [1] "Female first author team size 2018 geometric mean: 3.8532037796516"
## [1] "Male first author team size 2018 geometric mean: 3.23592949620164"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 130000, p-value = 0.001
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 3.79134432190694"
## [1] "Male last author team size 2018 geometric mean: 3.26162594334582"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 1e+05, p-value = 0.01
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##                     GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.039  1           1.019
## LastAuthorFemale  1.027  1           1.013
## UniqueAuthors     1.069  4           1.008
## Year              1.062 16           1.002
```



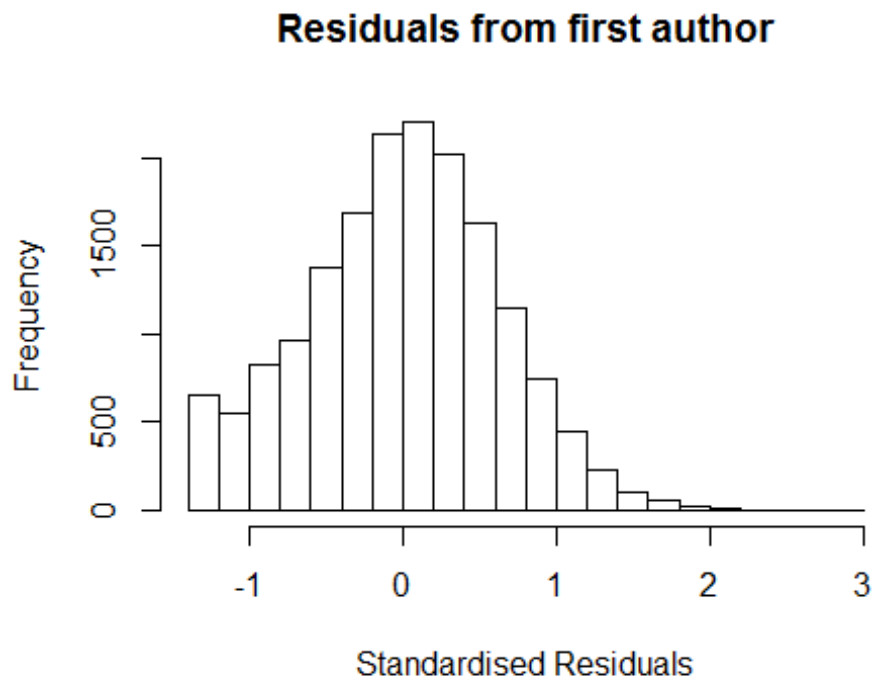

```
## [1] "List of  2 outliers with residuals above  2.5"
##         ScopusId  NLCS Year OneField Fields residuals
## 8246  0141749837 4.080 1998     3107      1     2.826
## 14654 0035288601 3.646 2001     3101      4     2.656
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
UniqueAuthors +
##     Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##     k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.6600 -0.4066  0.0176  0.4110  2.8255
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)         0.986960   0.026419   37.36  < 2e-16 ***
## FirstAuthorFemale1 -0.028090   0.014977   -1.88  0.06074 .
## LastAuthorFemale1  -0.019287   0.017184   -1.12  0.26171
## UniqueAuthors2      0.275859   0.015405   17.91  < 2e-16 ***
## UniqueAuthors3      0.379542   0.016124   23.54  < 2e-16 ***
## UniqueAuthors4      0.430720   0.017621   24.44  < 2e-16 ***
## UniqueAuthors5      0.540789   0.015948   33.91  < 2e-16 ***
## Year1997           -0.000984   0.035454   -0.03  0.97786
## Year1998           -0.008365   0.032912   -0.25  0.79936
```

```
## Year1999            0.132244   0.103831    1.27  0.20281
## Year2000           -0.046380   0.031877   -1.45  0.14570
## Year2001            0.003430   0.034380    0.10  0.92053
## Year2002           -0.038088   0.031789   -1.20  0.23088
## Year2003           -0.105009   0.032162   -3.27  0.00110 **
## Year2004            0.033311   0.089567    0.37  0.70996
## Year2005           -0.104806   0.028981   -3.62  0.00030 ***
## Year2006            0.048317   0.063342    0.76  0.44559
## Year2007           -0.093591   0.028621   -3.27  0.00108 **
## Year2008           -0.108538   0.028771   -3.77  0.00016 ***
## Year2009           -0.136687   0.028785   -4.75  2.1e-06 ***
## Year2010           -0.125374   0.028671   -4.37  1.2e-05 ***
## Year2011           -0.148557   0.028811   -5.16  2.5e-07 ***
## Year2012           -0.166553   0.028725   -5.80  6.8e-09 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.604
## Multiple R-squared:  0.0846, Adjusted R-squared:  0.0834
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
##  1386 weights are ~= 1. The remaining 15435 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.000   0.865   0.950   0.907   0.986   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          5.94e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.019  1           1.009
## LastAuthorFemale  1.012  1           1.006
## Year              1.013 16           1.000
```



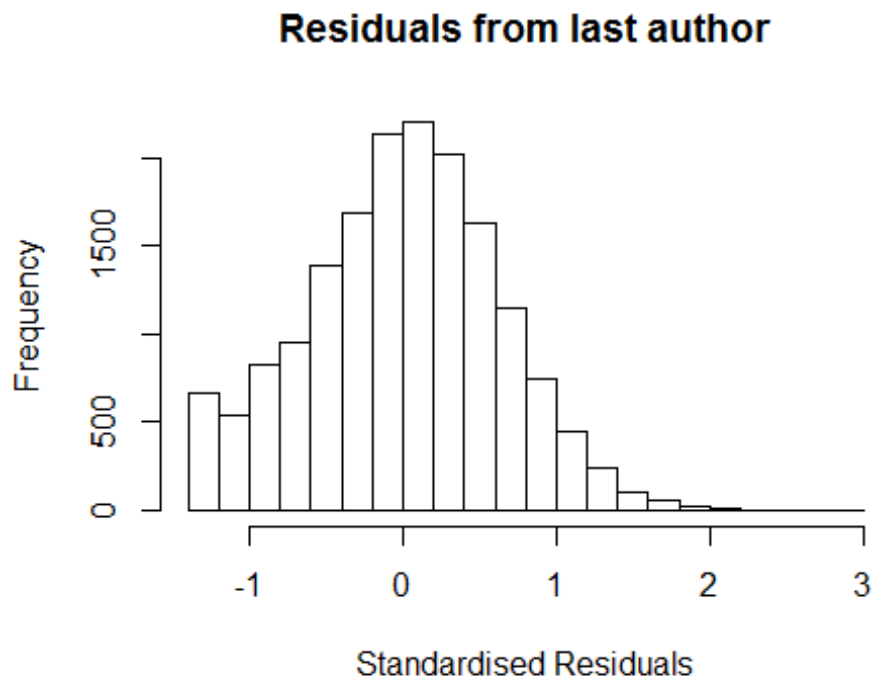

```
## [1] "List of  2 outliers with residuals above  2.5"
##        ScopusId  NLCS Year OneField Fields residuals
## 2164 4243216277 3.863 1996     3107      1     2.619
## 8246 0141749837 4.080 1998     3107      1     2.832
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##    Min     1Q Median     3Q    Max
## -1.384 -0.419  0.024  0.422  2.832
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)          1.24398    0.02459   50.59   <2e-16 ***
## FirstAuthorFemale1   0.02229    0.01531    1.46   0.1455
## LastAuthorFemale1    0.00656    0.01794    0.37   0.7147
## Year1997            -0.00202    0.03728   -0.05   0.9567
## Year1998             0.00404    0.03366    0.12   0.9044
## Year1999             0.13999    0.09749    1.44   0.1510
## Year2000            -0.05009    0.03252   -1.54   0.1235
## Year2001             0.02777    0.03529    0.79   0.4313
## Year2002            -0.01544    0.03269   -0.47   0.6367
## Year2003            -0.07303    0.03329   -2.19   0.0282 *
## Year2004             0.08386    0.07871    1.07   0.2867
```

```
## Year2005           -0.04840    0.02953   -1.64   0.1012
## Year2006            0.09357    0.05758    1.62   0.1042
## Year2007           -0.01695    0.02913   -0.58   0.5605
## Year2008           -0.03121    0.02936   -1.06   0.2878
## Year2009           -0.05469    0.02933   -1.86   0.0623 .
## Year2010           -0.03535    0.02928   -1.21   0.2274
## Year2011           -0.07597    0.02953   -2.57   0.0101 *
## Year2012           -0.07617    0.02937   -2.59   0.0095 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.624
## Multiple R-squared:  0.0031, Adjusted R-squared:  0.00203
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
##  1416 weights are ~= 1. The remaining 15405 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0039  0.8660  0.9500  0.9050  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          5.94e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##                   GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.01  1           1.005
## Year              1.01 16           1.000
```




```
## [1] "List of  2 outliers with residuals above  2.5"
##        ScopusId  NLCS Year OneField Fields residuals
## 2164 4243216277 3.863 1996     3107      1     2.619
## 8246 0141749837 4.080 1998     3107      1     2.832
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.3846 -0.4191  0.0236  0.4217  2.8317
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.24434    0.02457   50.65   <2e-16 ***
## FirstAuthorFemale1  0.02324    0.01528    1.52   0.1284
## Year1997           -0.00215    0.03729   -0.06   0.9540
## Year1998            0.00399    0.03366    0.12   0.9056
## Year1999            0.14024    0.09755    1.44   0.1506
## Year2000           -0.05006    0.03253   -1.54   0.1238
## Year2001            0.02780    0.03529    0.79   0.4308
## Year2002           -0.01540    0.03269   -0.47   0.6375
## Year2003           -0.07290    0.03328   -2.19   0.0285 *
## Year2004            0.08399    0.07876    1.07   0.2863
## Year2005           -0.04834    0.02953   -1.64   0.1017
```

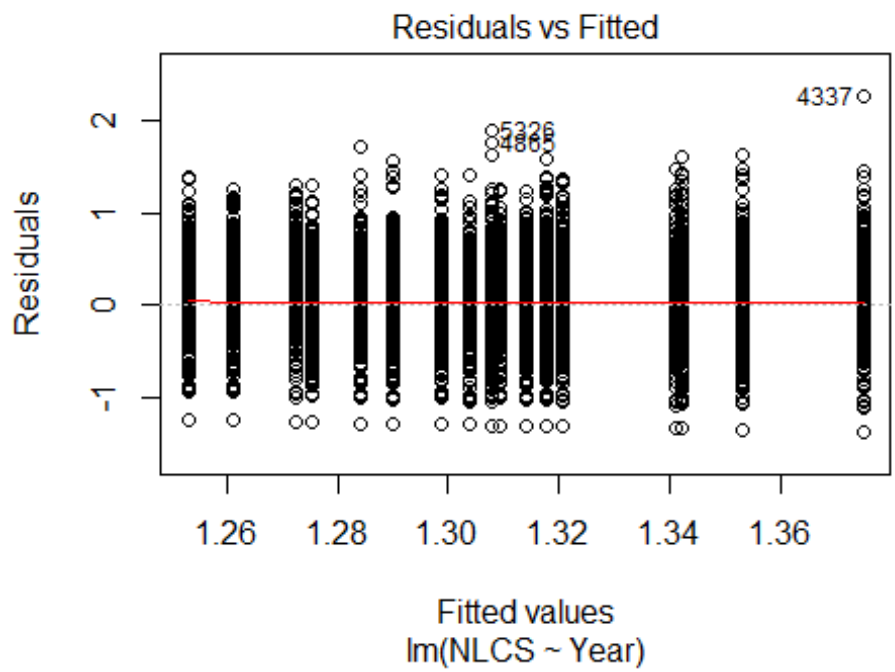
```
## Year2006            0.09340    0.05757    1.62   0.1047
## Year2007           -0.01691    0.02913   -0.58   0.5615
## Year2008           -0.03115    0.02936   -1.06   0.2888
## Year2009           -0.05462    0.02933   -1.86   0.0626 .
## Year2010           -0.03520    0.02928   -1.20   0.2293
## Year2011           -0.07590    0.02953   -2.57   0.0102 *
## Year2012           -0.07598    0.02937   -2.59   0.0097 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.624
## Multiple R-squared:  0.00309,    Adjusted R-squared:  0.00208
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
##  1414 weights are ~= 1. The remaining 15407 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0039  0.8660  0.9500  0.9050  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          5.94e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##                   GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.004  1           1.002
## Year             1.004 16           1.000
```



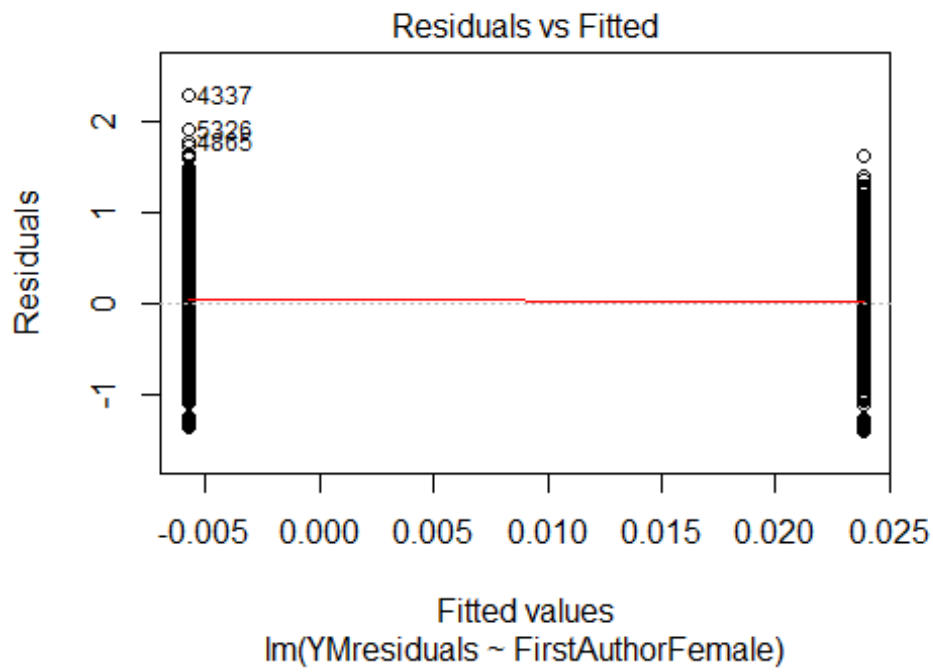

```
## [1] "List of  2 outliers with residuals above  2.5"
##        ScopusId  NLCS Year OneField Fields residuals
## 2164 4243216277 3.863 1996     3107      1     2.619
## 8246 0141749837 4.080 1998     3107      1     2.832
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.3877 -0.4180  0.0244  0.4232  2.8310
##
## Coefficients:
##                   Estimate Std. Error t value Pr(>|t|)
## (Intercept)        1.24540    0.02458   50.67   <2e-16 ***
## LastAuthorFemale1  0.01092    0.01787    0.61    0.541
## Year1997          -0.00224    0.03731   -0.06    0.952
## Year1998           0.00362    0.03366    0.11    0.914
## Year1999           0.14232    0.09713    1.47    0.143
## Year2000          -0.04960    0.03253   -1.52    0.127
## Year2001           0.02862    0.03528    0.81    0.417
## Year2002          -0.01484    0.03271   -0.45    0.650
## Year2003          -0.07272    0.03330   -2.18    0.029 *
## Year2004           0.08367    0.07878    1.06    0.288
## Year2005          -0.04758    0.02953   -1.61    0.107
```

```
## Year2006           0.09423    0.05754    1.64    0.102
## Year2007          -0.01607    0.02912   -0.55    0.581
## Year2008          -0.02981    0.02935   -1.02    0.310
## Year2009          -0.05352    0.02932   -1.83    0.068 .
## Year2010          -0.03421    0.02927   -1.17    0.243
## Year2011          -0.07473    0.02951   -2.53    0.011 *
## Year2012          -0.07476    0.02934   -2.55    0.011 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.624
## Multiple R-squared:  0.00298,    Adjusted R-squared:  0.00197
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
##  1389 weights are ~= 1. The remaining 15432 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.004   0.866   0.950   0.906   0.985   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          5.94e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                    psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Sample size for the above analysis:  16821"
## [1] ""
## [1] ""
## [1] "####################################"
## [1] "Analysis of AJSC 3108"
## [1] "####################################"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  730  722  629  764  662  715  747  545  482  582  774  678  727  590  662
## 2011 2012
##  716  867
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  411  433  380  335  375  313  431  320  308  369  442  441  468  382  430
```

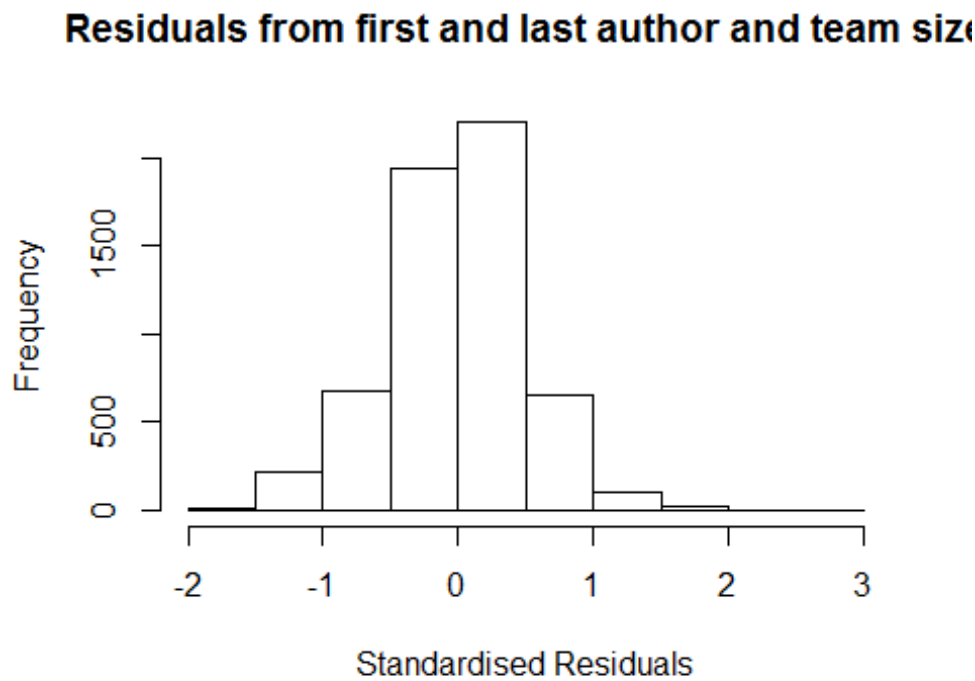
```
## 2011 2012
##  461  558
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  349  381  339  292  316  266  379  266  264  318  377  368  390  313  366
## 2011 2012
##  380  462
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
##  Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 60, df = 16, p-value = 5e-07
```




```
##
##  Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 20, df = 1, p-value = 7e-06
```

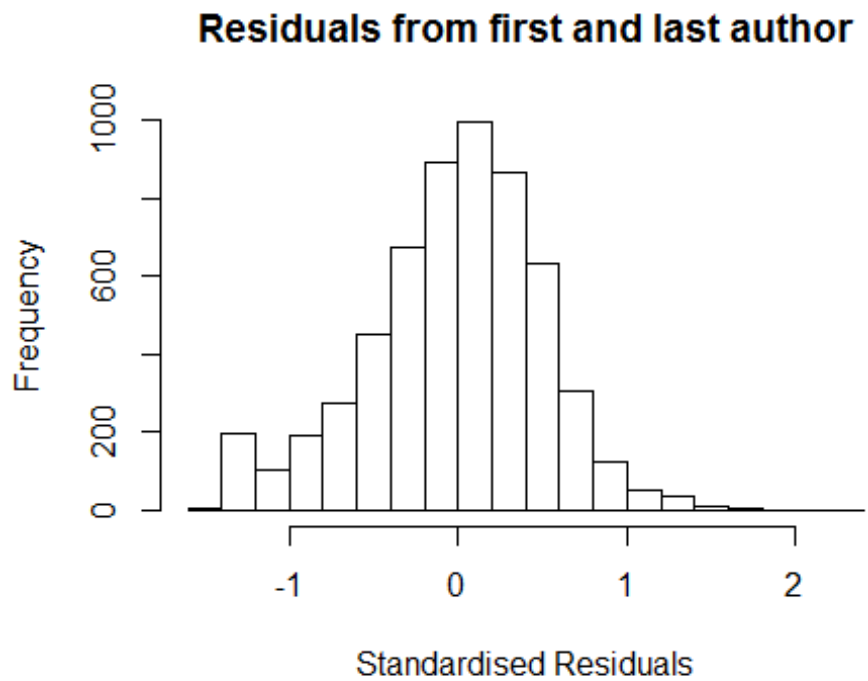



```
## [1] "Female first author team size 2018 geometric mean: 5.67457761784722"
## [1] "Male first author team size 2018 geometric mean: 4.52188270232901"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 13000, p-value = 0.02
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 6.67360124256601"
## [1] "Male last author team size 2018 geometric mean: 4.43982728079954"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 12000, p-value = 6e-05
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##                     GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.045  1           1.022
## LastAuthorFemale  1.033  1           1.016
## UniqueAuthors     1.119  4           1.014
## Year              1.136 16           1.004
```



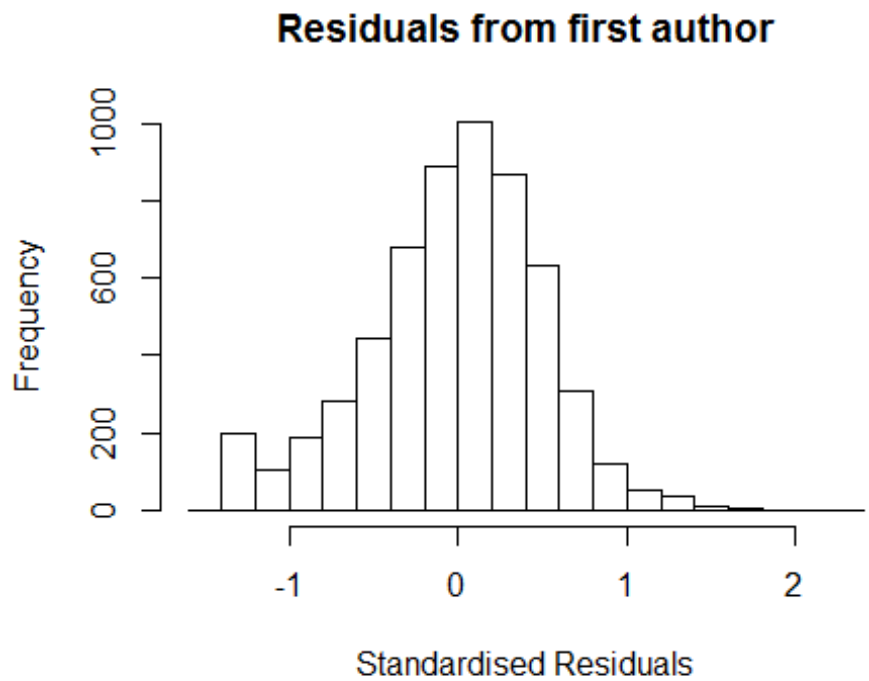

```
## [1] "List of  1 outliers with residuals above  2.5"
##        ScopusId  NLCS Year OneField Fields residuals
## 4337 0035288601 3.646 2001     3101      4     2.548
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
UniqueAuthors +
##     Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##     k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.5415 -0.3076  0.0156  0.3143  2.5482
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.05879    0.04301   24.62  < 2e-16 ***
## FirstAuthorFemale1 -0.01862    0.01562   -1.19  0.23331
## LastAuthorFemale1   0.00636    0.01878    0.34  0.73475
## UniqueAuthors2      0.13330    0.03703    3.60  0.00032 ***
## UniqueAuthors3      0.28913    0.03648    7.92  2.7e-15 ***
## UniqueAuthors4      0.31108    0.03526    8.82  < 2e-16 ***
## UniqueAuthors5      0.44364    0.03136   14.15  < 2e-16 ***
## Year1997            0.01225    0.04230    0.29  0.77203
## Year1998           -0.00179    0.04367   -0.04  0.96722
## Year1999           -0.08269    0.04229   -1.96  0.05058 .
```

```
## Year2000           -0.00806    0.04058   -0.20  0.84253
## Year2001            0.03904    0.04693    0.83  0.40549
## Year2002           -0.05864    0.04174   -1.40  0.16008
## Year2003           -0.06716    0.04278   -1.57  0.11647
## Year2004           -0.10998    0.04317   -2.55  0.01087 *
## Year2005           -0.06794    0.04079   -1.67  0.09586 .
## Year2006           -0.05542    0.03957   -1.40  0.16141
## Year2007           -0.06409    0.03865   -1.66  0.09734 .
## Year2008           -0.04327    0.03920   -1.10  0.26976
## Year2009           -0.07713    0.04168   -1.85  0.06430 .
## Year2010           -0.09948    0.04039   -2.46  0.01380 *
## Year2011           -0.13574    0.03966   -3.42  0.00063 ***
## Year2012           -0.12009    0.03886   -3.09  0.00201 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.46
## Multiple R-squared:  0.0897, Adjusted R-squared:  0.0862
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
##  observation 1851 is an outlier with |weight| = 0 ( < 1.7e-05);
##  518 weights are ~= 1. The remaining 5307 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0084  0.8630  0.9500  0.8920  0.9860  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.72e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.022  1           1.011
## LastAuthorFemale  1.030  1           1.015
## Year              1.042 16           1.001
```



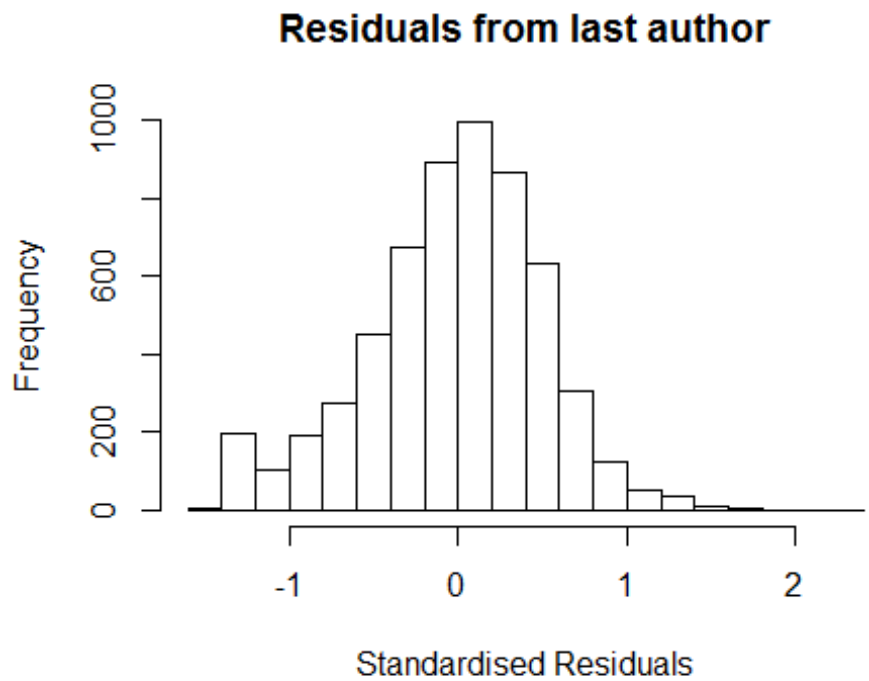

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.4333 -0.3299  0.0252  0.3239  2.2562
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)          1.35473    0.03253   41.64   <2e-16 ***
## FirstAuthorFemale1   0.02518    0.01593    1.58    0.114
## LastAuthorFemale1    0.01825    0.01945    0.94    0.348
## Year1997             0.01366    0.04357    0.31    0.754
## Year1998             0.00701    0.04442    0.16    0.875
## Year1999            -0.07322    0.04297   -1.70    0.088 .
## Year2000            -0.01026    0.04163   -0.25    0.805
## Year2001             0.03510    0.04820    0.73    0.467
## Year2002            -0.03136    0.04239   -0.74    0.459
## Year2003            -0.02570    0.04434   -0.58    0.562
## Year2004            -0.06755    0.04510   -1.50    0.134
## Year2005            -0.03980    0.04312   -0.92    0.356
```

```
## Year2006           -0.04120    0.04109   -1.00    0.316
## Year2007           -0.01528    0.04064   -0.38    0.707
## Year2008           -0.02534    0.04048   -0.63    0.531
## Year2009           -0.01900    0.04304   -0.44    0.659
## Year2010           -0.04280    0.04148   -1.03    0.302
## Year2011           -0.08426    0.04090   -2.06    0.039 *
## Year2012           -0.06841    0.04042   -1.69    0.091 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.478
## Multiple R-squared:  0.00442,    Adjusted R-squared:  0.00134
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
##  observation 1851 is an outlier with |weight| = 0 ( < 1.7e-05);
##  509 weights are ~= 1. The remaining 5316 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0886  0.8630  0.9490  0.8920  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.72e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.018  1           1.009
## Year              1.018 16           1.001
```




```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.4188 -0.3319  0.0253  0.3238  2.2540
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)          1.35692    0.03241   41.87   <2e-16 ***
## FirstAuthorFemale1   0.02678    0.01591    1.68    0.092 .
## Year1997             0.01326    0.04354    0.30    0.761
## Year1998             0.00689    0.04445    0.16    0.877
## Year1999            -0.07474    0.04288   -1.74    0.081 .
## Year2000            -0.01101    0.04161   -0.26    0.791
## Year2001             0.03511    0.04818    0.73    0.466
## Year2002            -0.03232    0.04238   -0.76    0.446
## Year2003            -0.02563    0.04434   -0.58    0.563
## Year2004            -0.06755    0.04508   -1.50    0.134
## Year2005            -0.03944    0.04310   -0.92    0.360
## Year2006            -0.04088    0.04109   -0.99    0.320
```

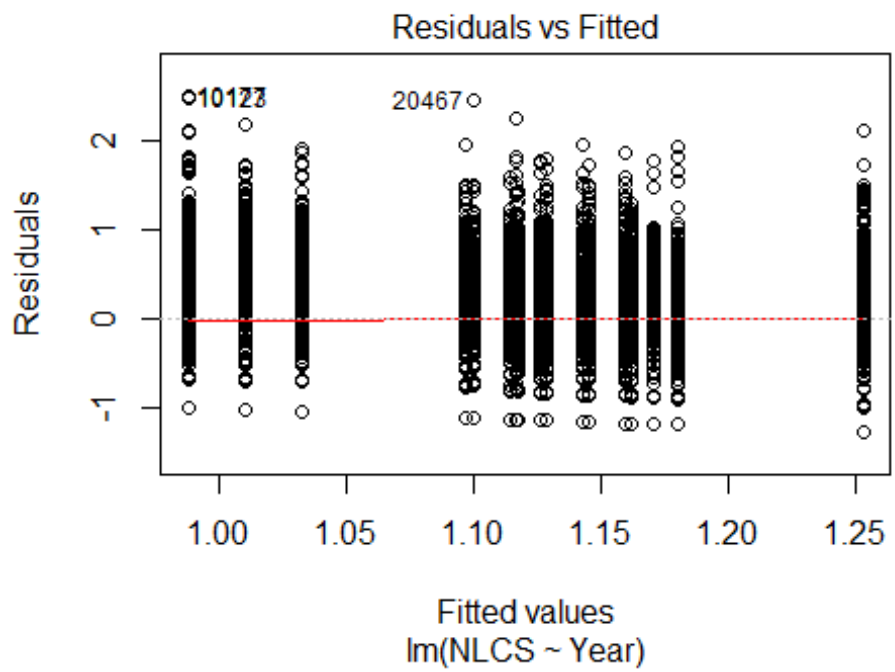
```
## Year2007           -0.01520    0.04064   -0.37    0.708
## Year2008           -0.02496    0.04050   -0.62    0.538
## Year2009           -0.01833    0.04304   -0.43    0.670
## Year2010           -0.04290    0.04147   -1.03    0.301
## Year2011           -0.08371    0.04091   -2.05    0.041 *
## Year2012           -0.06730    0.04043   -1.66    0.096 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.478
## Multiple R-squared:  0.00426,    Adjusted R-squared:  0.00135
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
##  observation 1851 is an outlier with |weight| = 0 ( < 1.7e-05);
##  507 weights are ~= 1. The remaining 5318 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0891  0.8640  0.9490  0.8920  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.72e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##                   GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.026  1           1.013
## Year             1.026 16           1.001
```



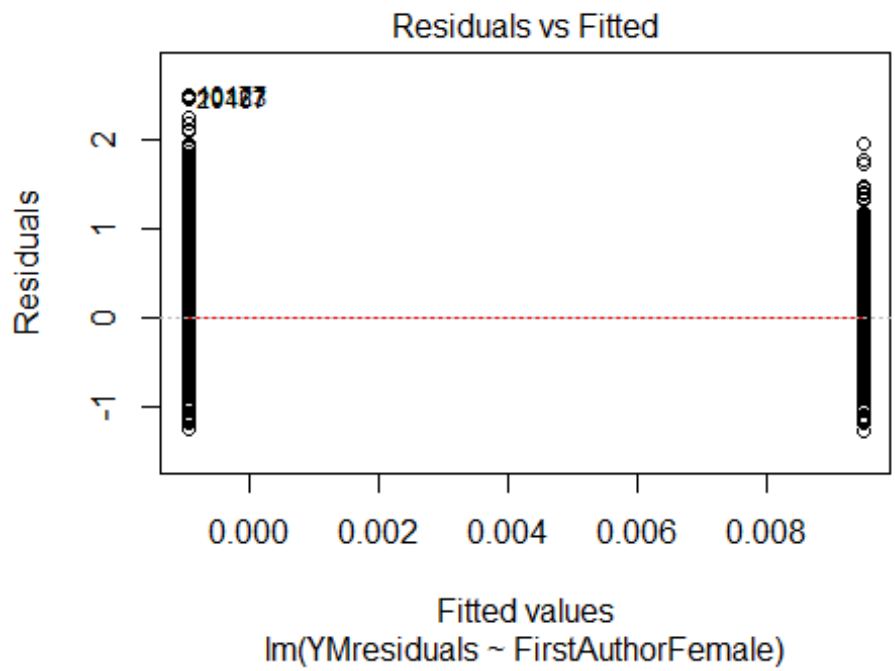

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.4158 -0.3328  0.0275  0.3227  2.2515
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.35919    0.03237   41.99   <2e-16 ***
## LastAuthorFemale1   0.02130    0.01941    1.10    0.273
## Year1997            0.01349    0.04359    0.31    0.757
## Year1998            0.00760    0.04443    0.17    0.864
## Year1999           -0.07285    0.04293   -1.70    0.090 .
## Year2000           -0.00991    0.04163   -0.24    0.812
## Year2001            0.03530    0.04825    0.73    0.464
## Year2002           -0.03239    0.04235   -0.76    0.444
## Year2003           -0.02618    0.04438   -0.59    0.555
## Year2004           -0.06777    0.04510   -1.50    0.133
## Year2005           -0.04020    0.04311   -0.93    0.351
## Year2006           -0.04076    0.04110   -0.99    0.321
```

```
## Year2007          -0.01553    0.04063   -0.38    0.702
## Year2008          -0.02513    0.04045   -0.62    0.534
## Year2009          -0.01718    0.04300   -0.40    0.690
## Year2010          -0.04203    0.04151   -1.01    0.311
## Year2011          -0.08325    0.04085   -2.04    0.042 *
## Year2012          -0.06677    0.04040   -1.65    0.098 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.478
## Multiple R-squared:  0.00401,    Adjusted R-squared:  0.0011
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
##  observation 1851 is an outlier with |weight| = 0 ( < 1.7e-05);
##  523 weights are ~= 1. The remaining 5302 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0901  0.8640  0.9490  0.8920  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.72e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                    psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Sample size for the above analysis:  5826"
## [1] ""
## [1] ""
## [1] "####################################"
## [1] "Analysis of AJSC 3109"
## [1] "####################################"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 1038 1070  978  963  844 1212 1125 1224 1228  994  988  980  929 1021  901
## 2011 2012
##  921  850
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  493  500  464  319  381  470  568  581  561  532  500  536  549  578  501
```

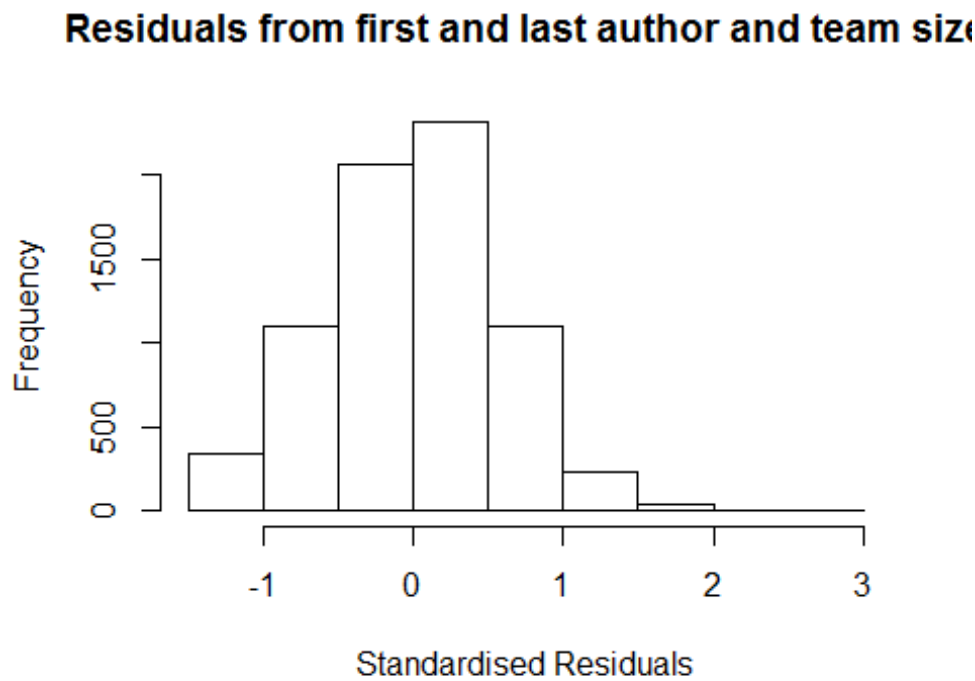
```
## 2011 2012
##  529  500
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  420  412  388  271  324  394  491  471  464  454  409  453  471  491  408
## 2011 2012
##  442  413
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
##  Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 200, df = 16, p-value <2e-16
```




```
##
##  Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 4.8, df = 1, p-value = 0.03
```

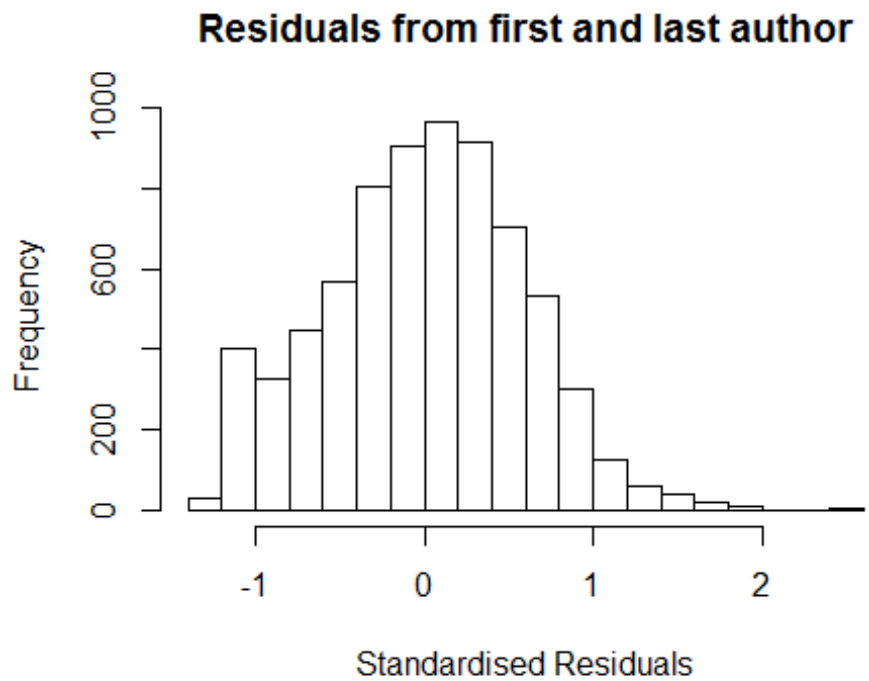



```
## [1] "Female first author team size 2018 geometric mean: 2.29954440448131"
## [1] "Male first author team size 2018 geometric mean: 2.10541019326293"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 6300, p-value = 0.5
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 2.15428853462788"
## [1] "Male last author team size 2018 geometric mean: 2.117522207611"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 5800, p-value = 0.8
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.120  1           1.058
## LastAuthorFemale  1.113  1           1.055
## UniqueAuthors     1.074  4           1.009
## Year              1.087 16           1.003
```



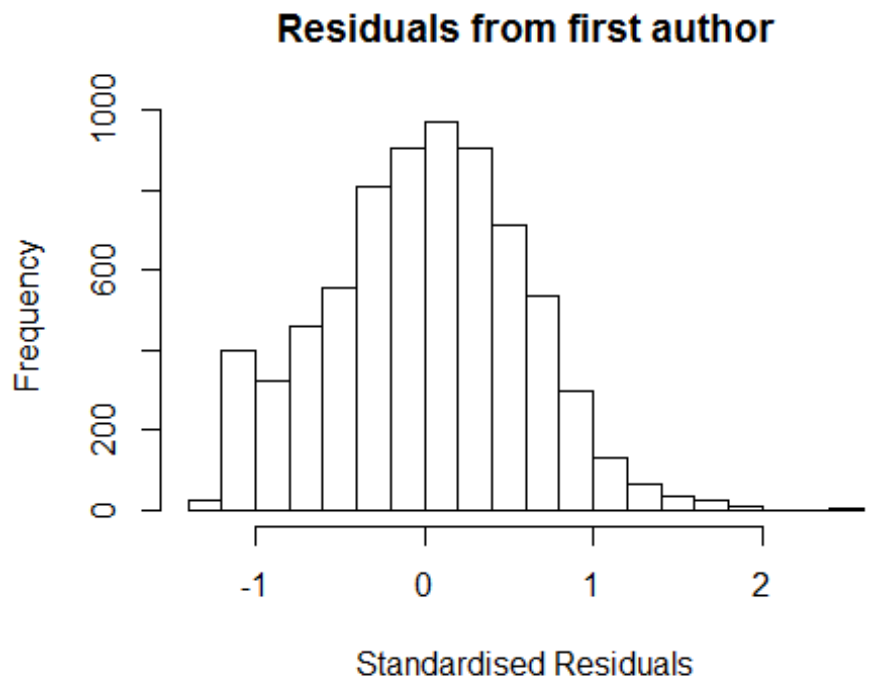

```
## [1] "List of  2 outliers with residuals above  2.5"
##          ScopusId  NLCS Year OneField Fields residuals
## 10123 11244344133 3.457 2004     2610      2     2.663
## 20467 84862158214 3.539 2012     2610      2     2.648
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
UniqueAuthors +
##     Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##     k.max = 1000))
##  \--> method = "MM"
## Residuals:
##    Min     1Q Median     3Q    Max
## -1.426 -0.388  0.021  0.385  2.663
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)          1.1108     0.0343   32.43  < 2e-16 ***
## FirstAuthorFemale1  -0.0354     0.0242   -1.47  0.14283
## LastAuthorFemale1    0.0435     0.0250    1.74  0.08223 .
## UniqueAuthors2       0.1954     0.0169   11.58  < 2e-16 ***
## UniqueAuthors3       0.3155     0.0204   15.47  < 2e-16 ***
## UniqueAuthors4       0.3704     0.0303   12.22  < 2e-16 ***
## UniqueAuthors5       0.4180     0.0318   13.16  < 2e-16 ***
## Year1997            -0.1190     0.0431   -2.76  0.00573 **
## Year1998            -0.1403     0.0431   -3.25  0.00114 **
```

```
## Year1999            -0.1060     0.0461   -2.30  0.02142 *
## Year2000            -0.1070     0.0422   -2.53  0.01129 *
## Year2001            -0.2126     0.0553   -3.84  0.00012 ***
## Year2002            -0.1208     0.0434   -2.78  0.00537 **
## Year2003            -0.2899     0.0491   -5.90  3.7e-09 ***
## Year2004            -0.3170     0.0482   -6.58  5.1e-11 ***
## Year2005            -0.1528     0.0413   -3.69  0.00022 ***
## Year2006            -0.1159     0.0419   -2.77  0.00570 **
## Year2007            -0.1538     0.0412   -3.73  0.00019 ***
## Year2008            -0.1781     0.0402   -4.43  9.5e-06 ***
## Year2009            -0.1781     0.0401   -4.44  9.3e-06 ***
## Year2010            -0.2051     0.0427   -4.81  1.6e-06 ***
## Year2011            -0.2029     0.0413   -4.91  9.1e-07 ***
## Year2012            -0.2203     0.0421   -5.24  1.7e-07 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.564
## Multiple R-squared:  0.067,  Adjusted R-squared:  0.0642
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
##  2 observations c(3178,6972) are outliers with |weight| = 0 ( < 1.4e-05);
##  621 weights are ~= 1. The remaining 6553 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0406  0.8680  0.9490  0.9040  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.39e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.089  1           1.044
## LastAuthorFemale  1.089  1           1.044
## Year              1.019 16           1.001
```



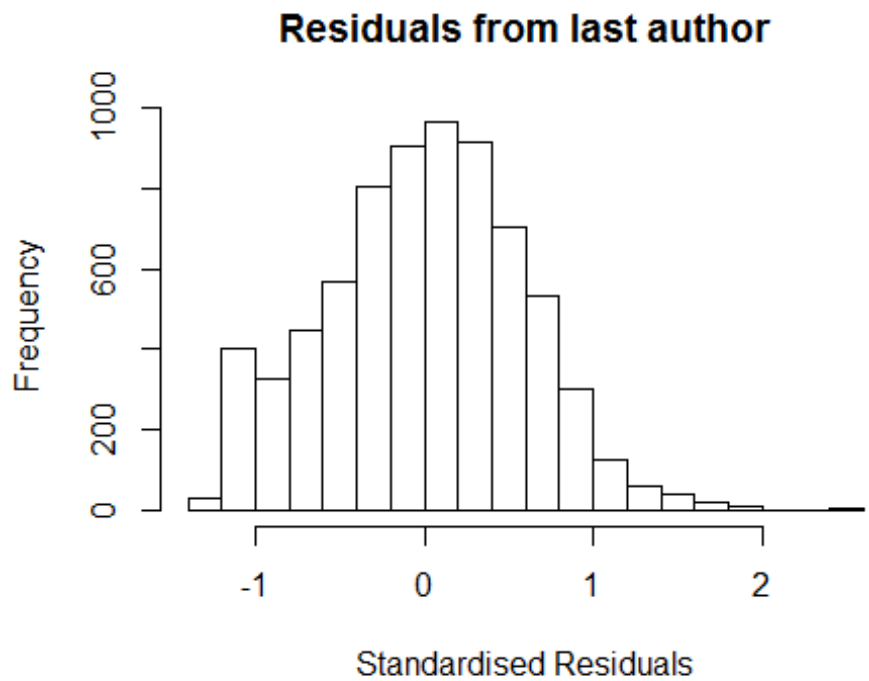

```
## [1] "List of  2 outliers with residuals above  2.5"
##          ScopusId  NLCS Year OneField Fields residuals
## 10123 11244344133 3.457 2004     2610      2     2.500
## 10177 41349117788 3.471 2004     2613      3     2.514
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.3229 -0.3921  0.0206  0.4039  2.5143
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)          1.25685    0.03336   37.67  < 2e-16 ***
## FirstAuthorFemale1   0.00568    0.02509    0.23  0.82080
## LastAuthorFemale1    0.06032    0.02621    2.30  0.02137 *
## Year1997            -0.12853    0.04455   -2.89  0.00392 **
## Year1998            -0.14918    0.04479   -3.33  0.00087 ***
## Year1999            -0.11932    0.04816   -2.48  0.01325 *
## Year2000            -0.10103    0.04374   -2.31  0.02093 *
## Year2001            -0.21433    0.05547   -3.86  0.00011 ***
## Year2002            -0.09804    0.04393   -2.23  0.02567 *
## Year2003            -0.27167    0.04945   -5.49  4.1e-08 ***
## Year2004            -0.30015    0.04809   -6.24  4.6e-10 ***
```

```
## Year2005           -0.13799    0.04221   -3.27  0.00108 **
## Year2006           -0.09904    0.04321   -2.29  0.02195 *
## Year2007           -0.13426    0.04253   -3.16  0.00160 **
## Year2008           -0.15362    0.04148   -3.70  0.00021 ***
## Year2009           -0.15672    0.04117   -3.81  0.00014 ***
## Year2010           -0.16880    0.04406   -3.83  0.00013 ***
## Year2011           -0.16356    0.04272   -3.83  0.00013 ***
## Year2012           -0.17274    0.04267   -4.05  5.2e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.59
## Multiple R-squared:  0.0131, Adjusted R-squared:  0.0106
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
##  638 weights are ~= 1. The remaining 6538 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0303  0.8680  0.9500  0.9070  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.39e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.011  1           1.005
## Year              1.011 16           1.000
```




```
## [1] "List of  2 outliers with residuals above  2.5"
##          ScopusId  NLCS Year OneField Fields residuals
## 10123 11244344133 3.457 2004     2610      2     2.500
## 10177 41349117788 3.471 2004     2613      3     2.514
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##    Min     1Q Median     3Q    Max
## -1.280 -0.391  0.019  0.404  2.512
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)           1.2597     0.0334   37.77  < 2e-16 ***
## FirstAuthorFemale1    0.0206     0.0240    0.86  0.39116
## Year1997             -0.1280     0.0446   -2.87  0.00407 **
## Year1998             -0.1487     0.0448   -3.32  0.00091 ***
## Year1999             -0.1188     0.0483   -2.46  0.01395 *
## Year2000             -0.1017     0.0438   -2.32  0.02025 *
## Year2001             -0.2141     0.0554   -3.86  0.00011 ***
## Year2002             -0.0991     0.0439   -2.26  0.02416 *
## Year2003             -0.2703     0.0495   -5.46  4.9e-08 ***
## Year2004             -0.3006     0.0481   -6.25  4.4e-10 ***
## Year2005             -0.1365     0.0422   -3.23  0.00124 **
```

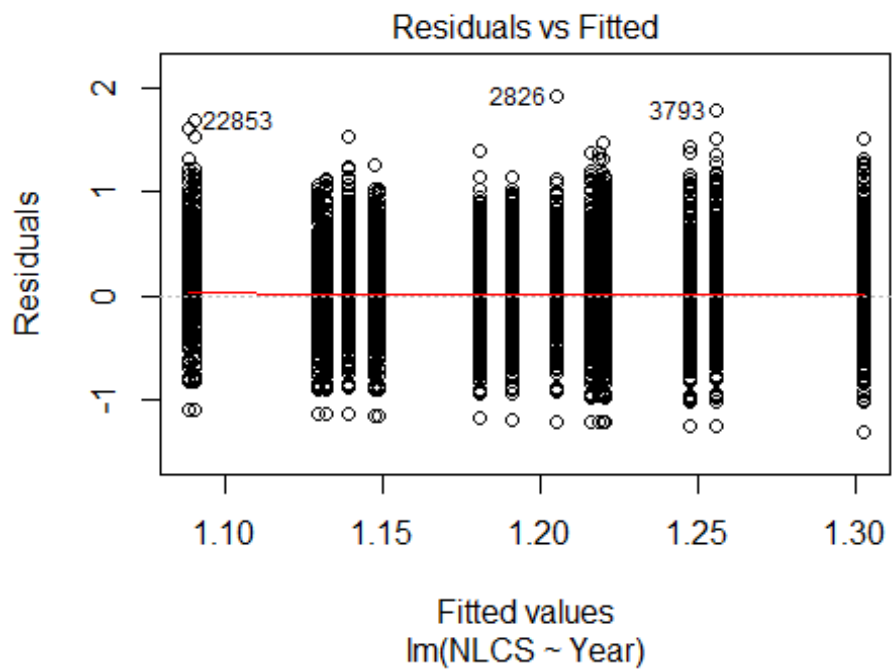
```
## Year2006            -0.1003     0.0432   -2.32  0.02037 *
## Year2007            -0.1325     0.0425   -3.12  0.00184 **
## Year2008            -0.1530     0.0415   -3.68  0.00023 ***
## Year2009            -0.1560     0.0412   -3.79  0.00015 ***
## Year2010            -0.1680     0.0441   -3.81  0.00014 ***
## Year2011            -0.1620     0.0427   -3.79  0.00015 ***
## Year2012            -0.1723     0.0427   -4.04  5.5e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.591
## Multiple R-squared:  0.0125, Adjusted R-squared:  0.0101
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
##  654 weights are ~= 1. The remaining 6522 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.031   0.869   0.949   0.907   0.984   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.39e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##                   GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.011  1           1.005
## Year             1.011 16           1.000
```



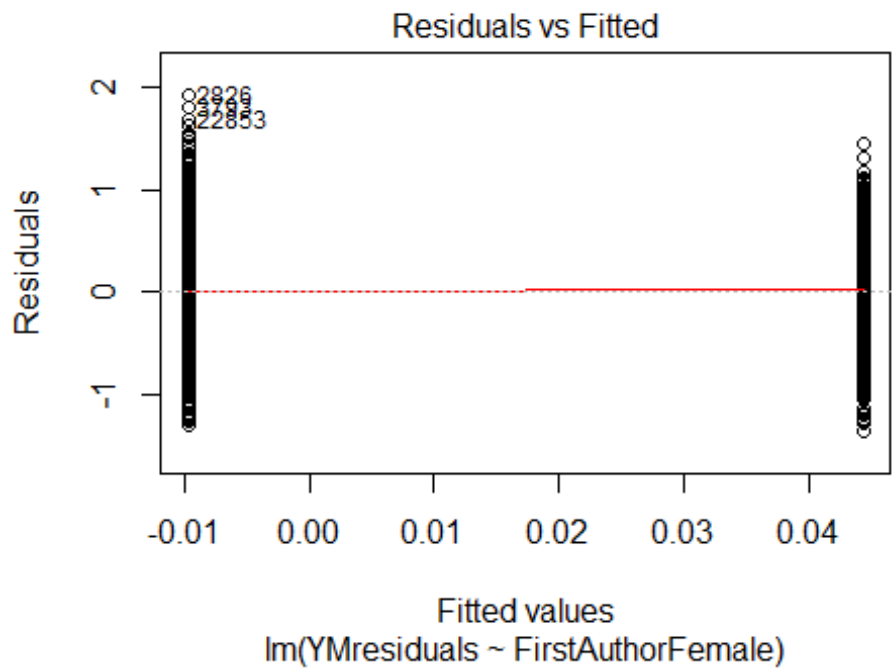

```
## [1] "List of  2 outliers with residuals above  2.5"
##          ScopusId  NLCS Year OneField Fields residuals
## 10123 11244344133 3.457 2004     2610      2     2.500
## 10177 41349117788 3.471 2004     2613      3     2.514
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.3192 -0.3927  0.0202  0.4036  2.5140
##
## Coefficients:
##                   Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.2572     0.0333   37.73  < 2e-16 ***
## LastAuthorFemale1   0.0620     0.0252    2.46  0.01402 *
## Year1997           -0.1285     0.0445   -2.89  0.00392 **
## Year1998           -0.1492     0.0448   -3.33  0.00087 ***
## Year1999           -0.1193     0.0482   -2.48  0.01326 *
## Year2000           -0.1011     0.0437   -2.31  0.02090 *
## Year2001           -0.2142     0.0555   -3.86  0.00011 ***
## Year2002           -0.0980     0.0439   -2.23  0.02578 *
## Year2003           -0.2716     0.0495   -5.49  4.1e-08 ***
## Year2004           -0.3002     0.0481   -6.24  4.6e-10 ***
## Year2005           -0.1380     0.0422   -3.27  0.00108 **
```

```
## Year2006           -0.0990     0.0432   -2.29  0.02196 *
## Year2007           -0.1340     0.0425   -3.15  0.00162 **
## Year2008           -0.1536     0.0415   -3.70  0.00021 ***
## Year2009           -0.1568     0.0412   -3.81  0.00014 ***
## Year2010           -0.1687     0.0441   -3.83  0.00013 ***
## Year2011           -0.1635     0.0427   -3.83  0.00013 ***
## Year2012           -0.1725     0.0427   -4.04  5.3e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.59
## Multiple R-squared:  0.0131, Adjusted R-squared:  0.0108
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
##  640 weights are ~= 1. The remaining 6536 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0302  0.8680  0.9500  0.9070  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.39e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Sample size for the above analysis:  7176"
## [1] ""
## [1] ""
## [1] "####################################"
## [1] "Analysis of AJSC 3110"
## [1] "####################################"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 1341 1285 1369 1319 1234 1114 1037 1104 1162 1236 1358 1307 1267 1130 1231
## 2011 2012
## 1128 1106
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  427  407  463  439  474  375  432  471  477  546  561  575  575  510  621
```

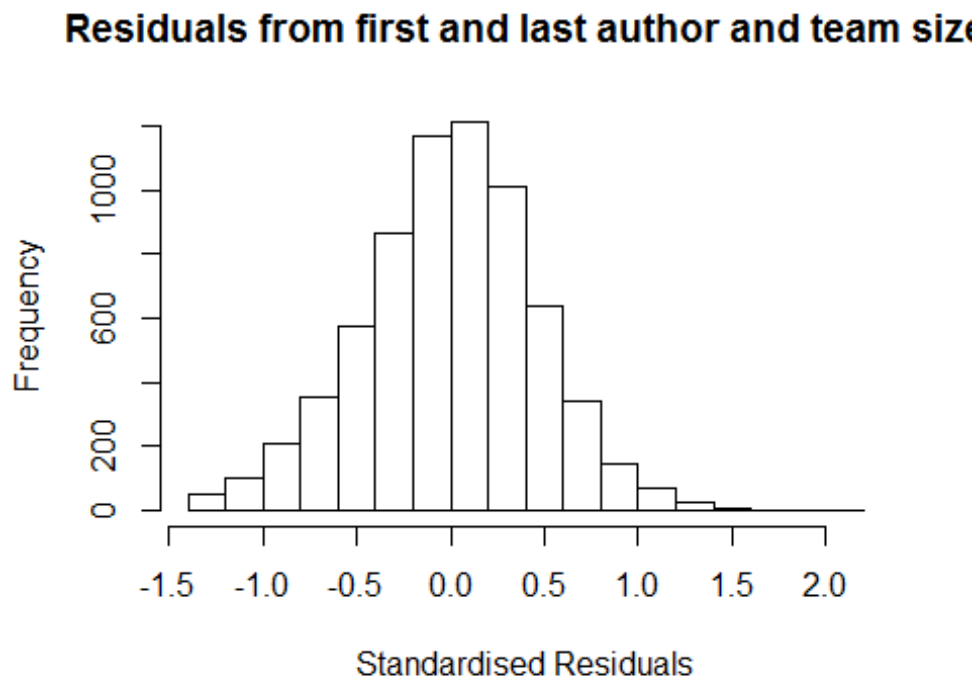
```
## 2011 2012
##  576  557
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  352  329  365  372  400  298  346  370  397  429  459  450  446  408  482
## 2011 2012
##  446  436
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
##  Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 90, df = 16, p-value = 3e-12
```




```
##
##  Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 17, df = 1, p-value = 3e-05
```

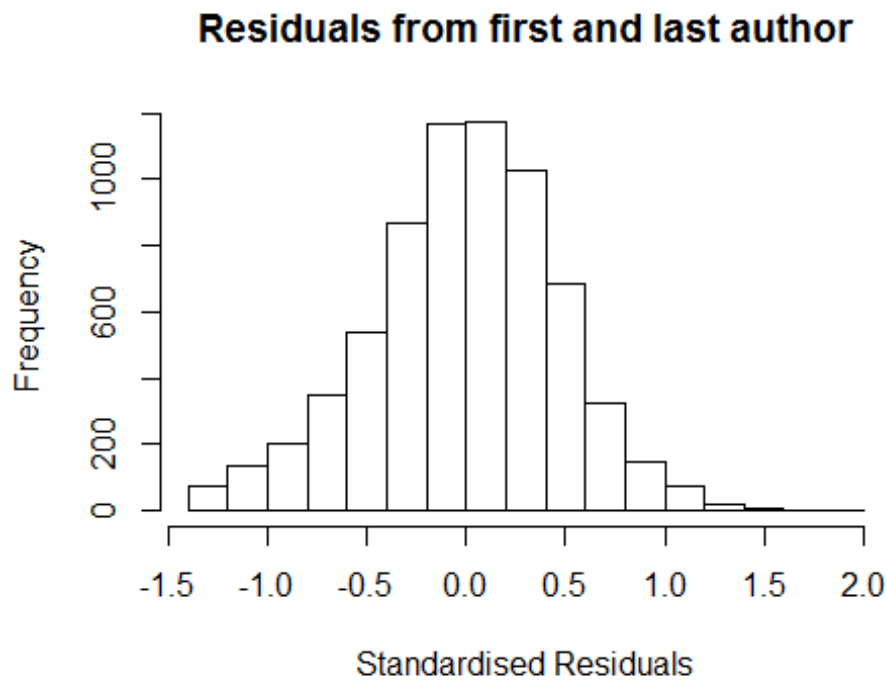



```
## [1] "Female first author team size 2018 geometric mean: 3.46090731729823"
## [1] "Male first author team size 2018 geometric mean: 3.70812604976103"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 7700, p-value = 0.3
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 3.69268623136181"
## [1] "Male last author team size 2018 geometric mean: 3.63842275827987"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 8200, p-value = 0.6
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##                     GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.038  1           1.019
## LastAuthorFemale  1.032  1           1.016
## UniqueAuthors     1.156  4           1.018
## Year              1.152 16           1.004
```



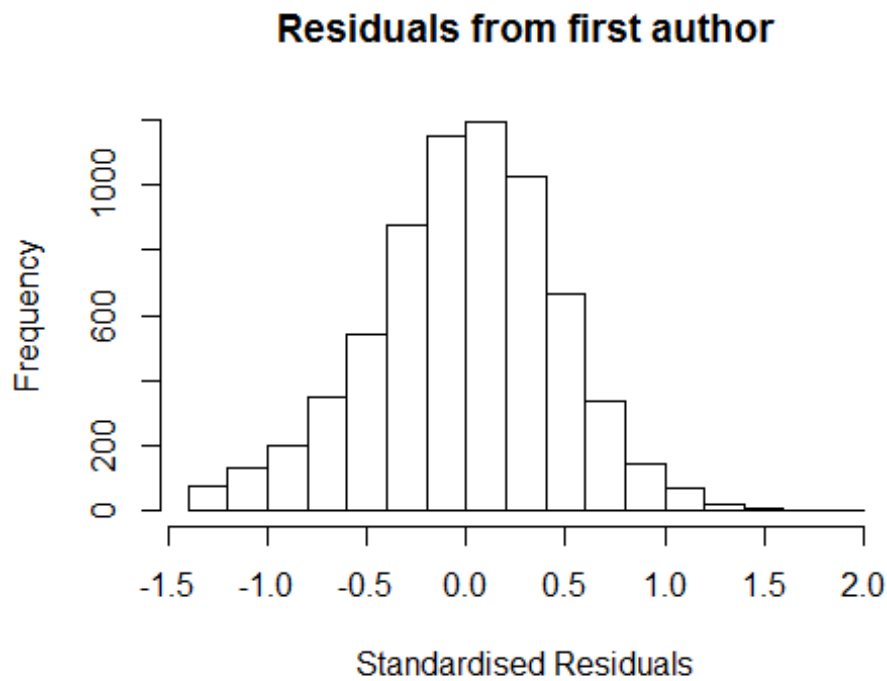

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
UniqueAuthors +
##     Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##     k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.3765 -0.2950  0.0105  0.2959  2.0972
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)          1.0946     0.0351   31.16  < 2e-16 ***
## FirstAuthorFemale1   0.0308     0.0140    2.21  0.02747 *
## LastAuthorFemale1    0.0313     0.0172    1.83  0.06790 .
## UniqueAuthors2       0.2119     0.0265    7.99  1.6e-15 ***
## UniqueAuthors3       0.2667     0.0269    9.91  < 2e-16 ***
## UniqueAuthors4       0.3022     0.0278   10.88  < 2e-16 ***
## UniqueAuthors5       0.3376     0.0277   12.20  < 2e-16 ***
## Year1997            -0.0737     0.0417   -1.77  0.07691 .
## Year1998            -0.0557     0.0421   -1.32  0.18555
## Year1999            -0.0828     0.0379   -2.19  0.02888 *
```

```
## Year2000            -0.1256     0.0385   -3.26  0.00113 **
## Year2001            -0.0820     0.0390   -2.10  0.03569 *
## Year2002            -0.0812     0.0380   -2.13  0.03290 *
## Year2003            -0.1233     0.0367   -3.36  0.00079 ***
## Year2004            -0.1249     0.0373   -3.35  0.00081 ***
## Year2005            -0.1597     0.0344   -4.64  3.6e-06 ***
## Year2006            -0.2174     0.0344   -6.32  2.8e-10 ***
## Year2007            -0.2211     0.0346   -6.39  1.8e-10 ***
## Year2008            -0.2254     0.0362   -6.23  4.8e-10 ***
## Year2009            -0.2006     0.0365   -5.49  4.1e-08 ***
## Year2010            -0.2152     0.0351   -6.13  9.2e-10 ***
## Year2011            -0.3087     0.0366   -8.43  < 2e-16 ***
## Year2012            -0.2719     0.0386   -7.05  2.0e-12 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.445
## Multiple R-squared:  0.0611, Adjusted R-squared:  0.058
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
##  observation 679 is an outlier with |weight| = 0 ( < 1.5e-05);
##  552 weights are ~= 1. The remaining 6232 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0152  0.8680  0.9510  0.9000  0.9860  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.47e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                    psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.014  1           1.007
## LastAuthorFemale  1.021  1           1.011
## Year              1.027 16           1.001
```



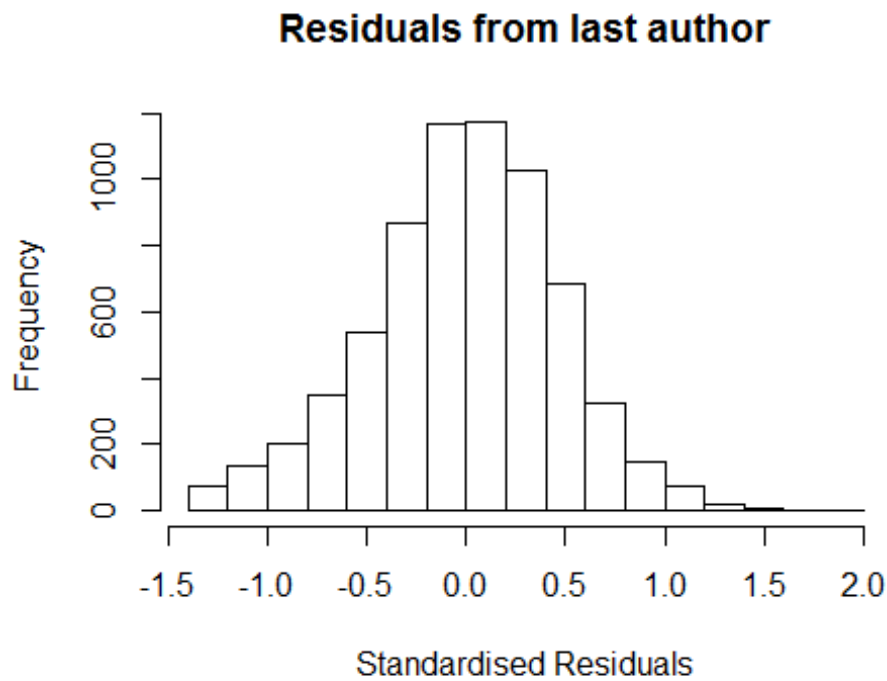

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.3345 -0.3015  0.0113  0.3063  1.8890
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)           1.2957     0.0286   45.31  < 2e-16 ***
## FirstAuthorFemale1    0.0540     0.0141    3.82  0.00013 ***
## LastAuthorFemale1     0.0388     0.0174    2.23  0.02571 *
## Year1997             -0.0667     0.0413   -1.61  0.10641
## Year1998             -0.0277     0.0409   -0.68  0.49800
## Year1999             -0.0611     0.0381   -1.60  0.10901
## Year2000             -0.1029     0.0389   -2.65  0.00808 **
## Year2001             -0.0594     0.0398   -1.49  0.13572
## Year2002             -0.0475     0.0379   -1.26  0.20934
## Year2003             -0.0770     0.0370   -2.08  0.03743 *
## Year2004             -0.0823     0.0374   -2.20  0.02788 *
## Year2005             -0.1192     0.0348   -3.43  0.00061 ***
```

```
## Year2006            -0.1750     0.0346   -5.06  4.3e-07 ***
## Year2007            -0.1685     0.0348   -4.83  1.4e-06 ***
## Year2008            -0.1715     0.0361   -4.75  2.0e-06 ***
## Year2009            -0.1486     0.0362   -4.11  4.0e-05 ***
## Year2010            -0.1567     0.0350   -4.47  7.8e-06 ***
## Year2011            -0.2467     0.0367   -6.72  1.9e-11 ***
## Year2012            -0.2052     0.0386   -5.32  1.1e-07 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.451
## Multiple R-squared:  0.0216, Adjusted R-squared:  0.019
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  532 weights are ~= 1. The remaining 6253 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0405  0.8660  0.9510  0.9000  0.9860  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.47e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##                   GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.01  1           1.005
## Year              1.01 16           1.000
```




```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##    Min     1Q Median     3Q    Max
## -1.299 -0.302  0.012  0.304  1.887
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)           1.2990     0.0285   45.54  < 2e-16 ***
## FirstAuthorFemale1    0.0583     0.0141    4.12  3.8e-05 ***
## Year1997             -0.0680     0.0413   -1.65  0.09944 .
## Year1998             -0.0294     0.0409   -0.72  0.47291
## Year1999             -0.0611     0.0381   -1.60  0.10909
## Year2000             -0.1015     0.0388   -2.61  0.00898 **
## Year2001             -0.0599     0.0399   -1.50  0.13296
## Year2002             -0.0467     0.0379   -1.23  0.21739
## Year2003             -0.0776     0.0370   -2.10  0.03619 *
## Year2004             -0.0822     0.0374   -2.20  0.02801 *
## Year2005             -0.1182     0.0348   -3.40  0.00068 ***
## Year2006             -0.1752     0.0346   -5.06  4.3e-07 ***
```

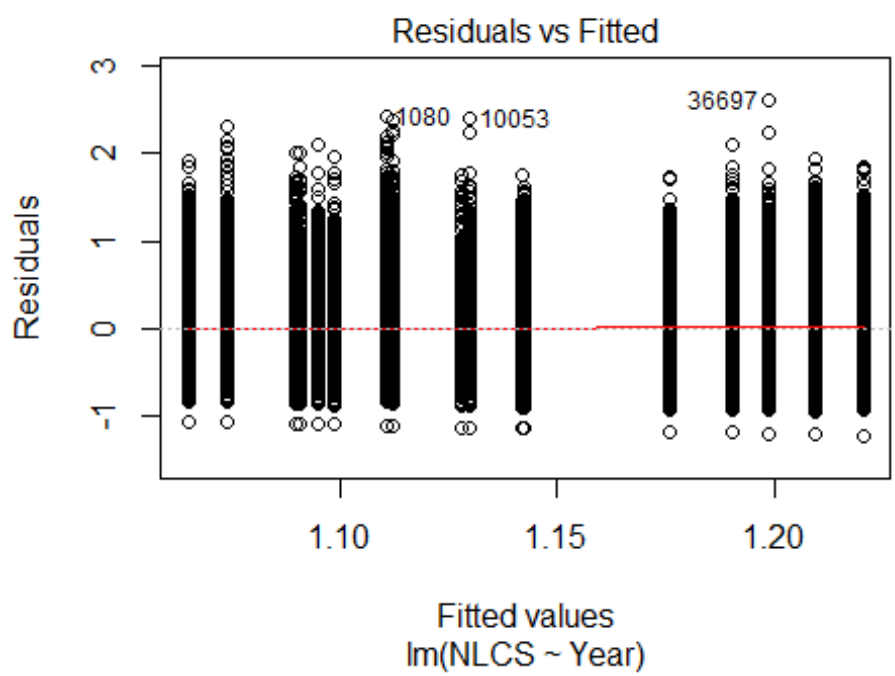
```
## Year2007            -0.1675     0.0349   -4.80  1.6e-06 ***
## Year2008            -0.1708     0.0361   -4.73  2.3e-06 ***
## Year2009            -0.1470     0.0361   -4.07  4.7e-05 ***
## Year2010            -0.1539     0.0350   -4.40  1.1e-05 ***
## Year2011            -0.2454     0.0367   -6.69  2.4e-11 ***
## Year2012            -0.2023     0.0386   -5.25  1.6e-07 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.452
## Multiple R-squared:  0.0209, Adjusted R-squared:  0.0184
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  543 weights are ~= 1. The remaining 6242 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0421  0.8650  0.9500  0.9000  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.47e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##                   GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.018  1           1.009
## Year             1.018 16           1.001
```



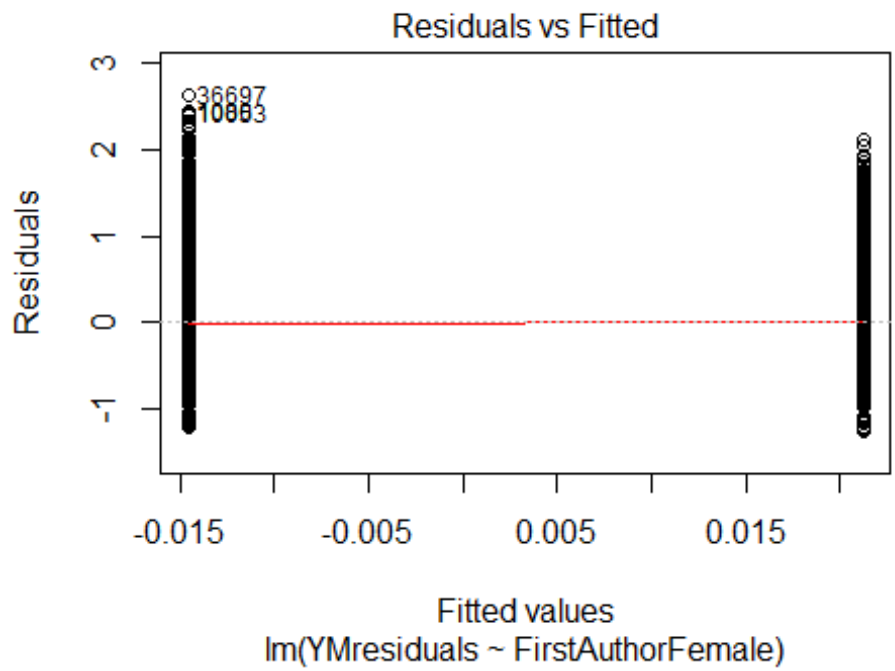

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.3497 -0.3010  0.0131  0.3043  1.8816
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)          1.3025     0.0286   45.58  < 2e-16 ***
## LastAuthorFemale1    0.0472     0.0174    2.71  0.00668 **
## Year1997            -0.0661     0.0416   -1.59  0.11170
## Year1998            -0.0259     0.0410   -0.63  0.52779
## Year1999            -0.0588     0.0382   -1.54  0.12373
## Year2000            -0.1027     0.0389   -2.64  0.00829 **
## Year2001            -0.0585     0.0399   -1.47  0.14281
## Year2002            -0.0456     0.0379   -1.20  0.22883
## Year2003            -0.0723     0.0371   -1.95  0.05121 .
## Year2004            -0.0794     0.0375   -2.12  0.03418 *
## Year2005            -0.1178     0.0348   -3.38  0.00072 ***
## Year2006            -0.1727     0.0346   -4.99  6.2e-07 ***
```

```
## Year2007           -0.1652     0.0349   -4.73  2.2e-06 ***
## Year2008           -0.1679     0.0361   -4.65  3.4e-06 ***
## Year2009           -0.1470     0.0362   -4.06  4.9e-05 ***
## Year2010           -0.1518     0.0351   -4.33  1.5e-05 ***
## Year2011           -0.2441     0.0367   -6.64  3.3e-11 ***
## Year2012           -0.2006     0.0387   -5.19  2.2e-07 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.452
## Multiple R-squared:  0.0197, Adjusted R-squared:  0.0172
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  542 weights are ~= 1. The remaining 6243 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0437  0.8660  0.9510  0.9000  0.9860  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.47e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Sample size for the above analysis:  6785"
## [1] ""
## [1] ""
## [1] "####################################"
## [1] "Analysis of AJSC 3200"
## [1] "####################################"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 2003 1771 1819 1550 1792 1982 1943 1703 2049 2044 2443 2706 2666 2199 2355
## 2011 2012
## 2298 2323
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 1724 1508 1510 1239 1332 1359 1687 1512 1798 1761 2122 2342 2325 1921 2039
## 2011 2012
```

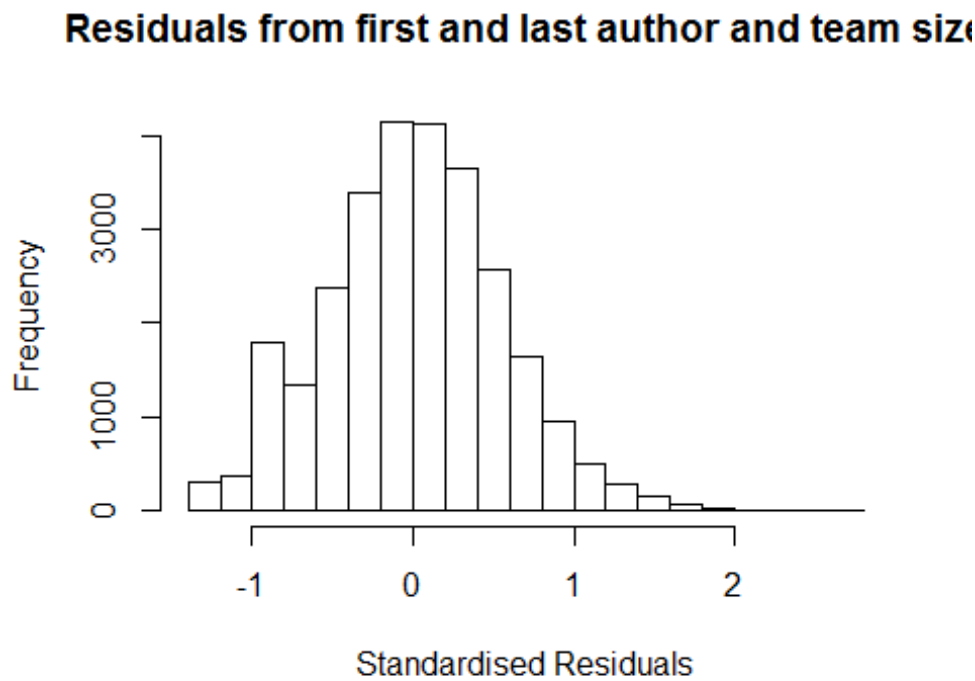
```
## 1961 2002
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 1606 1424 1393 1155 1240 1272 1558 1394 1637 1596 1947 2130 2127 1759 1857
## 2011 2012
## 1787 1832
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
##  Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 330, df = 16, p-value <2e-16
```




```
##
##  Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 95, df = 1, p-value <2e-16
```



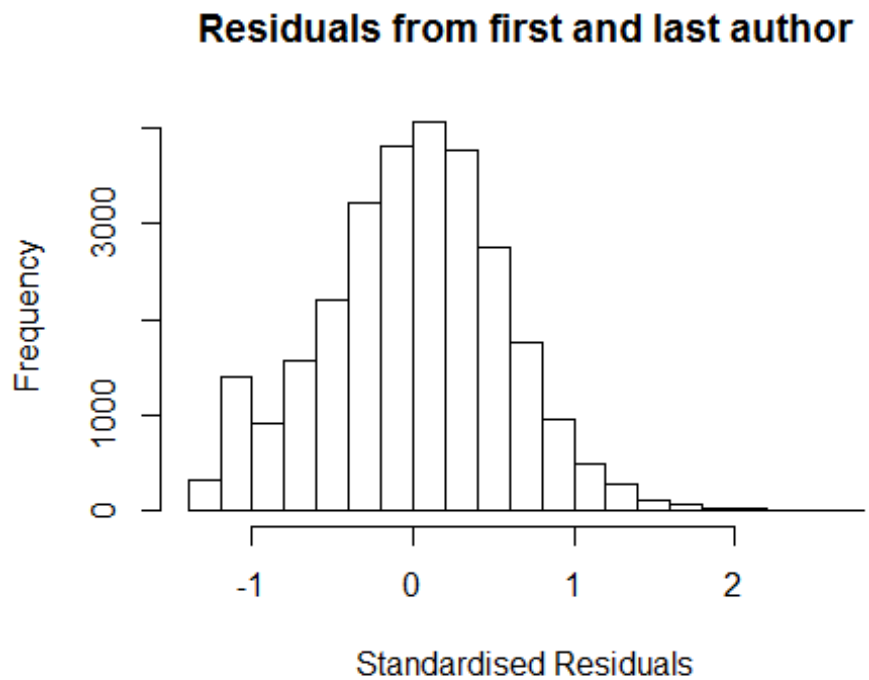

```
## [1] "Female first author team size 2018 geometric mean: 2.98408022498213"
## [1] "Male first author team size 2018 geometric mean: 2.53080591013358"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 560000, p-value = 5e-09
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 2.83522015426535"
## [1] "Male last author team size 2018 geometric mean: 2.69712178749186"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 510000, p-value = 0.08
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##                     GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.102  1           1.050
## LastAuthorFemale  1.081  1           1.040
## UniqueAuthors     1.066  4           1.008
## Year              1.063 16           1.002
```




```
## [1] "List of  4 outliers with residuals above  2.5"
##          ScopusId  NLCS Year OneField Fields residuals
## 1080   0030194684 3.543 1996     3200      1     2.654
## 3142   0030638716 3.482 1997     3200      1     2.574
## 5561   0032215040 3.376 1998     3200      1     2.516
## 36697 79960392344 3.807 2011     3200      1     2.550
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
UniqueAuthors +
##     Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##     k.max = 1000))
##  \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.39353 -0.35288  0.00743  0.35804  2.65366
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)          0.88934    0.01755   50.67   <2e-16 ***
## FirstAuthorFemale1   0.01790    0.00700    2.56   0.0105 *
## LastAuthorFemale1   -0.01092    0.00705   -1.55   0.1216
## UniqueAuthors2       0.26751    0.00956   27.98   <2e-16 ***
## UniqueAuthors3       0.32989    0.01033   31.92   <2e-16 ***
## UniqueAuthors4       0.36628    0.01156   31.68   <2e-16 ***
## UniqueAuthors5       0.41234    0.01100   37.50   <2e-16 ***
```

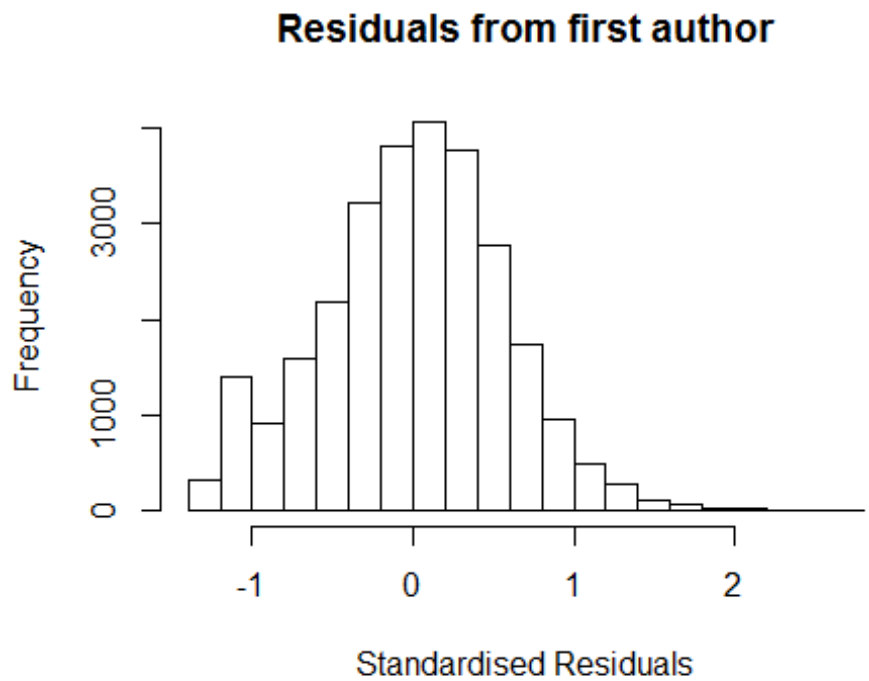
```
## Year1997            0.01912    0.02305    0.83   0.4069
## Year1998           -0.02889    0.02273   -1.27   0.2038
## Year1999            0.04132    0.02265    1.82   0.0681 .
## Year2000            0.02464    0.02261    1.09   0.2757
## Year2001           -0.00922    0.02297   -0.40   0.6882
## Year2002           -0.01809    0.02217   -0.82   0.4144
## Year2003           -0.06340    0.02290   -2.77   0.0056 **
## Year2004           -0.04416    0.02136   -2.07   0.0387 *
## Year2005           -0.04168    0.02076   -2.01   0.0447 *
## Year2006           -0.00993    0.02005   -0.50   0.6202
## Year2007           -0.00436    0.02014   -0.22   0.8285
## Year2008            0.02279    0.01952    1.17   0.2430
## Year2009            0.06660    0.02141    3.11   0.0019 **
## Year2010            0.05263    0.02116    2.49   0.0129 *
## Year2011            0.03787    0.02121    1.79   0.0742 .
## Year2012            0.07395    0.02206    3.35   0.0008 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.528
## Multiple R-squared:  0.0825, Adjusted R-squared:  0.0818
## Convergence in 10 IRWLS iterations
##
## Robustness weights:
##  4 observations c(762,2180,3669,24684)
##   are outliers with |weight| = 0 ( < 3.6e-06);
##  2363 weights are ~= 1. The remaining 25347 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0001  0.8670  0.9510  0.9020  0.9860  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          3.61e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.076  1           1.037
```

```
## LastAuthorFemale  1.065  1           1.032
## Year              1.019 16           1.001
```



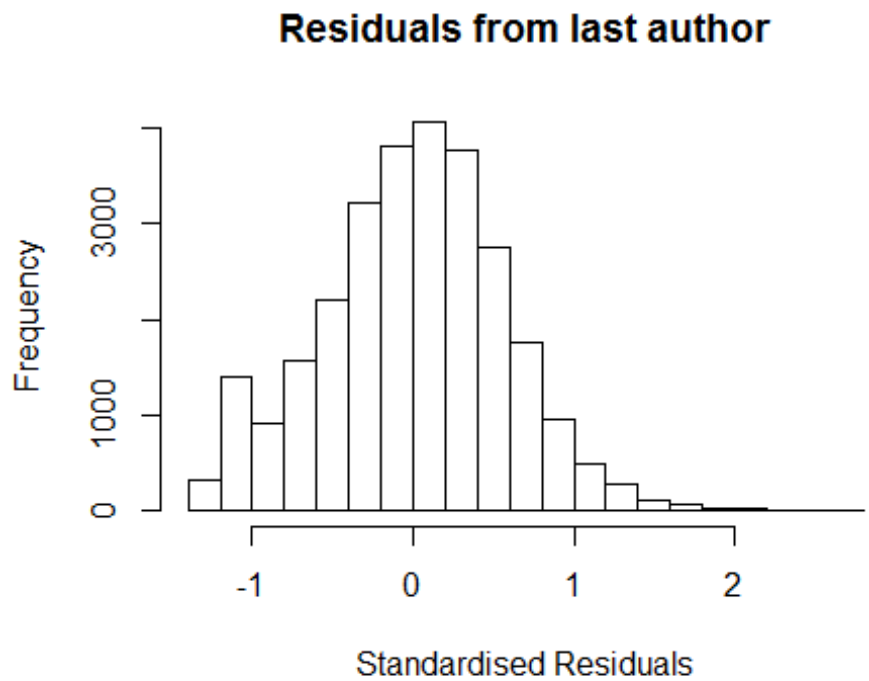

```
## [1] "List of  1 outliers with residuals above  2.5"
##          ScopusId  NLCS Year OneField Fields residuals
## 36697 79960392344 3.807 2011     3200      1     2.638
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.2548 -0.3654  0.0204  0.3730  2.6382
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.06e+00   1.74e-02   61.16  < 2e-16 ***
## FirstAuthorFemale1  4.68e-02   7.19e-03    6.50  8.0e-11 ***
## LastAuthorFemale1   2.44e-03   7.28e-03    0.33   0.7377
## Year1997            1.37e-02   2.38e-02    0.57   0.5655
## Year1998           -2.71e-02   2.40e-02   -1.13   0.2577
## Year1999            4.68e-02   2.33e-02    2.00   0.0450 *
## Year2000            3.40e-02   2.32e-02    1.47   0.1425
## Year2001           -9.96e-05   2.39e-02    0.00   0.9967
## Year2002            4.31e-03   2.33e-02    0.19   0.8531
```

```
## Year2003           -2.98e-02   2.40e-02   -1.24   0.2142
## Year2004            5.29e-03   2.23e-02    0.24   0.8125
## Year2005            1.51e-02   2.16e-02    0.70   0.4839
## Year2006            5.69e-02   2.07e-02    2.75   0.0059 **
## Year2007            6.01e-02   2.07e-02    2.90   0.0037 **
## Year2008            9.02e-02   2.01e-02    4.49  7.1e-06 ***
## Year2009            1.23e-01   2.22e-02    5.54  3.1e-08 ***
## Year2010            1.05e-01   2.18e-02    4.84  1.3e-06 ***
## Year2011            1.05e-01   2.19e-02    4.80  1.6e-06 ***
## Year2012            1.42e-01   2.27e-02    6.25  4.2e-10 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.548
## Multiple R-squared:  0.0107, Adjusted R-squared:  0.01
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
##  observation 24684 is an outlier with |weight| = 0 ( < 3.6e-06);
##  2302 weights are ~= 1. The remaining 25411 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0045  0.8650  0.9510  0.9020  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          3.61e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##                     GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.016  1           1.008
## Year              1.016 16           1.000
```




```
## [1] "List of  1 outliers with residuals above  2.5"
##          ScopusId  NLCS Year OneField Fields residuals
## 36697 79960392344 3.807 2011     3200      1     2.638
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.2539 -0.3649  0.0201  0.3730  2.6376
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.06e+00   1.73e-02   61.62  < 2e-16 ***
## FirstAuthorFemale1  4.76e-02   7.03e-03    6.77  1.3e-11 ***
## Year1997            1.37e-02   2.39e-02    0.57   0.5664
## Year1998           -2.71e-02   2.40e-02   -1.13   0.2576
## Year1999            4.68e-02   2.33e-02    2.01   0.0448 *
## Year2000            3.40e-02   2.32e-02    1.47   0.1423
## Year2001           -2.93e-05   2.39e-02    0.00   0.9990
## Year2002            4.32e-03   2.33e-02    0.19   0.8525
## Year2003           -2.98e-02   2.40e-02   -1.24   0.2150
## Year2004            5.33e-03   2.23e-02    0.24   0.8111
## Year2005            1.51e-02   2.16e-02    0.70   0.4832
## Year2006            5.70e-02   2.07e-02    2.75   0.0059 **
```

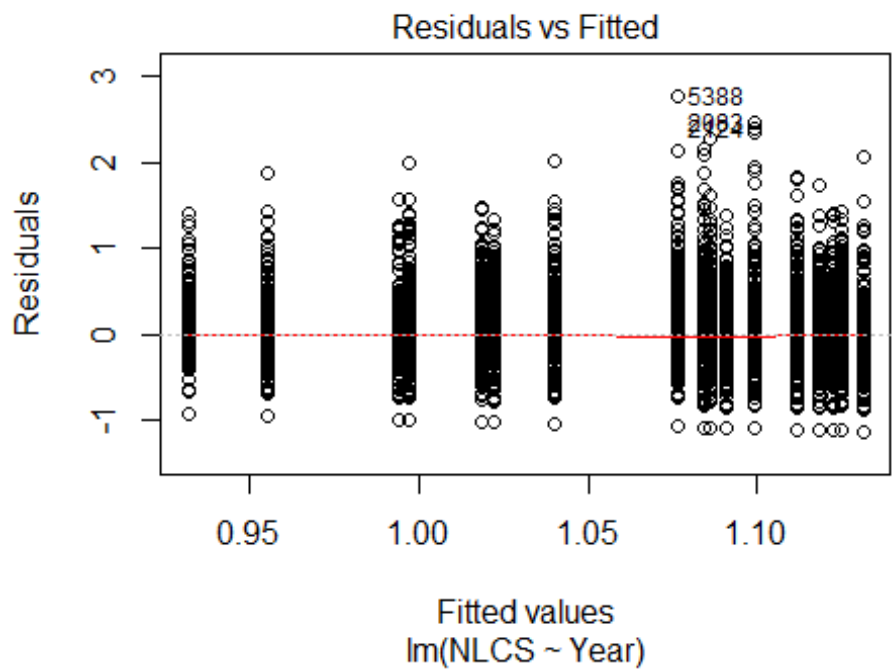
```
## Year2007            6.01e-02   2.07e-02    2.90   0.0037 **
## Year2008            9.02e-02   2.01e-02    4.49  7.1e-06 ***
## Year2009            1.23e-01   2.22e-02    5.54  3.0e-08 ***
## Year2010            1.06e-01   2.18e-02    4.84  1.3e-06 ***
## Year2011            1.05e-01   2.19e-02    4.80  1.6e-06 ***
## Year2012            1.42e-01   2.27e-02    6.26  4.0e-10 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.548
## Multiple R-squared:  0.0107, Adjusted R-squared:  0.01
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
##  observation 24684 is an outlier with |weight| = 0 ( < 3.6e-06);
##  2294 weights are ~= 1. The remaining 25419 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0046  0.8650  0.9510  0.9020  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          3.61e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                    psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##                   GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.006  1           1.003
## Year             1.006 16           1.000
```



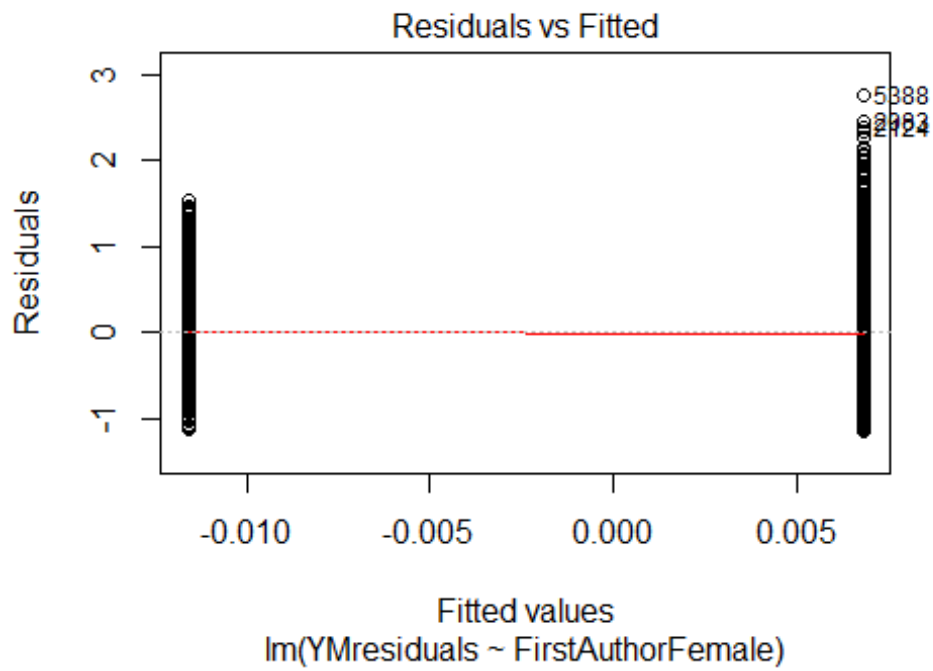

```
## [1] "List of  1 outliers with residuals above  2.5"
##          ScopusId  NLCS Year OneField Fields residuals
## 36697 79960392344 3.807 2011     3200      1     2.638
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.2396 -0.3643  0.0228  0.3726  2.6222
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)        1.075369   0.017245   62.36  < 2e-16 ***
## LastAuthorFemale1  0.018470   0.007108    2.60   0.0094 **
## Year1997           0.012945   0.023820    0.54   0.5868
## Year1998          -0.027917   0.023925   -1.17   0.2433
## Year1999           0.044804   0.023365    1.92   0.0552 .
## Year2000           0.033621   0.023203    1.45   0.1474
## Year2001           0.000879   0.023896    0.04   0.9707
## Year2002           0.004314   0.023268    0.19   0.8529
## Year2003          -0.029456   0.024025   -1.23   0.2202
## Year2004           0.007611   0.022304    0.34   0.7329
## Year2005           0.016622   0.021597    0.77   0.4415
## Year2006           0.059208   0.020686    2.86   0.0042 **
```

```
## Year2007           0.062321   0.020709    3.01   0.0026 **
## Year2008           0.094031   0.020068    4.69  2.8e-06 ***
## Year2009           0.125068   0.022202    5.63  1.8e-08 ***
## Year2010           0.108038   0.021776    4.96  7.0e-07 ***
## Year2011           0.109424   0.021876    5.00  5.7e-07 ***
## Year2012           0.145773   0.022673    6.43  1.3e-10 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.549
## Multiple R-squared:  0.00922,    Adjusted R-squared:  0.00861
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
##  observation 24684 is an outlier with |weight| = 0 ( < 3.6e-06);
##  2358 weights are ~= 1. The remaining 25355 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0061  0.8640  0.9500  0.9020  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          3.61e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Sample size for the above analysis:  27714"
## [1] ""
## [1] ""
## [1] "####################################"
## [1] "Analysis of AJSC 3201"
## [1] "####################################"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  297  291  300  306  270  259  290  192  185  209  219  283  288  338  370
## 2011 2012
##  355  347
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  261  253  246  238  204  188  233  172  167  174  197  251  241  294  307
```

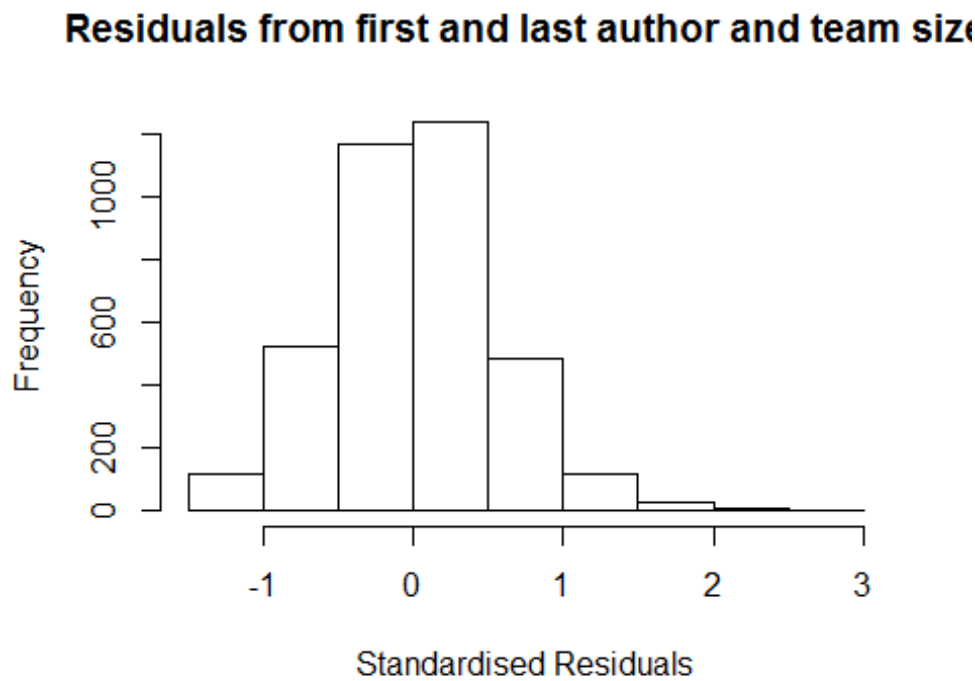
```
## 2011 2012
##  289  280
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  248  239  229  227  192  175  208  160  154  158  178  222  215  267  281
## 2011 2012
##  257  253
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
##  Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 71, df = 16, p-value = 6e-09
```




```
##
##  Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 42, df = 1, p-value = 1e-10
```

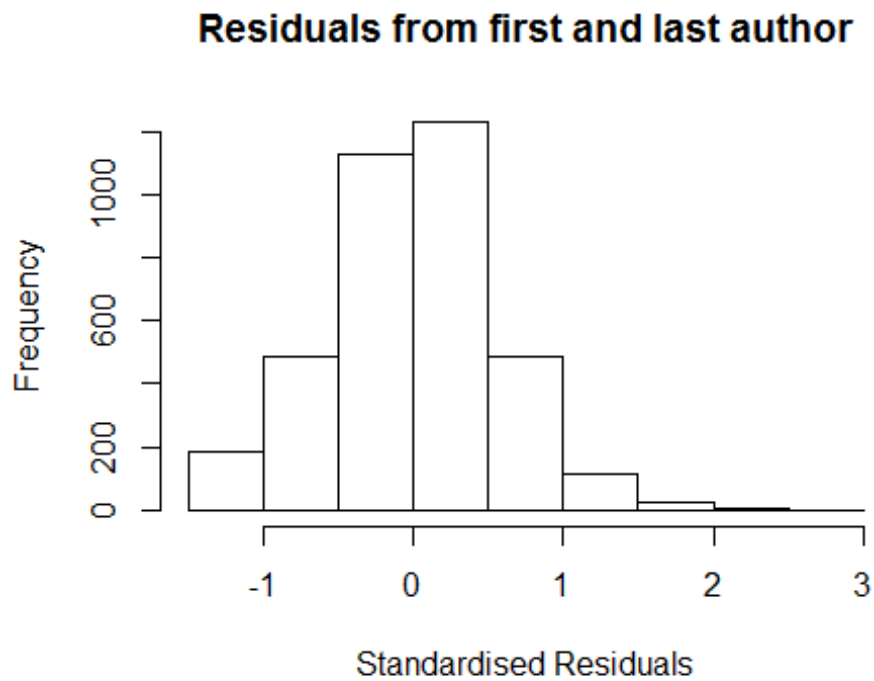



```
## [1] "Female first author team size 2018 geometric mean: 2.71969388948289"
## [1] "Male first author team size 2018 geometric mean: 2.35968214127431"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 8900, p-value = 0.08
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 2.72888581686227"
## [1] "Male last author team size 2018 geometric mean: 2.39992040315389"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 8700, p-value = 0.07
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.232  1           1.110
## LastAuthorFemale  1.232  1           1.110
## UniqueAuthors     1.140  4           1.017
## Year              1.142 16           1.004
```



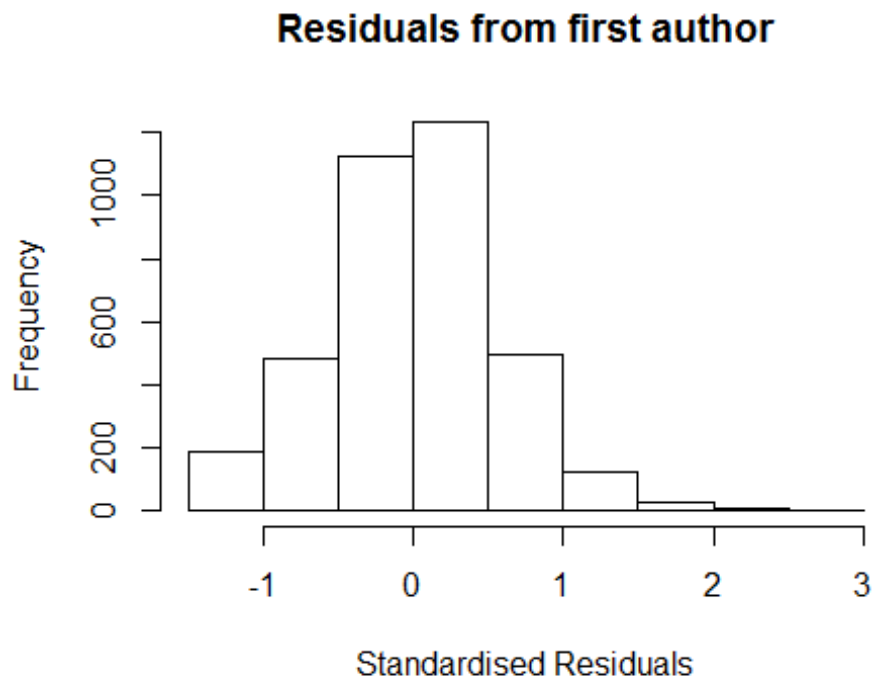

```
## [1] "List of  1 outliers with residuals above  2.5"
##         ScopusId  NLCS Year OneField Fields residuals
## 5388 84857285608 3.843 2012     1213      3     2.783
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
UniqueAuthors +
##     Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##     k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.2597 -0.3492  0.0102  0.3501  2.7832
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         0.96965    0.04207   23.05  < 2e-16 ***
## FirstAuthorFemale1 -0.00643    0.02042   -0.31    0.753
## LastAuthorFemale1  -0.03178    0.02050   -1.55    0.121
## UniqueAuthors2      0.14100    0.02343    6.02  1.9e-09 ***
## UniqueAuthors3      0.16897    0.02668    6.33  2.7e-10 ***
## UniqueAuthors4      0.22388    0.03455    6.48  1.0e-10 ***
## UniqueAuthors5      0.20499    0.03377    6.07  1.4e-09 ***
## Year1997           -0.02731    0.05607   -0.49    0.626
## Year1998           -0.07728    0.05604   -1.38    0.168
## Year1999           -0.10391    0.05326   -1.95    0.051 .
```

```
## Year2000           -0.01322    0.05844   -0.23    0.821
## Year2001           -0.10210    0.05729   -1.78    0.075 .
## Year2002           -0.01065    0.05870   -0.18    0.856
## Year2003           -0.06792    0.05708   -1.19    0.234
## Year2004            0.04540    0.05815    0.78    0.435
## Year2005           -0.00115    0.05635   -0.02    0.984
## Year2006            0.04400    0.05835    0.75    0.451
## Year2007            0.03035    0.05450    0.56    0.578
## Year2008            0.04485    0.05177    0.87    0.386
## Year2009            0.08509    0.05295    1.61    0.108
## Year2010           -0.03602    0.05178   -0.70    0.487
## Year2011           -0.00734    0.05110   -0.14    0.886
## Year2012           -0.05086    0.05864   -0.87    0.386
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.526
## Multiple R-squared:  0.0363, Adjusted R-squared:  0.0305
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  2 observations c(1372,3656)
##   are outliers with |weight| <= 2.5e-05 ( < 2.7e-05);
##  325 weights are ~= 1. The remaining 3336 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0109  0.8650  0.9500  0.9000  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          2.73e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.204  1           1.097
## LastAuthorFemale  1.208  1           1.099
## Year              1.053 16           1.002
```



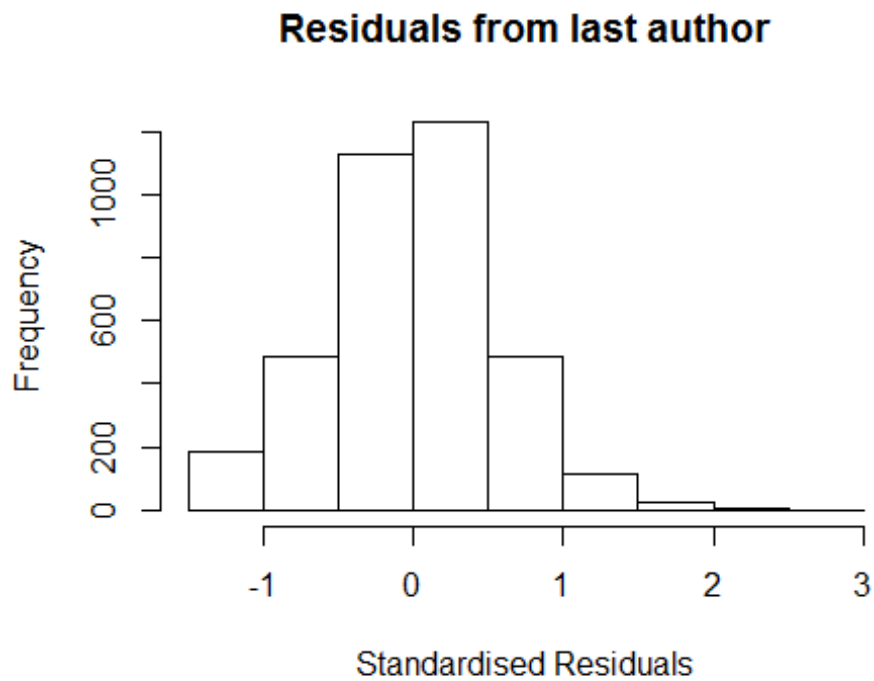

```
## [1] "List of  2 outliers with residuals above  2.5"
##         ScopusId  NLCS Year OneField Fields residuals
## 2083  0036595423 3.556 2002     3201      1     2.509
## 5388 84857285608 3.843 2012     1213      3     2.821
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.1675 -0.3520  0.0153  0.3502  2.8206
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)          1.04268    0.04118   25.32   <2e-16 ***
## FirstAuthorFemale1   0.01290    0.02048    0.63    0.529
## LastAuthorFemale1   -0.01489    0.02060   -0.72    0.470
## Year1997            -0.02451    0.05550   -0.44    0.659
## Year1998            -0.06860    0.05674   -1.21    0.227
## Year1999            -0.10277    0.05361   -1.92    0.055 .
## Year2000            -0.01783    0.05762   -0.31    0.757
## Year2001            -0.10944    0.05752   -1.90    0.057 .
## Year2002             0.00409    0.05933    0.07    0.945
## Year2003            -0.06552    0.05806   -1.13    0.259
## Year2004             0.07381    0.05930    1.24    0.213
```

```
## Year2005            0.00990    0.05659    0.17    0.861
## Year2006            0.07295    0.05917    1.23    0.218
## Year2007            0.06585    0.05414    1.22    0.224
## Year2008            0.07186    0.05177    1.39    0.165
## Year2009            0.11192    0.05343    2.09    0.036 *
## Year2010           -0.02369    0.05234   -0.45    0.651
## Year2011            0.02516    0.05144    0.49    0.625
## Year2012           -0.02032    0.05986   -0.34    0.734
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.533
## Multiple R-squared:  0.0131, Adjusted R-squared:  0.00827
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  2 observations c(1372,3656) are outliers with |weight| = 0 ( < 2.7e-05);
##  288 weights are ~= 1. The remaining 3373 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0015  0.8640  0.9520  0.9020  0.9860  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          2.73e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.027  1           1.014
## Year              1.027 16           1.001
```




```
## [1] "List of  2 outliers with residuals above  2.5"
##         ScopusId  NLCS Year OneField Fields residuals
## 2083  0036595423 3.556 2002     3201      1     2.509
## 5388 84857285608 3.843 2012     1213      3     2.821
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.1580 -0.3528  0.0146  0.3521  2.8241
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)          1.03976    0.04108   25.31   <2e-16 ***
## FirstAuthorFemale1   0.00651    0.01895    0.34    0.731
## Year1997            -0.02420    0.05556   -0.44    0.663
## Year1998            -0.06849    0.05679   -1.21    0.228
## Year1999            -0.10282    0.05366   -1.92    0.055 .
## Year2000            -0.01859    0.05758   -0.32    0.747
## Year2001            -0.10998    0.05751   -1.91    0.056 .
## Year2002             0.00294    0.05934    0.05    0.960
## Year2003            -0.06572    0.05819   -1.13    0.259
## Year2004             0.07400    0.05933    1.25    0.212
## Year2005             0.01007    0.05665    0.18    0.859
```

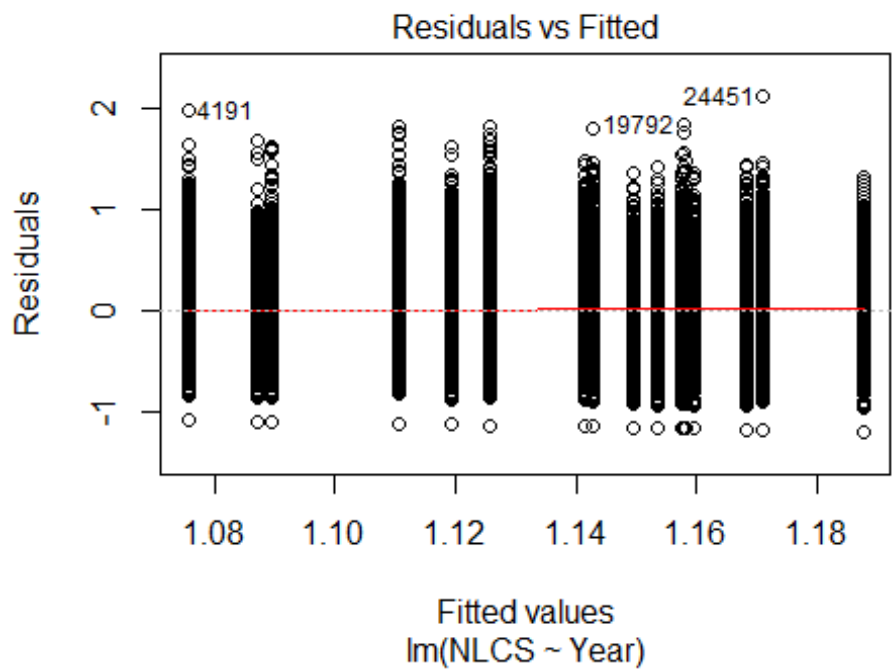
```
## Year2006            0.07271    0.05922    1.23    0.220
## Year2007            0.06564    0.05419    1.21    0.226
## Year2008            0.07148    0.05180    1.38    0.168
## Year2009            0.11173    0.05348    2.09    0.037 *
## Year2010           -0.02407    0.05236   -0.46    0.646
## Year2011            0.02494    0.05152    0.48    0.628
## Year2012           -0.02087    0.05988   -0.35    0.728
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.533
## Multiple R-squared:  0.013,  Adjusted R-squared:  0.0084
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  2 observations c(1372,3656) are outliers with |weight| = 0 ( < 2.7e-05);
##  296 weights are ~= 1. The remaining 3365 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0012  0.8640  0.9520  0.9020  0.9860  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          2.73e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##                  GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.03  1           1.015
## Year             1.03 16           1.001
```



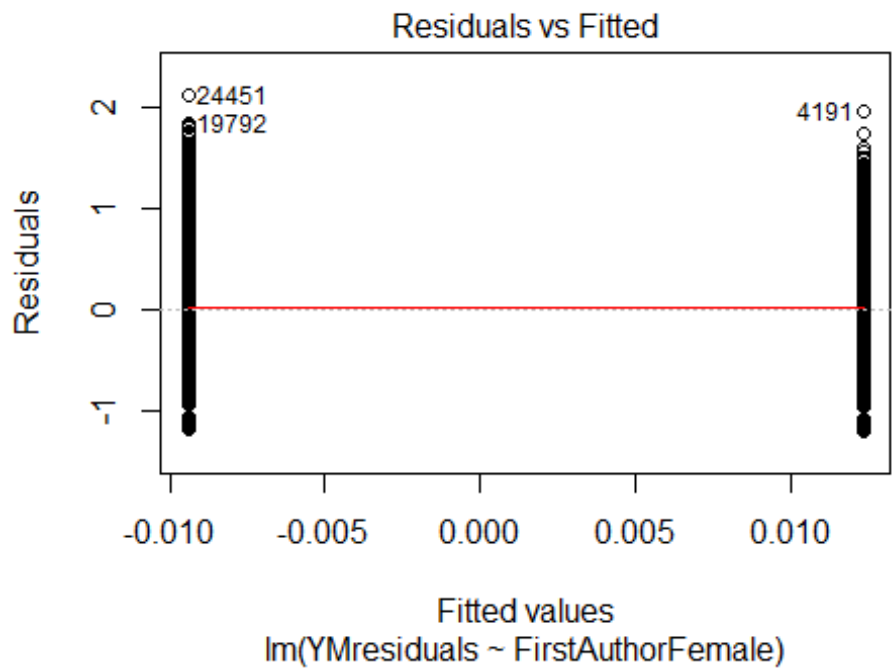

```
## [1] "List of  2 outliers with residuals above  2.5"
##         ScopusId  NLCS Year OneField Fields residuals
## 2083  0036595423 3.556 2002     3201      1     2.509
## 5388 84857285608 3.843 2012     1213      3     2.821
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##    Min     1Q Median     3Q    Max
## -1.158 -0.352  0.015  0.354  2.818
##
## Coefficients:
##                   Estimate Std. Error t value Pr(>|t|)
## (Intercept)        1.04499    0.04098   25.50   <2e-16 ***
## LastAuthorFemale1 -0.00937    0.01904   -0.49    0.623
## Year1997          -0.02399    0.05547   -0.43    0.665
## Year1998          -0.06835    0.05672   -1.21    0.228
## Year1999          -0.10300    0.05364   -1.92    0.055 .
## Year2000          -0.01672    0.05758   -0.29    0.772
## Year2001          -0.10974    0.05750   -1.91    0.056 .
## Year2002           0.00510    0.05927    0.09    0.931
## Year2003          -0.06394    0.05794   -1.10    0.270
## Year2004           0.07400    0.05926    1.25    0.212
## Year2005           0.01057    0.05660    0.19    0.852
```

```
## Year2006           0.07420    0.05913    1.25    0.210
## Year2007           0.06582    0.05413    1.22    0.224
## Year2008           0.07298    0.05168    1.41    0.158
## Year2009           0.11282    0.05344    2.11    0.035 *
## Year2010          -0.02317    0.05235   -0.44    0.658
## Year2011           0.02599    0.05148    0.50    0.614
## Year2012          -0.01966    0.05987   -0.33    0.743
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.533
## Multiple R-squared:  0.013,  Adjusted R-squared:  0.00842
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  2 observations c(1372,3656) are outliers with |weight| = 0 ( < 2.7e-05);
##  287 weights are ~= 1. The remaining 3374 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0019  0.8640  0.9530  0.9020  0.9860  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          2.73e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                  psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Sample size for the above analysis:  3663"
## [1] ""
## [1] ""
## [1] "##################################"
## [1] "Analysis of AJSC 3202"
## [1] "##################################"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 1135 1047 1047 1019 1437 1567 1469 1086 1312 1340 1532 1704 1814 2052 2192
## 2011 2012
## 2405 2431
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
```

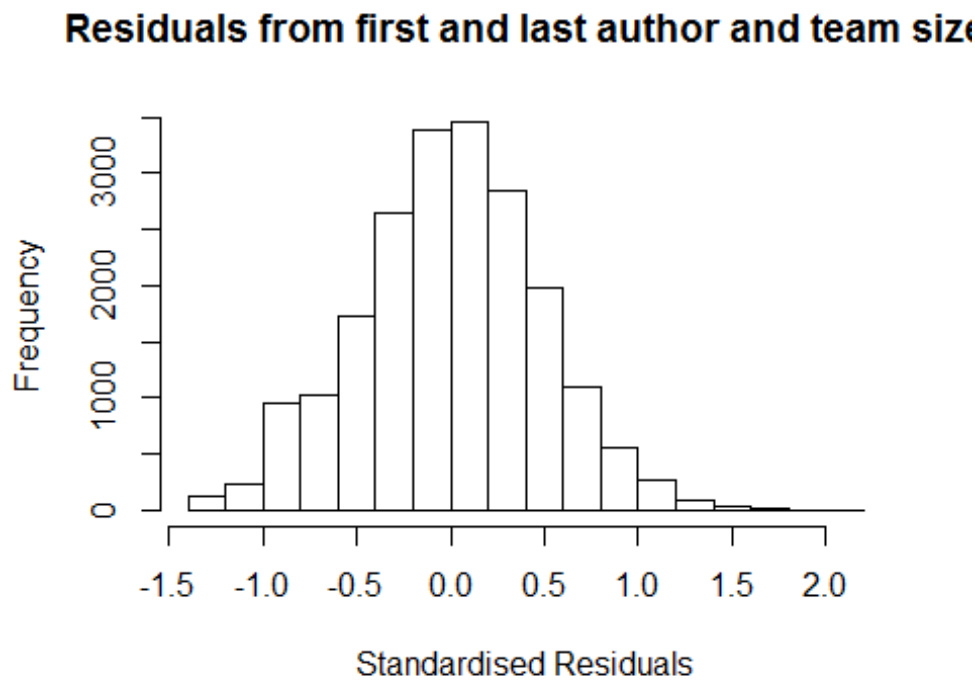
```
##  987  915  892  880 1110 1160 1281  965 1140 1185 1335 1485 1569 1728 1850
## 2011 2012
## 2030 2047
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  925  861  831  809 1000 1077 1153  863  993 1081 1213 1366 1403 1570 1654
## 2011 2012
## 1834 1835
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
##  Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 150, df = 16, p-value <2e-16
```




```
##
##  Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 55, df = 1, p-value = 1e-13
```

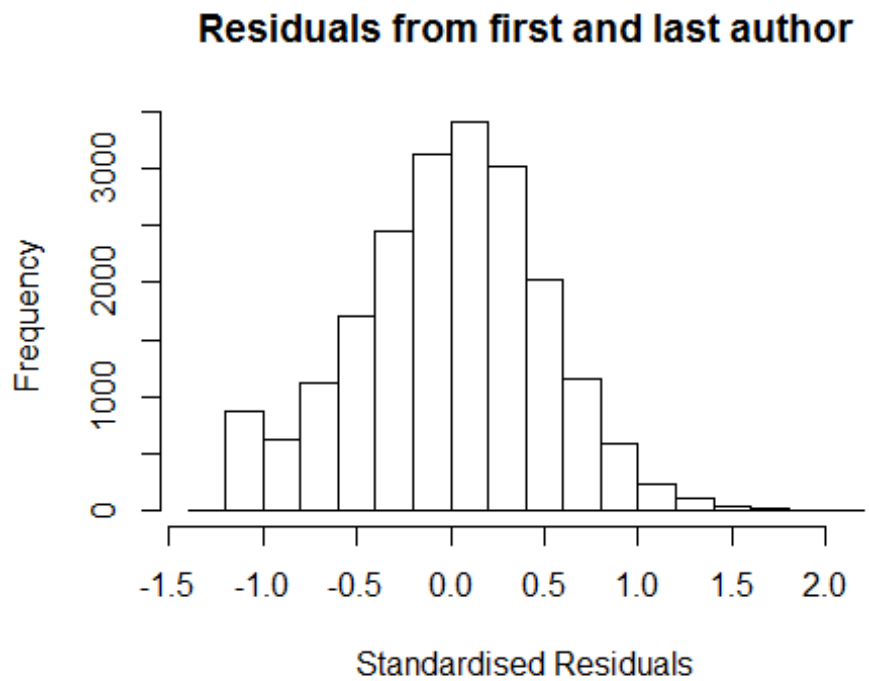



```
## [1] "Female first author team size 2018 geometric mean: 3.18199167588844"
## [1] "Male first author team size 2018 geometric mean: 2.69477451616669"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 520000, p-value = 1e-07
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 3.04706117147472"
## [1] "Male last author team size 2018 geometric mean: 2.8735569644023"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 480000, p-value = 0.1
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.102  1           1.050
## LastAuthorFemale  1.092  1           1.045
## UniqueAuthors     1.054  4           1.007
## Year              1.064 16           1.002
```



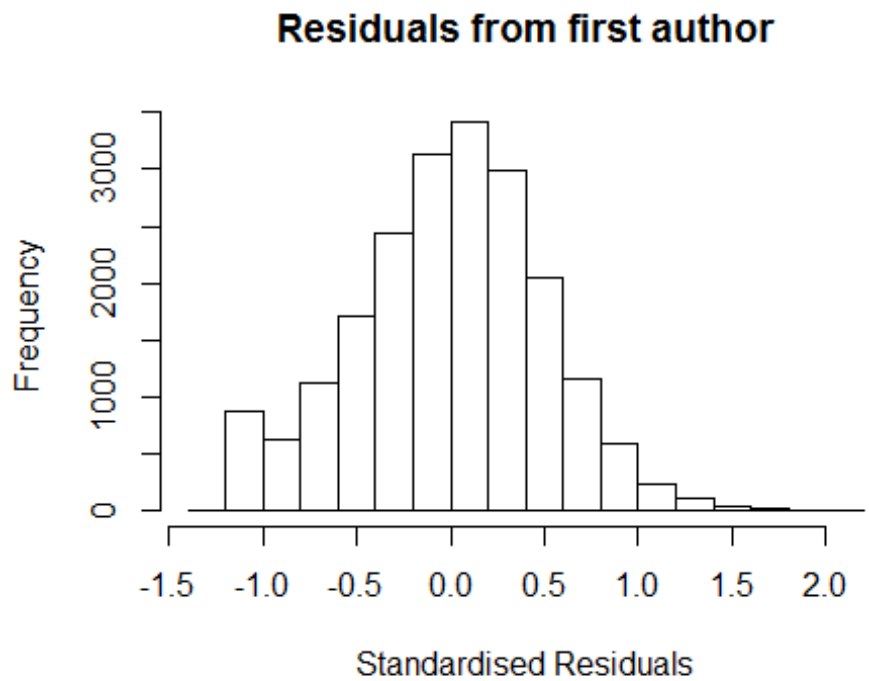

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
UniqueAuthors +
##     Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##     k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.3444 -0.3162  0.0067  0.3183  2.1818
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         0.92342    0.01866   49.50  < 2e-16 ***
## FirstAuthorFemale1  0.00365    0.00725    0.50   0.6144
## LastAuthorFemale1  -0.00165    0.00727   -0.23   0.8205
## UniqueAuthors2      0.25898    0.01028   25.20  < 2e-16 ***
## UniqueAuthors3      0.33131    0.01094   30.28  < 2e-16 ***
## UniqueAuthors4      0.35575    0.01251   28.43  < 2e-16 ***
## UniqueAuthors5      0.39311    0.01162   33.84  < 2e-16 ***
## Year1997           -0.04498    0.02461   -1.83   0.0676 .
## Year1998           -0.05220    0.02545   -2.05   0.0403 *
## Year1999           -0.04145    0.02388   -1.74   0.0826 .
```

```
## Year2000            0.01132    0.02292    0.49   0.6212
## Year2001            0.02584    0.02371    1.09   0.2756
## Year2002            0.01728    0.02300    0.75   0.4525
## Year2003            0.01609    0.02420    0.66   0.5061
## Year2004            0.02581    0.02283    1.13   0.2584
## Year2005           -0.02597    0.02217   -1.17   0.2415
## Year2006           -0.02462    0.02188   -1.13   0.2606
## Year2007           -0.00815    0.02141   -0.38   0.7035
## Year2008           -0.01268    0.02139   -0.59   0.5532
## Year2009           -0.03617    0.02154   -1.68   0.0932 .
## Year2010           -0.00585    0.02122   -0.28   0.7829
## Year2011           -0.06555    0.02112   -3.10   0.0019 **
## Year2012           -0.09115    0.02166   -4.21  2.6e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.471
## Multiple R-squared:  0.0877, Adjusted R-squared:  0.0867
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  1760 weights are ~= 1. The remaining 18708 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0004  0.8650  0.9500  0.9020  0.9860  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          4.89e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.085  1           1.042
## LastAuthorFemale  1.080  1           1.039
## Year              1.024 16           1.001
```



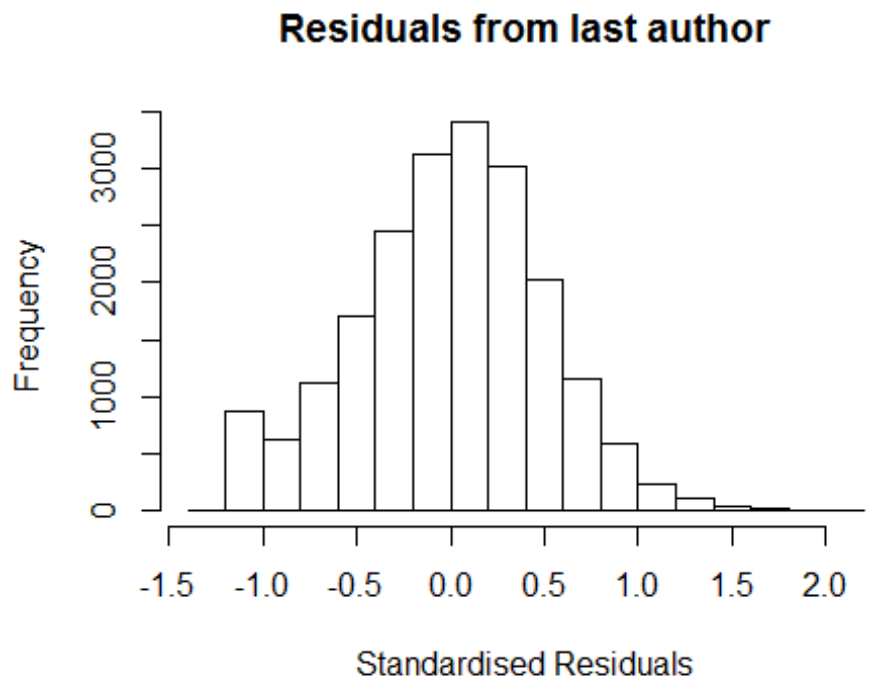

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.2149 -0.3300  0.0201  0.3297  2.1185
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)          1.11105    0.01868   59.49   <2e-16 ***
## FirstAuthorFemale1   0.02433    0.00748    3.25   0.0012 **
## LastAuthorFemale1    0.00511    0.00752    0.68   0.4969
## Year1997            -0.04496    0.02586   -1.74   0.0821 .
## Year1998            -0.05285    0.02709   -1.95   0.0511 .
## Year1999            -0.04030    0.02497   -1.61   0.1065
## Year2000             0.03156    0.02415    1.31   0.1913
## Year2001             0.02998    0.02500    1.20   0.2305
## Year2002             0.05684    0.02425    2.34   0.0191 *
## Year2003             0.05299    0.02553    2.08   0.0379 *
## Year2004             0.07441    0.02385    3.12   0.0018 **
## Year2005             0.02823    0.02307    1.22   0.2212
```

```
## Year2006            0.03780    0.02282    1.66   0.0976 .
## Year2007            0.04501    0.02249    2.00   0.0454 *
## Year2008            0.03974    0.02258    1.76   0.0785 .
## Year2009            0.01252    0.02261    0.55   0.5798
## Year2010            0.05448    0.02231    2.44   0.0146 *
## Year2011           -0.00420    0.02220   -0.19   0.8499
## Year2012           -0.02394    0.02289   -1.05   0.2956
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.487
## Multiple R-squared:  0.00569,    Adjusted R-squared:  0.00481
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  1701 weights are ~= 1. The remaining 18767 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0195  0.8660  0.9500  0.9010  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          4.89e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.017  1           1.008
## Year              1.017 16           1.001
```




```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.2128 -0.3306  0.0199  0.3292  2.1170
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)          1.11227    0.01855   59.95  < 2e-16 ***
## FirstAuthorFemale1   0.02604    0.00727    3.58  0.00034 ***
## Year1997            -0.04510    0.02586   -1.74  0.08115 .
## Year1998            -0.05283    0.02709   -1.95  0.05119 .
## Year1999            -0.04016    0.02497   -1.61  0.10788
## Year2000             0.03161    0.02415    1.31  0.19067
## Year2001             0.03006    0.02500    1.20  0.22930
## Year2002             0.05689    0.02425    2.35  0.01897 *
## Year2003             0.05311    0.02553    2.08  0.03754 *
## Year2004             0.07445    0.02385    3.12  0.00180 **
## Year2005             0.02825    0.02308    1.22  0.22096
## Year2006             0.03795    0.02282    1.66  0.09634 .
```

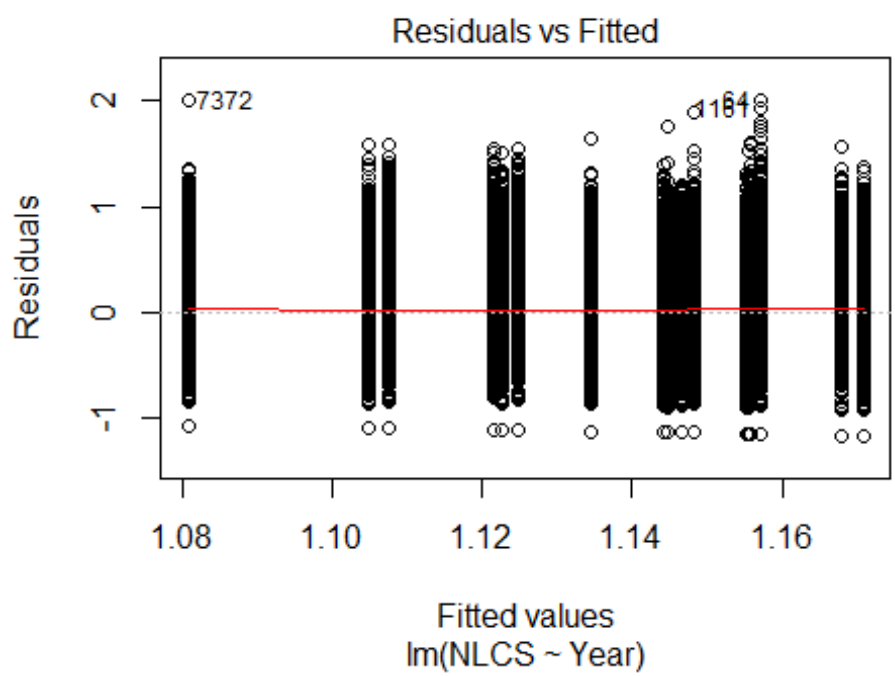
```
## Year2007            0.04527    0.02249    2.01  0.04416 *
## Year2008            0.03978    0.02259    1.76  0.07824 .
## Year2009            0.01276    0.02261    0.56  0.57245
## Year2010            0.05474    0.02231    2.45  0.01416 *
## Year2011           -0.00399    0.02220   -0.18  0.85734
## Year2012           -0.02353    0.02288   -1.03  0.30382
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.488
## Multiple R-squared:  0.00566,    Adjusted R-squared:  0.00483
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  1700 weights are ~= 1. The remaining 18768 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0199  0.8660  0.9500  0.9010  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          4.89e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##                   GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.012  1           1.006
## Year             1.012 16           1.000
```



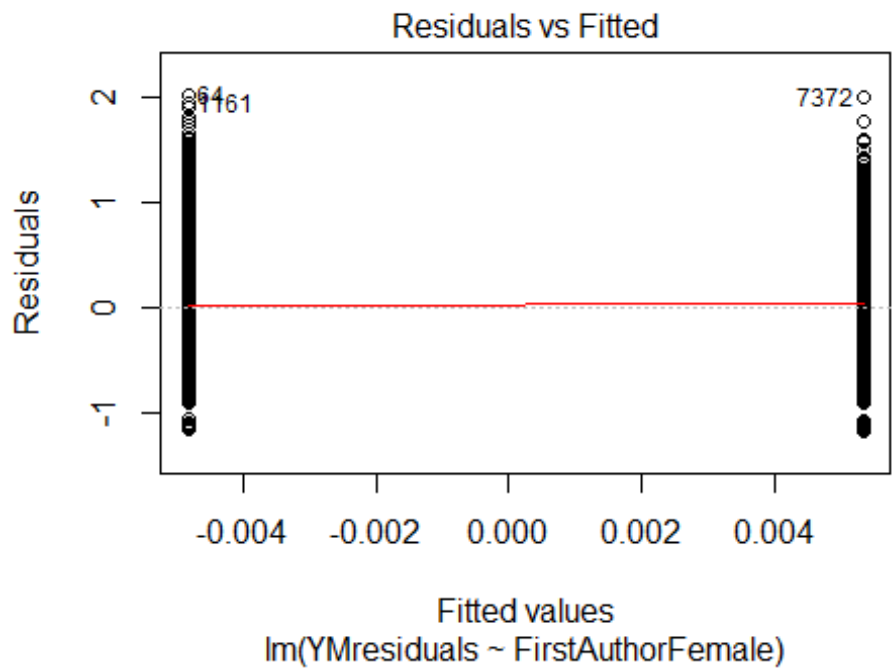

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.2066 -0.3300  0.0208  0.3308  2.1109
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)          1.1171     0.0186   60.17   <2e-16 ***
## LastAuthorFemale1    0.0134     0.0073    1.83   0.0670 .
## Year1997            -0.0452     0.0259   -1.75   0.0808 .
## Year1998            -0.0526     0.0271   -1.94   0.0521 .
## Year1999            -0.0410     0.0250   -1.64   0.1009
## Year2000             0.0322     0.0242    1.33   0.1831
## Year2001             0.0308     0.0250    1.23   0.2175
## Year2002             0.0578     0.0242    2.38   0.0171 *
## Year2003             0.0542     0.0255    2.12   0.0337 *
## Year2004             0.0761     0.0238    3.20   0.0014 **
## Year2005             0.0305     0.0231    1.32   0.1860
## Year2006             0.0390     0.0228    1.71   0.0869 .
```

```
## Year2007            0.0468     0.0225    2.08   0.0373 *
## Year2008            0.0419     0.0226    1.86   0.0632 .
## Year2009            0.0141     0.0226    0.63   0.5319
## Year2010            0.0560     0.0223    2.51   0.0120 *
## Year2011           -0.0019     0.0222   -0.09   0.9317
## Year2012           -0.0219     0.0228   -0.96   0.3381
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.488
## Multiple R-squared:  0.00519,    Adjusted R-squared:  0.00437
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  1728 weights are ~= 1. The remaining 18740 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0213  0.8660  0.9500  0.9010  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          4.89e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Sample size for the above analysis:  20468"
## [1] ""
## [1] ""
## [1] "####################################"
## [1] "Analysis of AJSC 3203"
## [1] "####################################"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 2009 1783 1799 1675 2129 2273 2171 1821 2108 2157 2471 2863 2885 3207 3336
## 2011 2012
## 3606 3717
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 1728 1530 1517 1410 1654 1604 1875 1621 1858 1902 2158 2504 2536 2803 2933
## 2011 2012
```

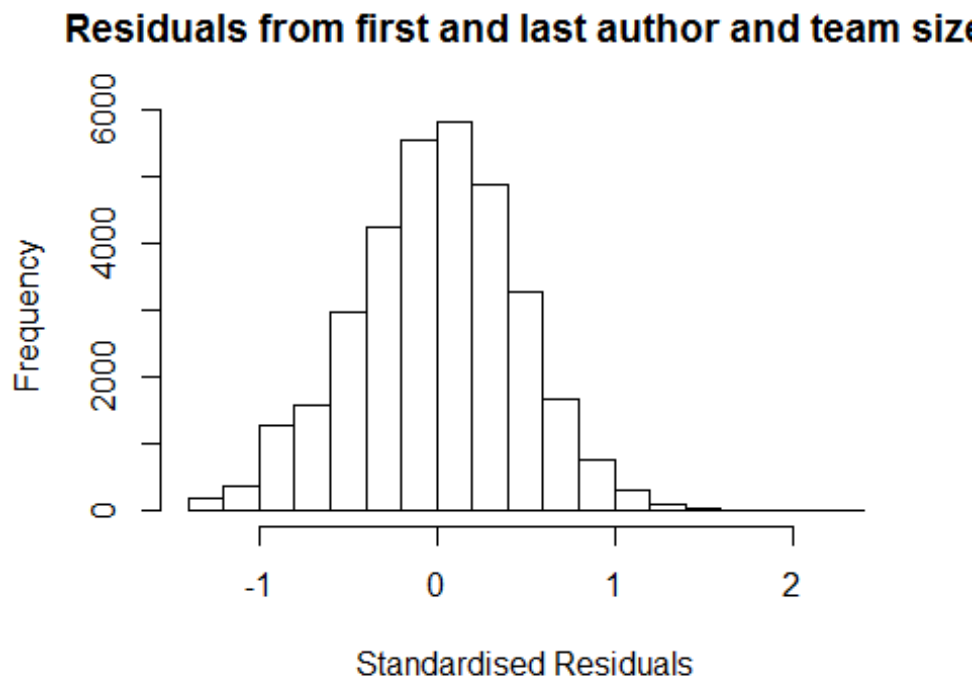
```
## 3159 3256
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 1614 1418 1403 1304 1510 1484 1711 1483 1695 1727 1974 2266 2311 2559 2652
## 2011 2012
## 2880 2959
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
##  Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 210, df = 16, p-value <2e-16
```




```
##
##  Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 51, df = 1, p-value = 9e-13
```

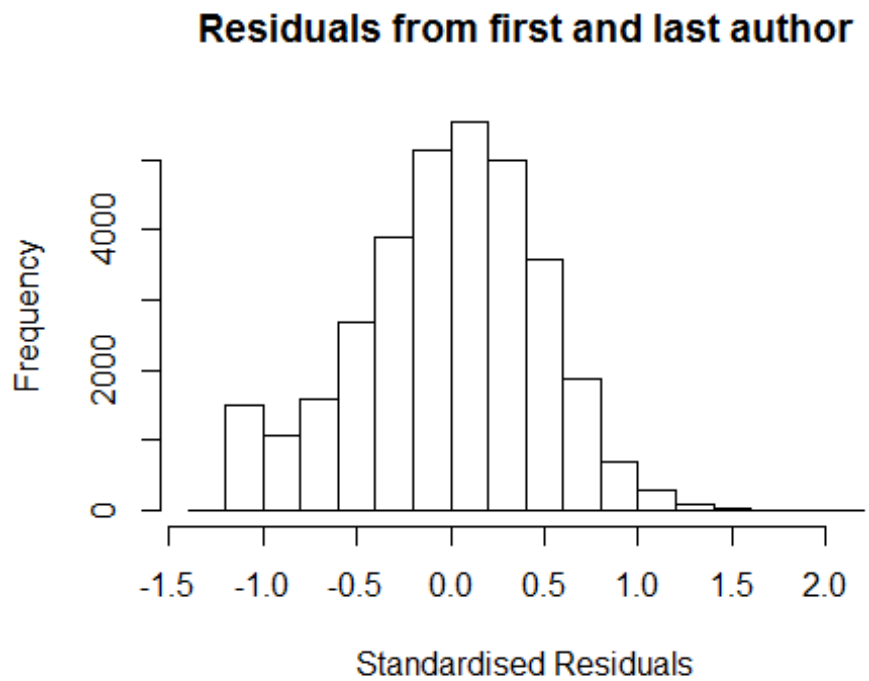



```
## [1] "Female first author team size 2018 geometric mean: 3.62944918319558"
## [1] "Male first author team size 2018 geometric mean: 3.20021371989917"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 1300000, p-value = 5e-06
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 3.44718522080462"
## [1] "Male last author team size 2018 geometric mean: 3.44417401662828"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 1300000, p-value = 0.7
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.067  1           1.033
## LastAuthorFemale  1.050  1           1.025
## UniqueAuthors     1.068  4           1.008
## Year              1.068 16           1.002
```



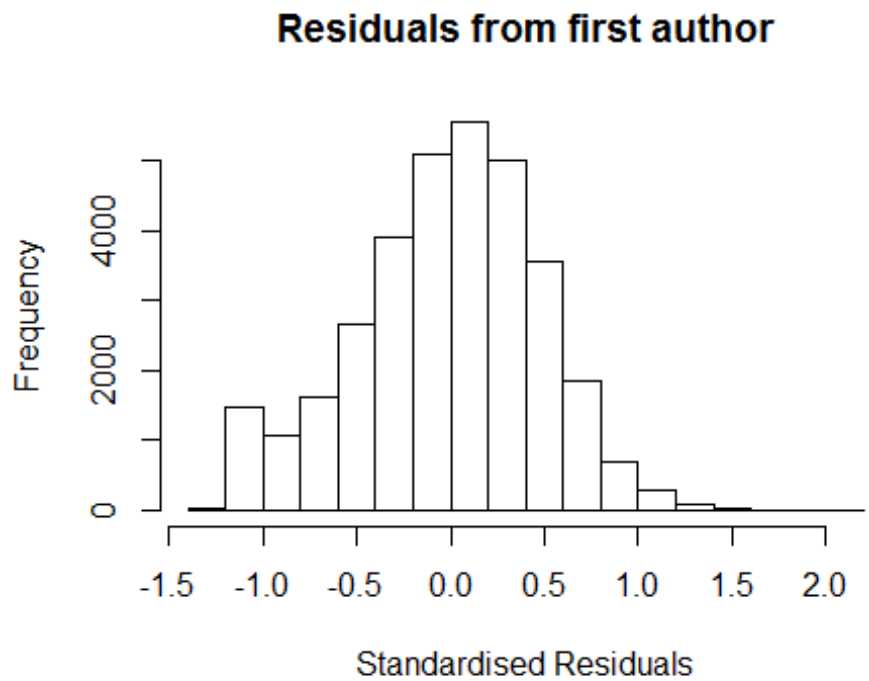

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
UniqueAuthors +
##     Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##     k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.3837 -0.3047  0.0117  0.3059  2.2785
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)         0.876596   0.015245   57.50  < 2e-16 ***
## FirstAuthorFemale1  0.000022   0.005351    0.00  0.99673
## LastAuthorFemale1  -0.021643   0.005397   -4.01  6.1e-05 ***
## UniqueAuthors2      0.336215   0.009101   36.94  < 2e-16 ***
## UniqueAuthors3      0.392685   0.009204   42.67  < 2e-16 ***
## UniqueAuthors4      0.439334   0.009501   46.24  < 2e-16 ***
## UniqueAuthors5      0.507110   0.008626   58.79  < 2e-16 ***
## Year1997           -0.033106   0.019732   -1.68  0.09339 .
## Year1998           -0.031568   0.019817   -1.59  0.11118
## Year1999           -0.053505   0.019470   -2.75  0.00600 **
```

```
## Year2000           -0.051449   0.018662   -2.76  0.00584 **
## Year2001           -0.017437   0.018807   -0.93  0.35383
## Year2002           -0.012394   0.018307   -0.68  0.49840
## Year2003           -0.028245   0.018209   -1.55  0.12087
## Year2004           -0.057067   0.017751   -3.21  0.00131 **
## Year2005           -0.062992   0.017273   -3.65  0.00027 ***
## Year2006           -0.074062   0.017119   -4.33  1.5e-05 ***
## Year2007           -0.060432   0.016748   -3.61  0.00031 ***
## Year2008           -0.081777   0.016589   -4.93  8.3e-07 ***
## Year2009           -0.054522   0.016747   -3.26  0.00113 **
## Year2010           -0.068209   0.016520   -4.13  3.7e-05 ***
## Year2011           -0.093979   0.016539   -5.68  1.3e-08 ***
## Year2012           -0.101268   0.016710   -6.06  1.4e-09 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.451
## Multiple R-squared:  0.134,  Adjusted R-squared:  0.134
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  observation 4860 is an outlier with |weight| = 0 ( < 3e-06);
##  2755 weights are ~= 1. The remaining 30194 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0126  0.8680  0.9500  0.9040  0.9860  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          3.03e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.037  1           1.018
## LastAuthorFemale  1.028  1           1.014
## Year              1.022 16           1.001
```



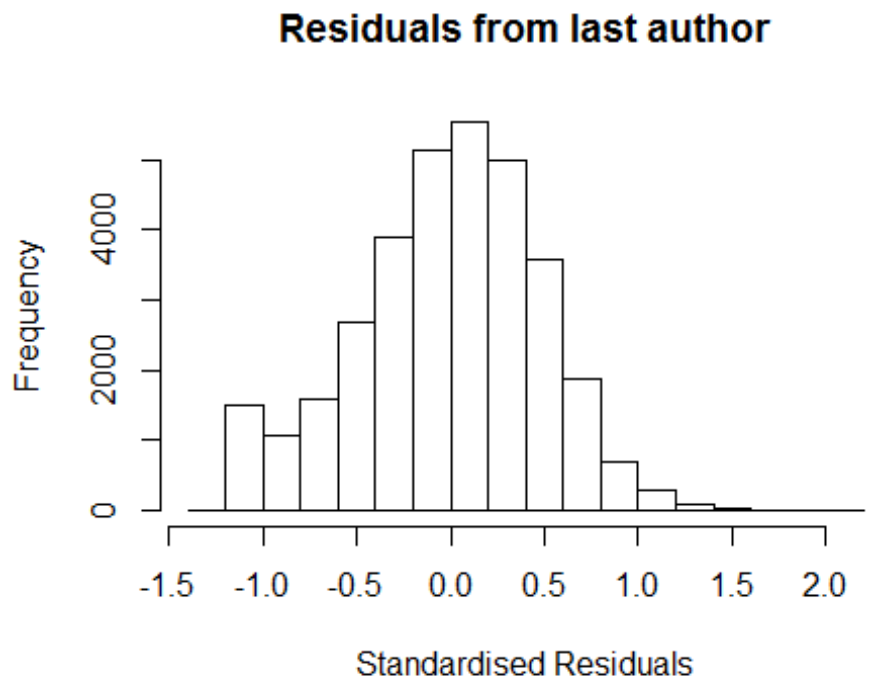

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.2134 -0.3219  0.0224  0.3251  2.0103
##
## Coefficients:
##                      Estimate Std. Error t value Pr(>|t|)
## (Intercept)          1.152680   0.014853   77.60  < 2e-16 ***
## FirstAuthorFemale1   0.025032   0.005616    4.46  8.3e-06 ***
## LastAuthorFemale1   -0.023657   0.005676   -4.17  3.1e-05 ***
## Year1997            -0.050734   0.020843   -2.43   0.0149 *
## Year1998            -0.019090   0.020719   -0.92   0.3569
## Year1999            -0.061944   0.020587   -3.01   0.0026 **
## Year2000            -0.047477   0.019778   -2.40   0.0164 *
## Year2001             0.005458   0.019906    0.27   0.7840
## Year2002             0.035700   0.019274    1.85   0.0640 .
## Year2003             0.028035   0.018975    1.48   0.1396
## Year2004             0.015522   0.018669    0.83   0.4057
## Year2005             0.009831   0.018189    0.54   0.5889
```

```
## Year2006           -0.000855   0.017965   -0.05   0.9620
## Year2007            0.003363   0.017693    0.19   0.8492
## Year2008           -0.011334   0.017578   -0.64   0.5191
## Year2009            0.016944   0.017709    0.96   0.3387
## Year2010           -0.002924   0.017478   -0.17   0.8671
## Year2011           -0.018965   0.017450   -1.09   0.2771
## Year2012           -0.023249   0.017645   -1.32   0.1876
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.475
## Multiple R-squared:  0.0033, Adjusted R-squared:  0.00275
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
##  2737 weights are ~= 1. The remaining 30213 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0334  0.8670  0.9490  0.9010  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          3.03e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.017  1           1.009
## Year              1.017 16           1.001
```




```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.2004 -0.3204  0.0226  0.3256  2.0168
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)          1.14620    0.01476   77.66  < 2e-16 ***
## FirstAuthorFemale1   0.01925    0.00559    3.44  0.00058 ***
## Year1997            -0.05026    0.02085   -2.41  0.01592 *
## Year1998            -0.01891    0.02073   -0.91  0.36158
## Year1999            -0.06183    0.02059   -3.00  0.00268 **
## Year2000            -0.04830    0.01978   -2.44  0.01458 *
## Year2001             0.00487    0.01991    0.24  0.80667
## Year2002             0.03493    0.01927    1.81  0.06986 .
## Year2003             0.02768    0.01898    1.46  0.14476
## Year2004             0.01499    0.01867    0.80  0.42195
## Year2005             0.00924    0.01819    0.51  0.61158
## Year2006            -0.00168    0.01797   -0.09  0.92531
```

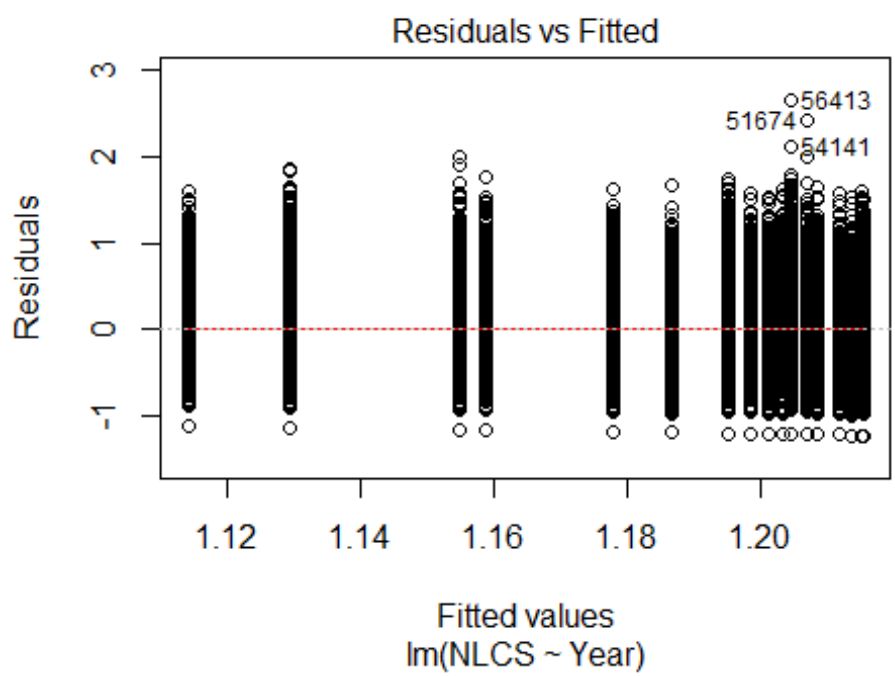
```
## Year2007            0.00282    0.01770    0.16  0.87346
## Year2008           -0.01219    0.01758   -0.69  0.48802
## Year2009            0.01622    0.01772    0.92  0.35977
## Year2010           -0.00353    0.01748   -0.20  0.83985
## Year2011           -0.02001    0.01745   -1.15  0.25154
## Year2012           -0.02461    0.01764   -1.40  0.16301
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.475
## Multiple R-squared:  0.00279,    Adjusted R-squared:  0.00228
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
##  2757 weights are ~= 1. The remaining 30193 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0316  0.8670  0.9490  0.9010  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          3.03e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##                   GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.008  1           1.004
## Year             1.008 16           1.000
```



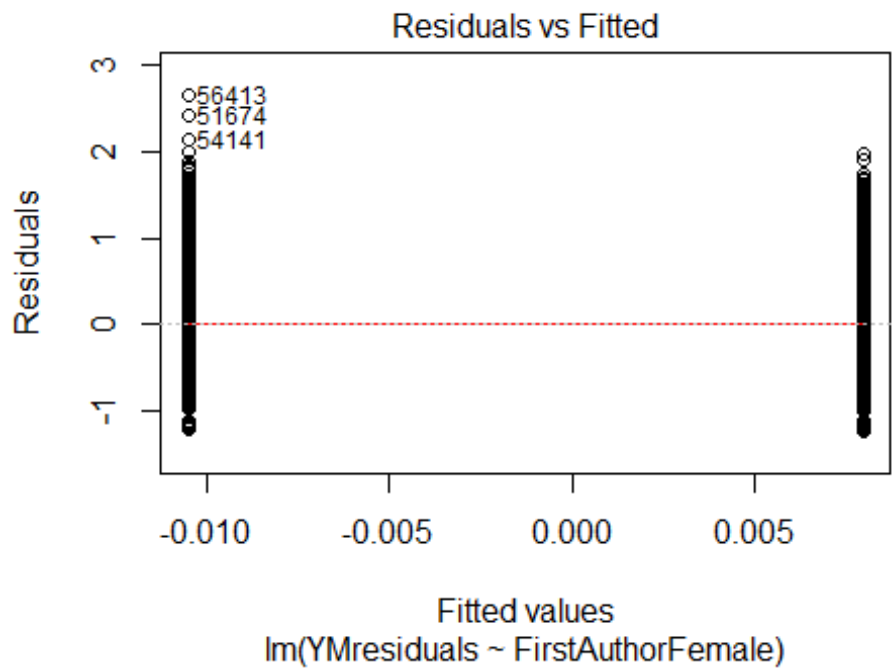

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.1967 -0.3216  0.0222  0.3258  2.0031
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.15995    0.01474   78.69   <2e-16 ***
## LastAuthorFemale1 -0.01742    0.00565   -3.08   0.0020 **
## Year1997          -0.05039    0.02084   -2.42   0.0156 *
## Year1998          -0.01831    0.02071   -0.88   0.3765
## Year1999          -0.06140    0.02060   -2.98   0.0029 **
## Year2000          -0.04645    0.01979   -2.35   0.0189 *
## Year2001           0.00668    0.01990    0.34   0.7370
## Year2002           0.03680    0.01929    1.91   0.0564 .
## Year2003           0.02954    0.01897    1.56   0.1195
## Year2004           0.01781    0.01866    0.95   0.3400
## Year2005           0.01202    0.01818    0.66   0.5086
## Year2006           0.00129    0.01796    0.07   0.9427
```

```
## Year2007           0.00590    0.01767    0.33   0.7385
## Year2008          -0.00856    0.01756   -0.49   0.6261
## Year2009           0.02000    0.01769    1.13   0.2582
## Year2010           0.00026    0.01746    0.01   0.9881
## Year2011          -0.01566    0.01742   -0.90   0.3688
## Year2012          -0.01940    0.01761   -1.10   0.2704
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.475
## Multiple R-squared:  0.00272,    Adjusted R-squared:  0.0022
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
##  2739 weights are ~= 1. The remaining 30211 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0356  0.8660  0.9490  0.9010  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          3.03e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Sample size for the above analysis:  32950"
## [1] ""
## [1] ""
## [1] "####################################"
## [1] "Analysis of AJSC 3204"
## [1] "####################################"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 2402 2100 2090 1988 2434 2467 2524 2164 2314 2361 2949 3360 3627 3950 3910
## 2011 2012
## 4232 4286
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 2115 1720 1742 1614 1951 1839 2224 1914 2013 2023 2565 2915 3140 3430 3362
## 2011 2012
```

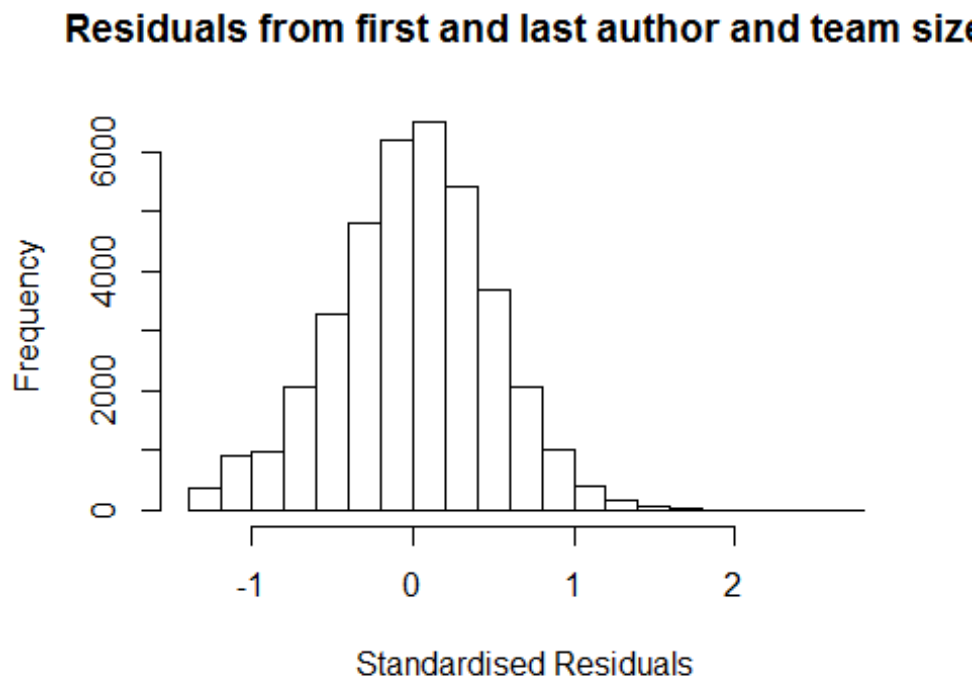
```
## 3615 3665
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 1991 1600 1626 1495 1797 1698 2025 1749 1843 1811 2306 2615 2817 3095 3007
## 2011 2012
## 3218 3236
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
##  Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 410, df = 16, p-value <2e-16
```




```
##
##  Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 35, df = 1, p-value = 3e-09
```

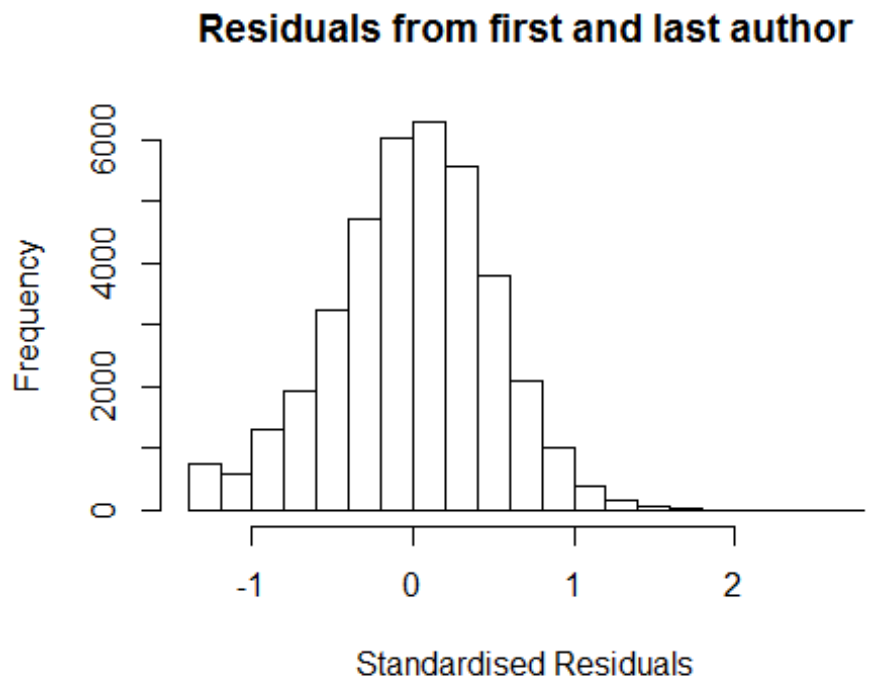



```
## [1] "Female first author team size 2018 geometric mean: 3.22147196413567"
## [1] "Male first author team size 2018 geometric mean: 2.87851195022437"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 1700000, p-value = 2e-06
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 3.07075098718664"
## [1] "Male last author team size 2018 geometric mean: 3.14853075855344"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 1600000, p-value = 0.09
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.082  1           1.040
## LastAuthorFemale  1.069  1           1.034
## UniqueAuthors     1.056  4           1.007
## Year              1.055 16           1.002
```



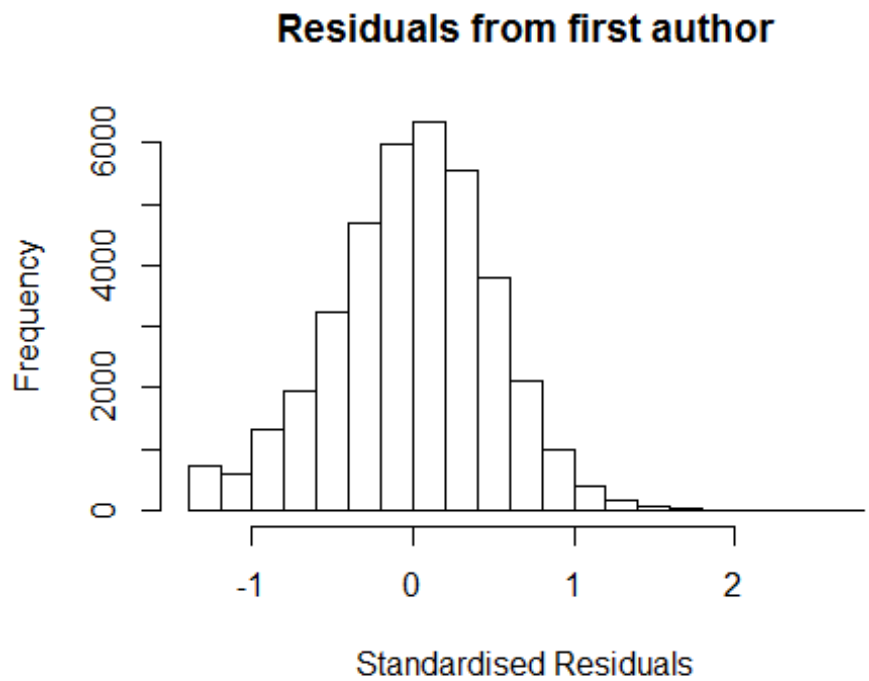

```
## [1] "List of  1 outliers with residuals above  2.5"
##          ScopusId  NLCS Year OneField Fields residuals
## 56413 84857285608 3.843 2012     1213      3     2.668
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
UniqueAuthors +
##     Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##     k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.3926 -0.3121  0.0107  0.3160  2.6676
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.04887    0.01481   70.82  < 2e-16 ***
## FirstAuthorFemale1  0.01556    0.00535    2.91  0.00366 **
## LastAuthorFemale1  -0.00721    0.00524   -1.38  0.16891
## UniqueAuthors2      0.17714    0.00804   22.04  < 2e-16 ***
## UniqueAuthors3      0.21428    0.00844   25.40  < 2e-16 ***
## UniqueAuthors4      0.26405    0.00936   28.21  < 2e-16 ***
## UniqueAuthors5      0.33167    0.00876   37.86  < 2e-16 ***
## Year1997           -0.05114    0.01956   -2.62  0.00892 **
## Year1998           -0.06638    0.01940   -3.42  0.00062 ***
## Year1999           -0.08248    0.01878   -4.39  1.1e-05 ***
```

```
## Year2000           -0.04804    0.01810   -2.65  0.00797 **
## Year2001           -0.03276    0.01839   -1.78  0.07483 .
## Year2002           -0.00347    0.01780   -0.19  0.84543
## Year2003           -0.02134    0.01780   -1.20  0.23062
## Year2004           -0.01285    0.01730   -0.74  0.45763
## Year2005           -0.02255    0.01745   -1.29  0.19624
## Year2006           -0.05654    0.01642   -3.44  0.00057 ***
## Year2007           -0.03345    0.01626   -2.06  0.03968 *
## Year2008           -0.03427    0.01603   -2.14  0.03249 *
## Year2009           -0.04161    0.01612   -2.58  0.00985 **
## Year2010           -0.02887    0.01635   -1.77  0.07750 .
## Year2011           -0.04307    0.01615   -2.67  0.00766 **
## Year2012           -0.05064    0.01659   -3.05  0.00227 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.468
## Multiple R-squared:  0.0514, Adjusted R-squared:  0.0508
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
##  2 observations c(36206,37735) are outliers with |weight| = 0 ( < 2.6e-
06);
##  3196 weights are ~= 1. The remaining 34731 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0342  0.8670  0.9510  0.9010  0.9860  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          2.64e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.069  1           1.034
## LastAuthorFemale  1.065  1           1.032
## Year              1.011 16           1.000
```



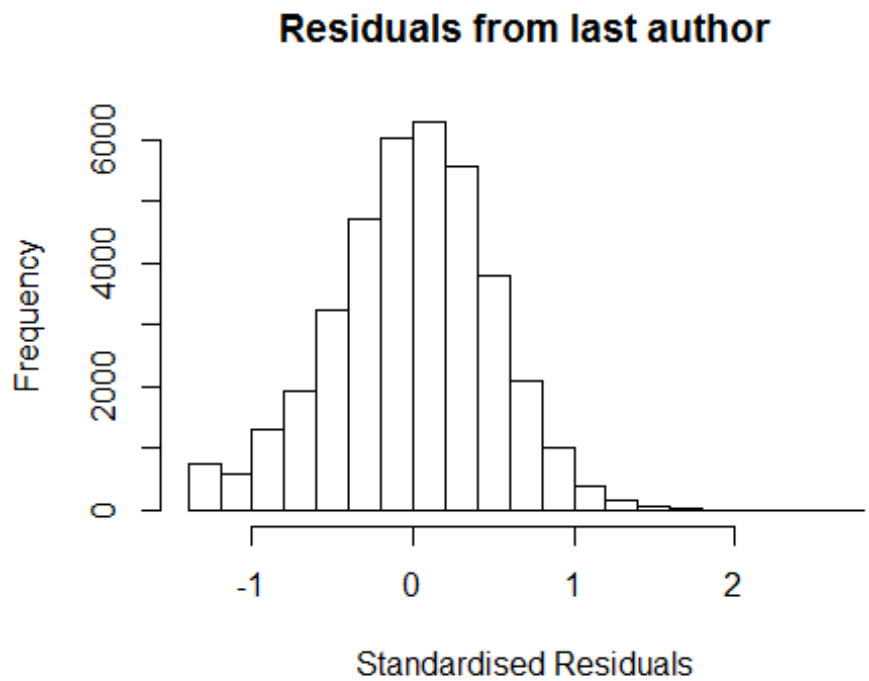

```
## [1] "List of  1 outliers with residuals above  2.5"
##          ScopusId  NLCS Year OneField Fields residuals
## 56413 84857285608 3.843 2012     1213      3     2.644
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.2480 -0.3201  0.0146  0.3217  2.6437
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)          1.19353    0.01437   83.05  < 2e-16 ***
## FirstAuthorFemale1  0.02804    0.00545    5.15  2.7e-07 ***
## LastAuthorFemale1  -0.01001    0.00535   -1.87  0.06149 .
## Year1997           -0.05971    0.02019   -2.96  0.00311 **
## Year1998           -0.07293    0.02007   -3.63  0.00028 ***
## Year1999           -0.07842    0.01935   -4.05  5.1e-05 ***
## Year2000           -0.03706    0.01862   -1.99  0.04657 *
## Year2001           -0.02207    0.01896   -1.16  0.24451
## Year2002            0.02645    0.01824    1.45  0.14710
## Year2003            0.00800    0.01826    0.44  0.66135
## Year2004            0.02529    0.01774    1.43  0.15396
## Year2005            0.01883    0.01783    1.06  0.29107
```

```
## Year2006           -0.01393    0.01685   -0.83  0.40827
## Year2007            0.01172    0.01666    0.70  0.48187
## Year2008            0.01158    0.01642    0.71  0.48047
## Year2009            0.00584    0.01655    0.35  0.72438
## Year2010            0.02073    0.01677    1.24  0.21624
## Year2011            0.00869    0.01659    0.52  0.60015
## Year2012            0.00576    0.01700    0.34  0.73474
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.478
## Multiple R-squared:  0.00441,    Adjusted R-squared:  0.00394
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
##  observation 37735 is an outlier with |weight| = 0 ( < 2.6e-06);
##  3143 weights are ~= 1. The remaining 34785 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0104  0.8670  0.9500  0.9000  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          2.64e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.009  1           1.004
## Year              1.009 16           1.000
```




```
## [1] "List of  1 outliers with residuals above  2.5"
##          ScopusId  NLCS Year OneField Fields residuals
## 56413 84857285608 3.843 2012     1213      3     2.644
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.2415 -0.3200  0.0147  0.3217  2.6474
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)          1.19040    0.01428   83.35  < 2e-16 ***
## FirstAuthorFemale1   0.02505    0.00530    4.72  2.3e-06 ***
## Year1997            -0.05955    0.02020   -2.95  0.00319 **
## Year1998            -0.07301    0.02008   -3.64  0.00028 ***
## Year1999            -0.07872    0.01935   -4.07  4.7e-05 ***
## Year2000            -0.03716    0.01862   -2.00  0.04597 *
## Year2001            -0.02238    0.01895   -1.18  0.23782
## Year2002             0.02609    0.01825    1.43  0.15268
## Year2003             0.00777    0.01826    0.43  0.67042
## Year2004             0.02497    0.01774    1.41  0.15912
## Year2005             0.01842    0.01783    1.03  0.30154
## Year2006            -0.01430    0.01684   -0.85  0.39592
```

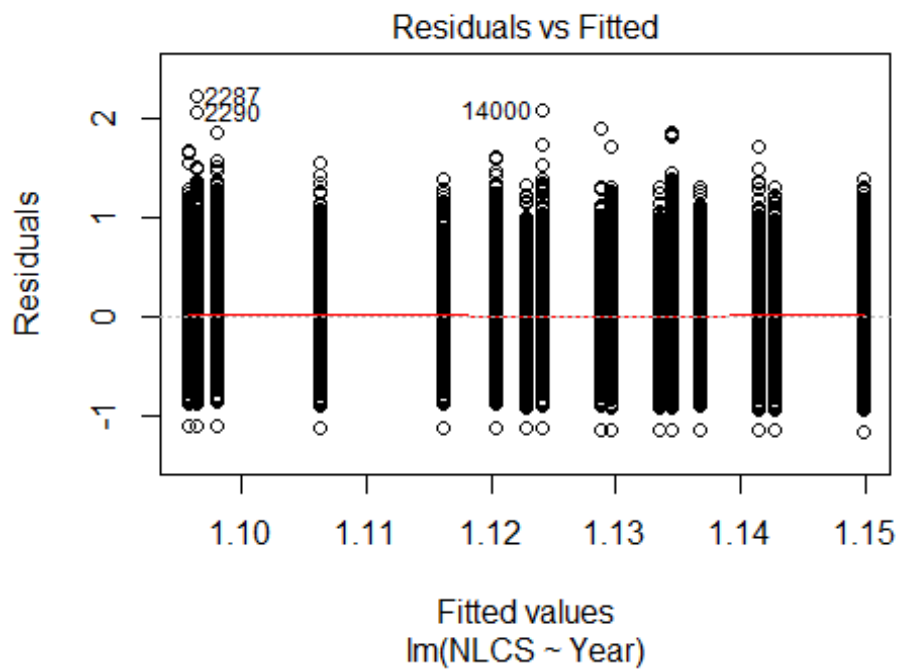
```
## Year2007            0.01149    0.01666    0.69  0.49018
## Year2008            0.01125    0.01642    0.68  0.49339
## Year2009            0.00557    0.01655    0.34  0.73646
## Year2010            0.02029    0.01676    1.21  0.22614
## Year2011            0.00824    0.01659    0.50  0.61946
## Year2012            0.00525    0.01700    0.31  0.75730
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.478
## Multiple R-squared:  0.00433,    Adjusted R-squared:  0.00388
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
##  observation 37735 is an outlier with |weight| = 0 ( < 2.6e-06);
##  3163 weights are ~= 1. The remaining 34765 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0098  0.8680  0.9500  0.9000  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          2.64e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##                   GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.005  1           1.002
## Year             1.005 16           1.000
```



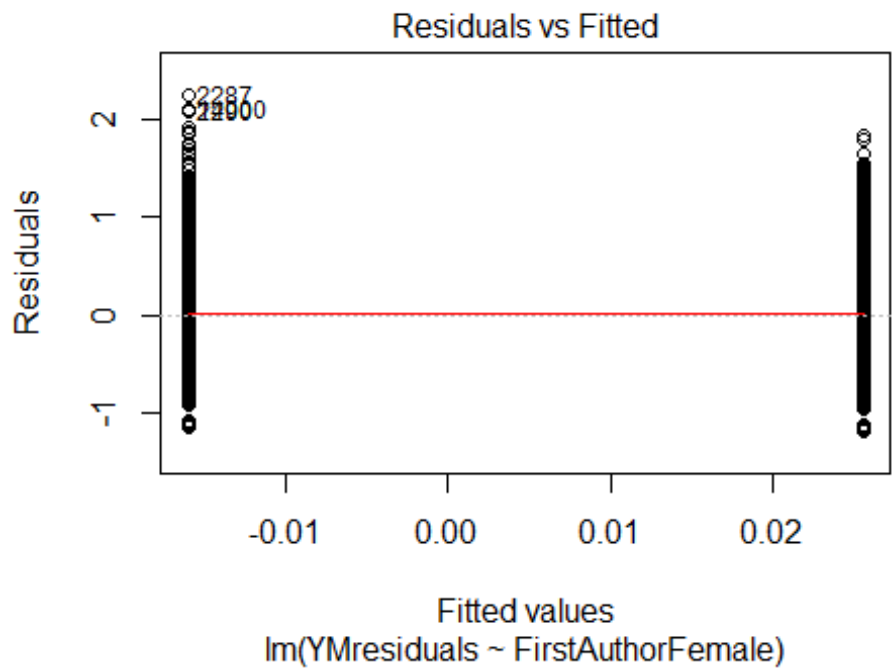

```
## [1] "List of  1 outliers with residuals above  2.5"
##          ScopusId  NLCS Year OneField Fields residuals
## 56413 84857285608 3.843 2012     1213      3     2.644
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.2321 -0.3201  0.0149  0.3234  2.6309
##
## Coefficients:
##                   Estimate Std. Error t value Pr(>|t|)
## (Intercept)        1.20354    0.01420   84.75  < 2e-16 ***
## LastAuthorFemale1 -0.00184    0.00521   -0.35  0.72468
## Year1997          -0.05935    0.02020   -2.94  0.00331 **
## Year1998          -0.07247    0.02008   -3.61  0.00031 ***
## Year1999          -0.07866    0.01935   -4.06  4.8e-05 ***
## Year2000          -0.03627    0.01862   -1.95  0.05148 .
## Year2001          -0.02100    0.01894   -1.11  0.26757
## Year2002           0.02852    0.01824    1.56  0.11800
## Year2003           0.00900    0.01826    0.49  0.62193
## Year2004           0.02731    0.01773    1.54  0.12359
## Year2005           0.02036    0.01784    1.14  0.25386
## Year2006          -0.01185    0.01684   -0.70  0.48171
```

```
## Year2007           0.01387    0.01666    0.83  0.40494
## Year2008           0.01442    0.01641    0.88  0.37945
## Year2009           0.00829    0.01654    0.50  0.61619
## Year2010           0.02324    0.01676    1.39  0.16541
## Year2011           0.01181    0.01658    0.71  0.47628
## Year2012           0.00855    0.01700    0.50  0.61511
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.478
## Multiple R-squared:  0.0037, Adjusted R-squared:  0.00326
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
##  observation 37735 is an outlier with |weight| = 0 ( < 2.6e-06);
##  3142 weights are ~= 1. The remaining 34786 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.013   0.868   0.950   0.900   0.985   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          2.64e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Sample size for the above analysis:  37929"
## [1] ""
## [1] ""
## [1] "####################################"
## [1] "Analysis of AJSC 3205"
## [1] "####################################"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 1535 1414 1223 1177 1254 1477 1420 1227 1354 1352 1525 1640 1819 1911 1904
## 2011 2012
## 2081 2002
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 1288 1173 1022  951  982 1033 1214 1048 1145 1129 1315 1409 1563 1627 1609
```

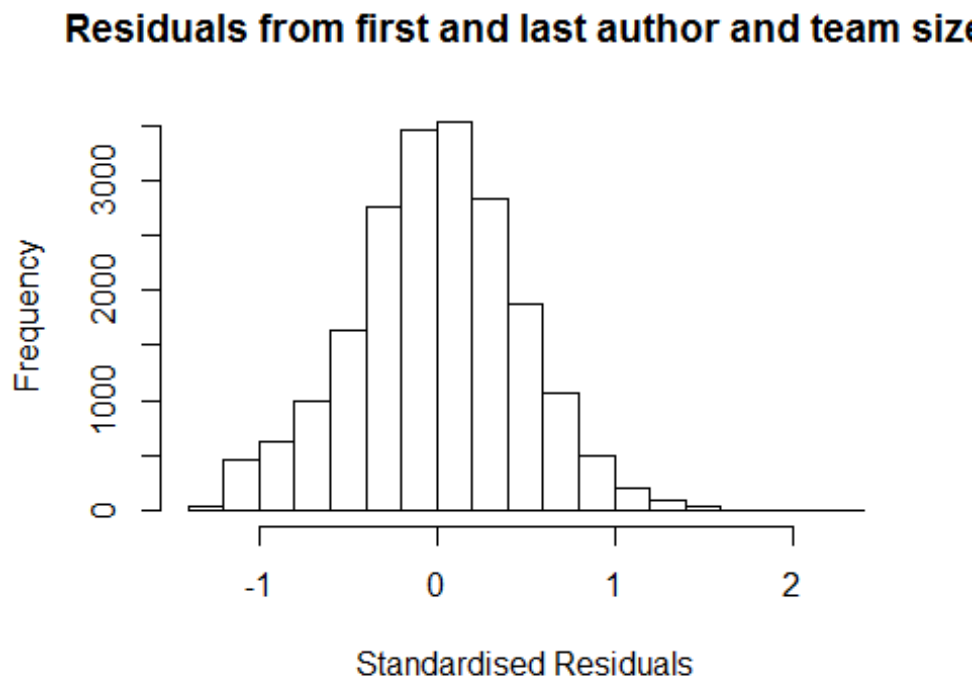
```
## 2011 2012
## 1780 1694
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 1175 1098  953  895  909  946 1093  963 1037 1026 1184 1285 1420 1487 1465
## 2011 2012
## 1611 1538
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
##  Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 270, df = 16, p-value <2e-16
```




```
##
##  Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 50, df = 1, p-value = 2e-12
```

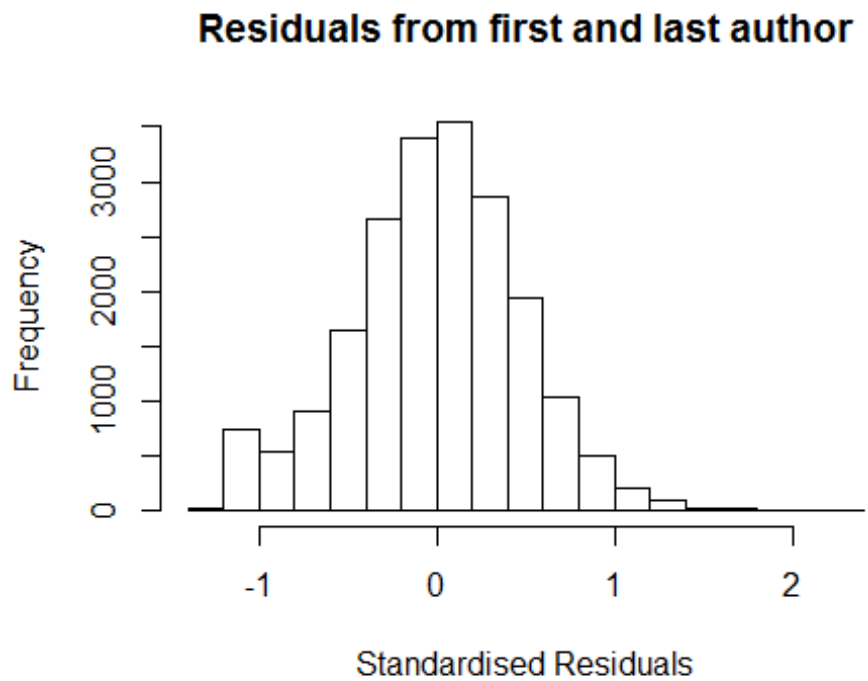



```
## [1] "Female first author team size 2018 geometric mean: 3.03453943069918"
## [1] "Male first author team size 2018 geometric mean: 2.64170818389654"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 4e+05, p-value = 3e-06
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 2.95443704395515"
## [1] "Male last author team size 2018 geometric mean: 2.74481382911957"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 380000, p-value = 0.006
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.074  1           1.036
## LastAuthorFemale  1.062  1           1.030
## UniqueAuthors     1.048  4           1.006
## Year              1.048 16           1.001
```



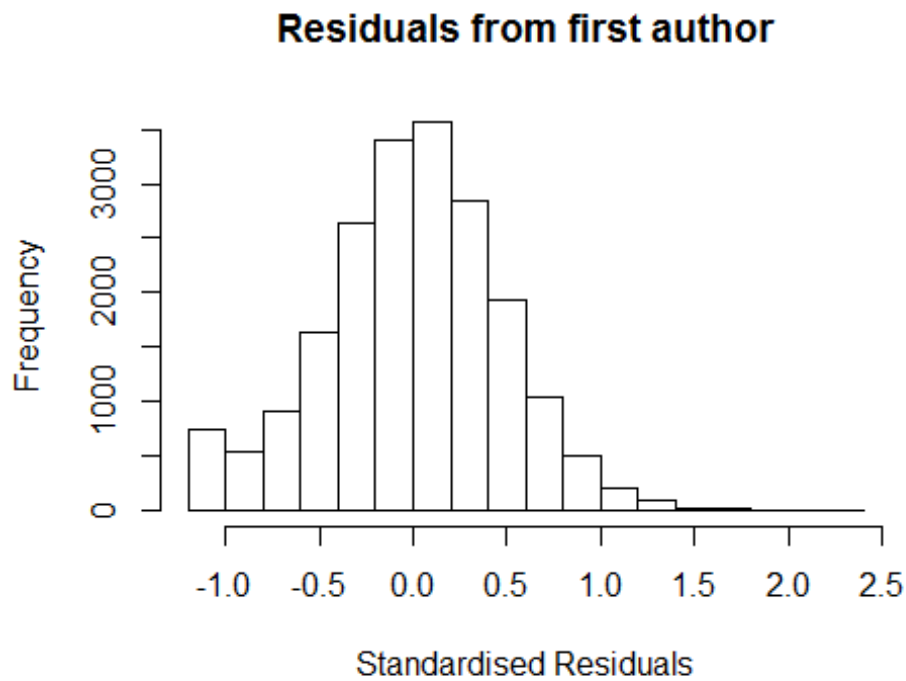

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
UniqueAuthors +
##     Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##     k.max = 1000))
##  \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.24932 -0.29912  0.00522  0.30548  2.33356
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)         0.972636   0.019121   50.87   <2e-16 ***
## FirstAuthorFemale1  0.023115   0.007042    3.28    0.001 **
## LastAuthorFemale1   0.013856   0.007197    1.93    0.054 .
## UniqueAuthors2      0.156950   0.010651   14.74   <2e-16 ***
## UniqueAuthors3      0.167258   0.011241   14.88   <2e-16 ***
## UniqueAuthors4      0.168717   0.012737   13.25   <2e-16 ***
## UniqueAuthors5      0.207591   0.012447   16.68   <2e-16 ***
## Year1997            0.008806   0.023876    0.37    0.712
## Year1998           -0.004737   0.024781   -0.19    0.848
## Year1999            0.034950   0.023396    1.49    0.135
```

```
## Year2000            0.024578   0.023987    1.02    0.306
## Year2001            0.019267   0.024988    0.77    0.441
## Year2002            0.055239   0.022915    2.41    0.016 *
## Year2003            0.021002   0.023175    0.91    0.365
## Year2004            0.000592   0.022570    0.03    0.979
## Year2005            0.021630   0.022355    0.97    0.333
## Year2006            0.008869   0.021655    0.41    0.682
## Year2007            0.000456   0.021222    0.02    0.983
## Year2008            0.010766   0.020728    0.52    0.603
## Year2009           -0.003795   0.020827   -0.18    0.855
## Year2010            0.008139   0.021009    0.39    0.698
## Year2011           -0.011988   0.020883   -0.57    0.566
## Year2012            0.003497   0.021704    0.16    0.872
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.45
## Multiple R-squared:  0.0257, Adjusted R-squared:  0.0246
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
##  2 observations c(1494,1497) are outliers with |weight| = 0 ( < 5e-06);
##  1657 weights are ~= 1. The remaining 18426 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0015  0.8660  0.9510  0.8990  0.9860  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          4.98e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.061  1           1.030
## LastAuthorFemale  1.057  1           1.028
## Year              1.015 16           1.000
```



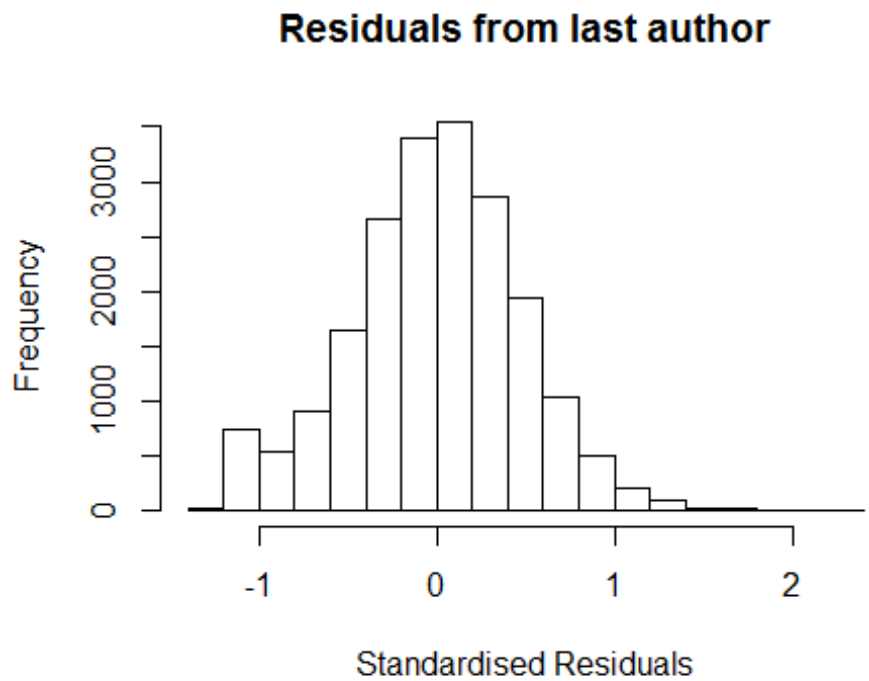

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.20942 -0.30404  0.00991  0.30548  2.22691
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)          1.08645    0.01785   60.85  < 2e-16 ***
## FirstAuthorFemale1   0.03692    0.00706    5.23  1.7e-07 ***
## LastAuthorFemale1    0.01970    0.00723    2.72   0.0065 **
## Year1997             0.00164    0.02414    0.07   0.9459
## Year1998            -0.00525    0.02528   -0.21   0.8356
## Year1999             0.03393    0.02370    1.43   0.1522
## Year2000             0.02967    0.02434    1.22   0.2228
## Year2001             0.02633    0.02511    1.05   0.2945
## Year2002             0.06636    0.02308    2.87   0.0040 **
## Year2003             0.03851    0.02328    1.65   0.0981 .
## Year2004             0.01642    0.02275    0.72   0.4704
## Year2005             0.04470    0.02259    1.98   0.0478 *
```

```
## Year2006            0.02537    0.02182    1.16   0.2450
## Year2007            0.02085    0.02136    0.98   0.3290
## Year2008            0.02908    0.02094    1.39   0.1650
## Year2009            0.01794    0.02105    0.85   0.3940
## Year2010            0.03174    0.02119    1.50   0.1342
## Year2011            0.01023    0.02108    0.49   0.6276
## Year2012            0.02907    0.02197    1.32   0.1858
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.453
## Multiple R-squared:  0.00355,    Adjusted R-squared:  0.00265
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
##  observation 1494 is an outlier with |weight| = 0 ( < 5e-06);
##  1711 weights are ~= 1. The remaining 18373 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0006  0.8650  0.9500  0.8980  0.9860  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          4.98e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                  psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.011  1           1.005
## Year              1.011 16           1.000
```




```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.19983 -0.30457  0.00921  0.30583  2.22229
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.09050    0.01775   61.43  < 2e-16 ***
## FirstAuthorFemale1  0.04221    0.00690    6.12  9.6e-10 ***
## Year1997            0.00221    0.02413    0.09   0.9271
## Year1998           -0.00482    0.02527   -0.19   0.8488
## Year1999            0.03435    0.02368    1.45   0.1469
## Year2000            0.03034    0.02430    1.25   0.2118
## Year2001            0.02748    0.02509    1.10   0.2734
## Year2002            0.06712    0.02308    2.91   0.0036 **
## Year2003            0.03930    0.02327    1.69   0.0912 .
## Year2004            0.01687    0.02273    0.74   0.4581
## Year2005            0.04524    0.02257    2.00   0.0450 *
## Year2006            0.02645    0.02180    1.21   0.2250
```

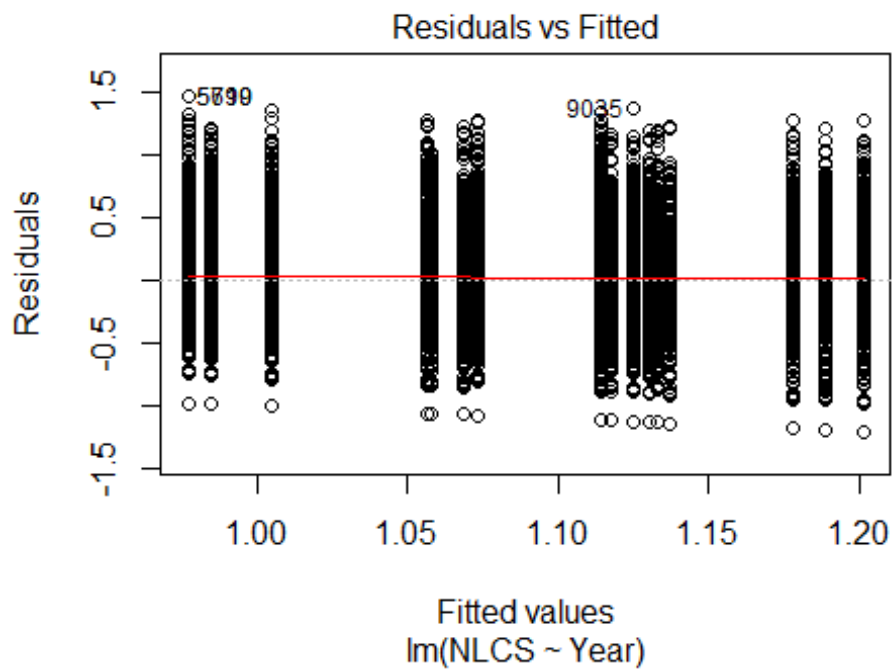
```
## Year2007            0.02177    0.02134    1.02   0.3077
## Year2008            0.02958    0.02093    1.41   0.1575
## Year2009            0.01890    0.02103    0.90   0.3688
## Year2010            0.03301    0.02117    1.56   0.1189
## Year2011            0.01107    0.02105    0.53   0.5990
## Year2012            0.03049    0.02196    1.39   0.1650
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.453
## Multiple R-squared:  0.00316,    Adjusted R-squared:  0.00232
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
##  observation 1494 is an outlier with |weight| = 0 ( < 5e-06);
##  1710 weights are ~= 1. The remaining 18374 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0009  0.8650  0.9500  0.8980  0.9860  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          4.98e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##                   GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.007  1           1.003
## Year             1.007 16           1.000
```



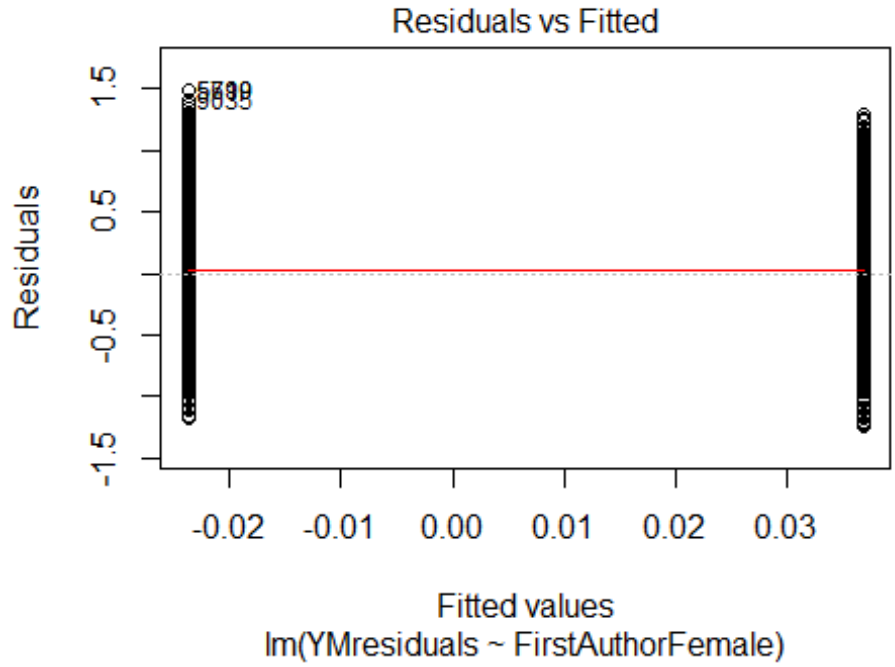

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.1930 -0.3047  0.0113  0.3047  2.2179
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.09575    0.01776   61.71  < 2e-16 ***
## LastAuthorFemale1   0.03006    0.00707    4.25  2.1e-05 ***
## Year1997            0.00135    0.02414    0.06   0.9553
## Year1998           -0.00574    0.02530   -0.23   0.8204
## Year1999            0.03319    0.02373    1.40   0.1619
## Year2000            0.02964    0.02437    1.22   0.2238
## Year2001            0.02747    0.02513    1.09   0.2743
## Year2002            0.06717    0.02310    2.91   0.0036 **
## Year2003            0.03916    0.02329    1.68   0.0927 .
## Year2004            0.01923    0.02277    0.84   0.3983
## Year2005            0.04645    0.02264    2.05   0.0402 *
## Year2006            0.02712    0.02184    1.24   0.2144
```

```
## Year2007           0.02304    0.02138    1.08   0.2810
## Year2008           0.03131    0.02096    1.49   0.1352
## Year2009           0.02089    0.02106    0.99   0.3213
## Year2010           0.03396    0.02119    1.60   0.1090
## Year2011           0.01355    0.02110    0.64   0.5207
## Year2012           0.03260    0.02197    1.48   0.1379
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.454
## Multiple R-squared:  0.00219,    Adjusted R-squared:  0.00134
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
##  observation 1494 is an outlier with |weight| = 0 ( < 5e-06);
##  1683 weights are ~= 1. The remaining 18401 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0015  0.8660  0.9500  0.8980  0.9860  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          4.98e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Sample size for the above analysis:  20085"
## [1] ""
## [1] ""
## [1] "####################################"
## [1] "Analysis of AJSC 3206"
## [1] "####################################"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  540  507  442  427  588  766  596  551  645  662  671  705  665  674  726
## 2011 2012
##  702  765
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  427  413  340  340  340  479  514  485  557  569  579  612  588  571  607
```

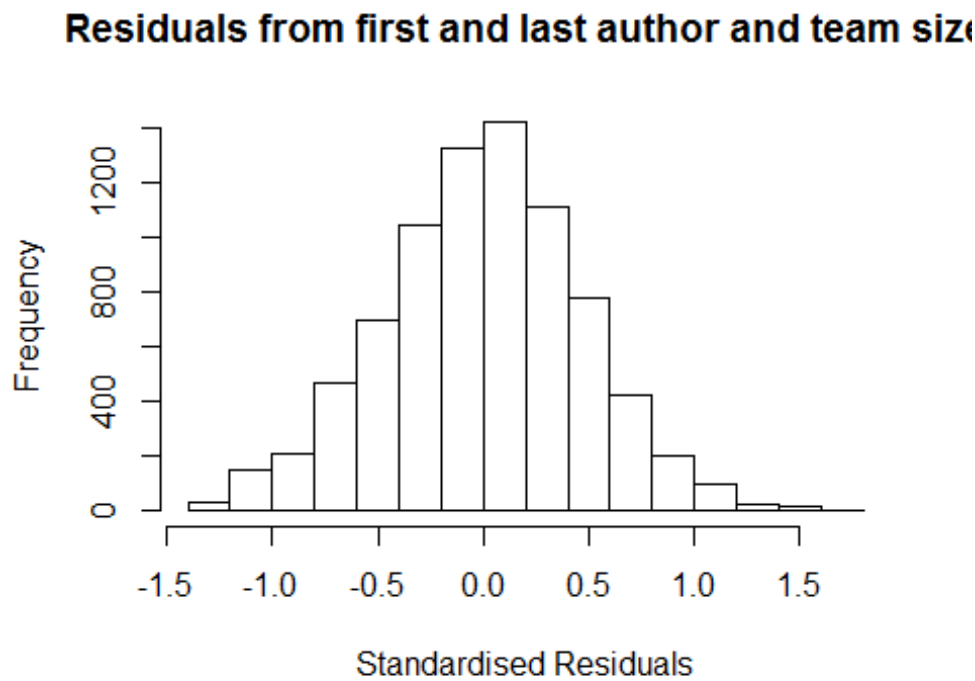
```
## 2011 2012
##  611  658
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  398  392  319  314  309  438  471  449  518  529  537  549  543  517  547
## 2011 2012
##  549  581
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
##  Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 170, df = 16, p-value <2e-16
```




```
##
##  Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 29, df = 1, p-value = 7e-08
```

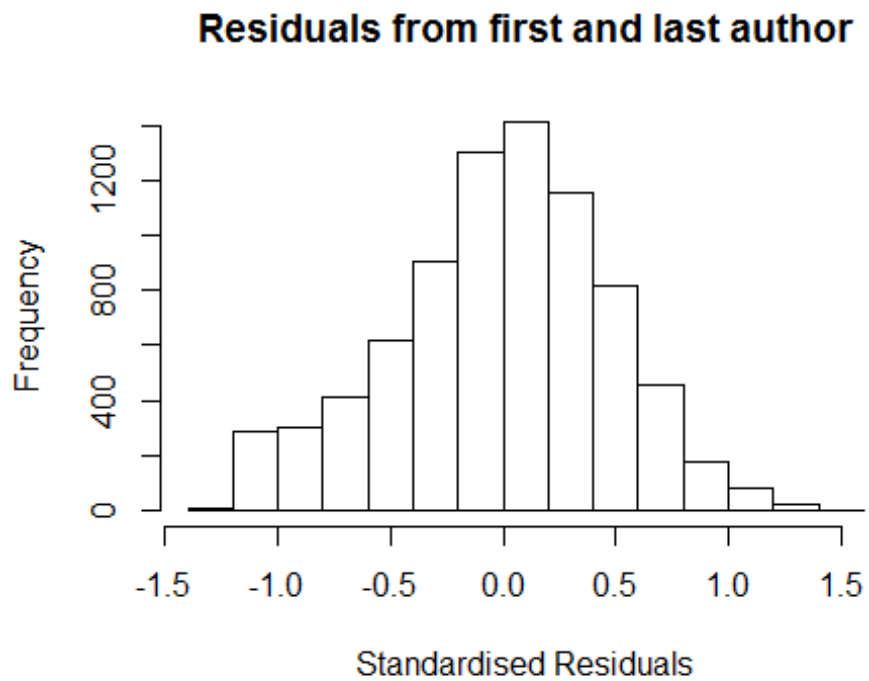



```
## [1] "Female first author team size 2018 geometric mean: 3.73510481922906"
## [1] "Male first author team size 2018 geometric mean: 3.27317332619426"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 50000, p-value = 0.003
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 3.35037270856445"
## [1] "Male last author team size 2018 geometric mean: 3.60967991857923"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 39000, p-value = 0.1
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##                     GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.033  1           1.017
## LastAuthorFemale  1.029  1           1.015
## UniqueAuthors     1.088  4           1.011
## Year              1.087 16           1.003
```



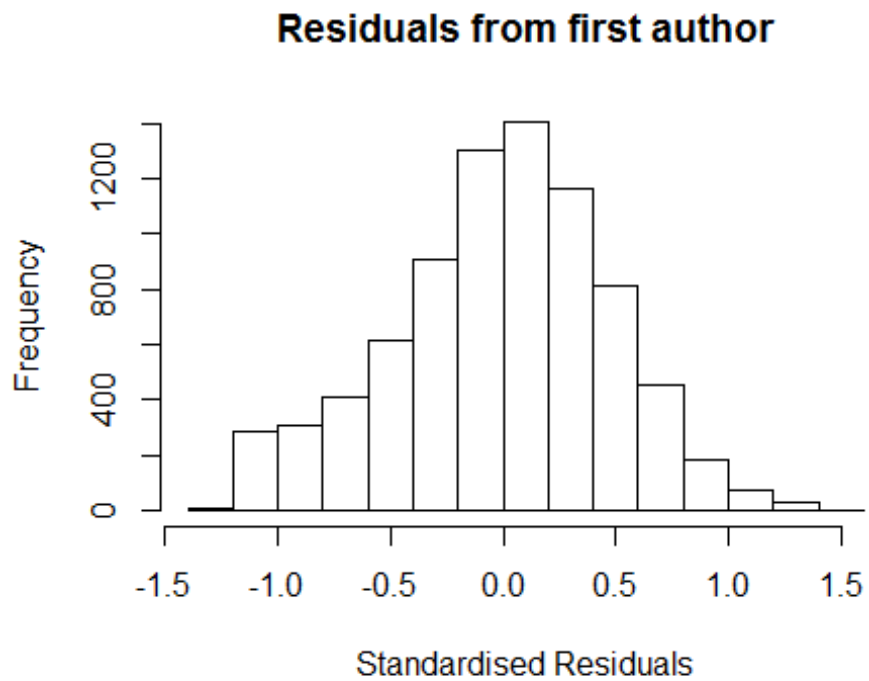

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
UniqueAuthors +
##     Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##     k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.3396 -0.3054  0.0104  0.3055  1.7147
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)          0.84663    0.03340   25.35  < 2e-16 ***
## FirstAuthorFemale1   0.03139    0.01086    2.89  0.00388 **
## LastAuthorFemale1    0.00278    0.01144    0.24  0.80819
## UniqueAuthors2       0.30240    0.02067   14.63  < 2e-16 ***
## UniqueAuthors3       0.36308    0.02075   17.50  < 2e-16 ***
## UniqueAuthors4       0.38014    0.02124   17.90  < 2e-16 ***
## UniqueAuthors5       0.42142    0.01993   21.15  < 2e-16 ***
## Year1997             0.07759    0.03868    2.01  0.04491 *
## Year1998             0.03744    0.04143    0.90  0.36621
## Year1999             0.00801    0.03816    0.21  0.83372
```

```
## Year2000            0.04146    0.03935    1.05  0.29199
## Year2001           -0.01521    0.03804   -0.40  0.68924
## Year2002           -0.05830    0.03947   -1.48  0.13969
## Year2003           -0.12606    0.03813   -3.31  0.00095 ***
## Year2004           -0.14187    0.03754   -3.78  0.00016 ***
## Year2005           -0.13417    0.03551   -3.78  0.00016 ***
## Year2006           -0.09147    0.03545   -2.58  0.00988 **
## Year2007           -0.05796    0.03471   -1.67  0.09503 .
## Year2008           -0.10309    0.03498   -2.95  0.00322 **
## Year2009           -0.05936    0.03456   -1.72  0.08589 .
## Year2010           -0.06734    0.03426   -1.97  0.04938 *
## Year2011           -0.05060    0.03409   -1.48  0.13784
## Year2012           -0.10878    0.03600   -3.02  0.00252 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.455
## Multiple R-squared:  0.106,  Adjusted R-squared:  0.104
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
##  707 weights are ~= 1. The remaining 7253 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.124   0.866   0.950   0.902   0.985   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.26e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.018  1           1.009
## LastAuthorFemale  1.017  1           1.008
## Year              1.026 16           1.001
```



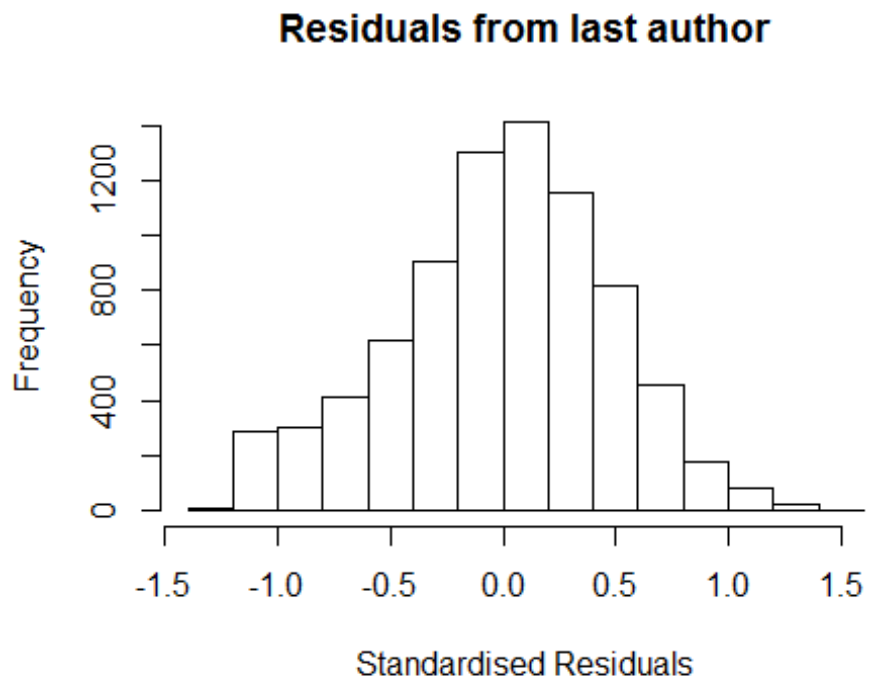

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.2655 -0.3122  0.0217  0.3195  1.4697
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.102046   0.028610   38.52  < 2e-16 ***
## FirstAuthorFemale1  0.062667   0.011179    5.61  2.1e-08 ***
## LastAuthorFemale1   0.007293   0.011745    0.62  0.53466
## Year1997            0.077645   0.037226    2.09  0.03703 *
## Year1998            0.057085   0.040741    1.40  0.16121
## Year1999            0.031139   0.037196    0.84  0.40253
## Year2000            0.093483   0.038378    2.44  0.01488 *
## Year2001            0.010293   0.038111    0.27  0.78710
## Year2002           -0.049934   0.039956   -1.25  0.21144
## Year2003           -0.125640   0.039861   -3.15  0.00163 **
## Year2004           -0.131737   0.039123   -3.37  0.00076 ***
## Year2005           -0.111117   0.036807   -3.02  0.00254 **
```

```
## Year2006           -0.061336   0.034749   -1.77  0.07758 .
## Year2007            0.004169   0.033355    0.12  0.90053
## Year2008           -0.054175   0.034232   -1.58  0.11355
## Year2009           -0.000478   0.033401   -0.01  0.98857
## Year2010           -0.004465   0.033122   -0.13  0.89277
## Year2011            0.008475   0.033004    0.26  0.79736
## Year2012           -0.043876   0.035023   -1.25  0.21032
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.462
## Multiple R-squared:  0.0202, Adjusted R-squared:  0.018
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
##  659 weights are ~= 1. The remaining 7301 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.291   0.859   0.949   0.898   0.986   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.26e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.015  1           1.007
## Year              1.015 16           1.000
```




```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.2613 -0.3119  0.0214  0.3194  1.4681
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.103800   0.028307   38.99  < 2e-16 ***
## FirstAuthorFemale1  0.064102   0.011263    5.69  1.3e-08 ***
## Year1997            0.077649   0.037210    2.09  0.03694 *
## Year1998            0.056663   0.040703    1.39  0.16393
## Year1999            0.030780   0.037150    0.83  0.40739
## Year2000            0.093366   0.038344    2.43  0.01492 *
## Year2001            0.010598   0.038101    0.28  0.78091
## Year2002           -0.049634   0.039947   -1.24  0.21409
## Year2003           -0.125885   0.039849   -3.16  0.00159 **
## Year2004           -0.131911   0.039108   -3.37  0.00075 ***
## Year2005           -0.111133   0.036794   -3.02  0.00253 **
## Year2006           -0.061129   0.034743   -1.76  0.07854 .
```

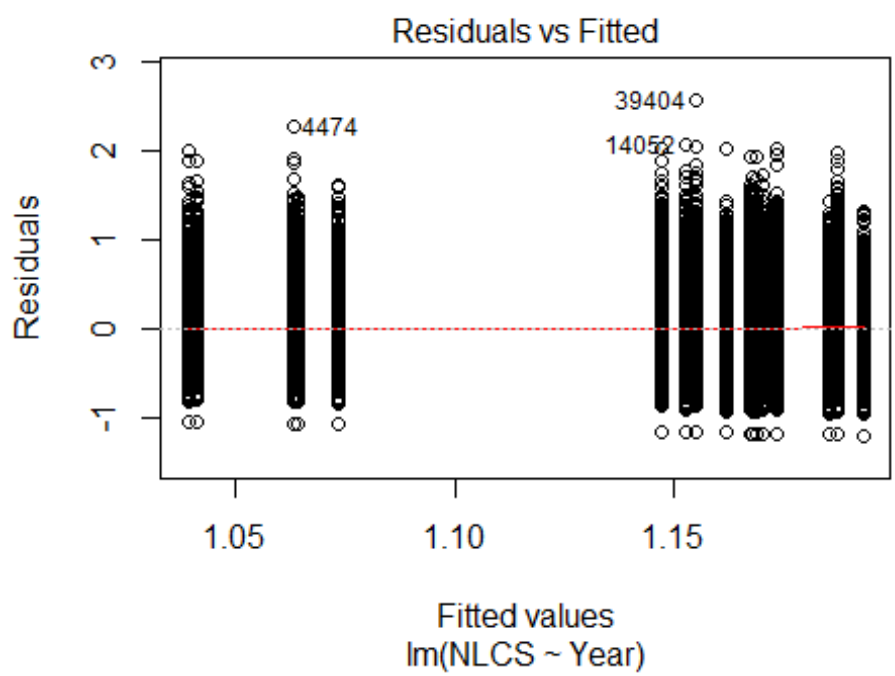
```
## Year2007            0.004236   0.033343    0.13  0.89890
## Year2008           -0.054118   0.034213   -1.58  0.11374
## Year2009           -0.000379   0.033386   -0.01  0.99093
## Year2010           -0.004533   0.033109   -0.14  0.89110
## Year2011            0.008580   0.032992    0.26  0.79483
## Year2012           -0.043322   0.035007   -1.24  0.21593
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.462
## Multiple R-squared:  0.0201, Adjusted R-squared:  0.018
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
##  665 weights are ~= 1. The remaining 7295 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.293   0.858   0.949   0.898   0.986   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.26e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##                   GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.016  1           1.008
## Year             1.016 16           1.000
```



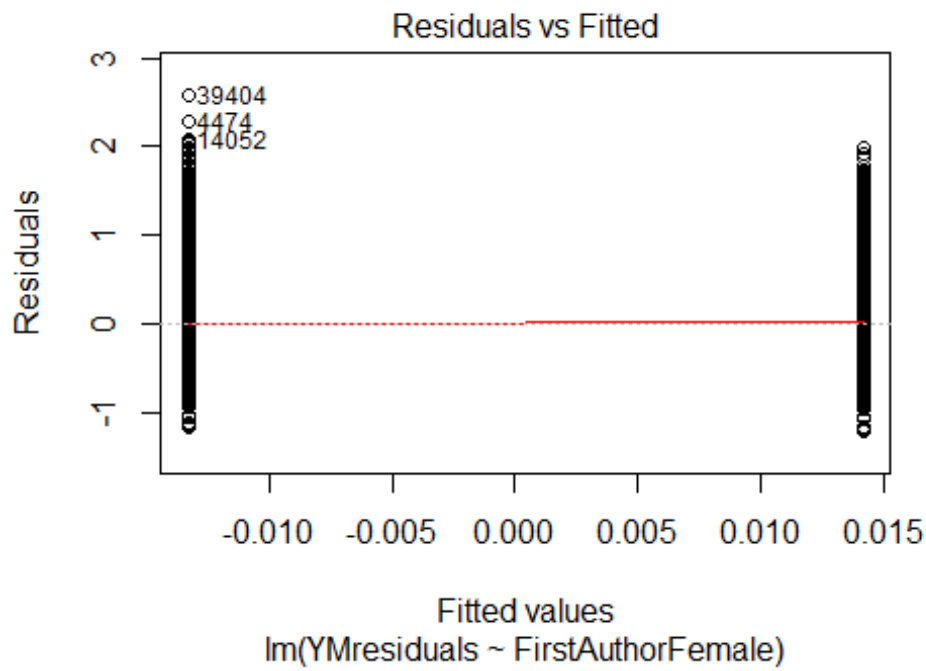

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.2346 -0.3107  0.0217  0.3198  1.4499
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.12351    0.02822   39.81  < 2e-16 ***
## LastAuthorFemale1   0.02072    0.01184    1.75  0.08013 .
## Year1997            0.07360    0.03725    1.98  0.04821 *
## Year1998            0.05557    0.04086    1.36  0.17382
## Year1999            0.02858    0.03704    0.77  0.44046
## Year2000            0.09035    0.03848    2.35  0.01891 *
## Year2001            0.00714    0.03812    0.19  0.85150
## Year2002           -0.05353    0.03998   -1.34  0.18062
## Year2003           -0.13101    0.04004   -3.27  0.00107 **
## Year2004           -0.13345    0.03921   -3.40  0.00067 ***
## Year2005           -0.11104    0.03699   -3.00  0.00270 **
## Year2006           -0.06151    0.03491   -1.76  0.07813 .
```

```
## Year2007           0.00333    0.03342    0.10  0.92063
## Year2008          -0.05286    0.03423   -1.54  0.12253
## Year2009           0.00497    0.03340    0.15  0.88177
## Year2010          -0.00369    0.03312   -0.11  0.91126
## Year2011           0.00989    0.03304    0.30  0.76460
## Year2012          -0.04300    0.03508   -1.23  0.22028
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.462
## Multiple R-squared:  0.0162, Adjusted R-squared:  0.0141
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
##  655 weights are ~= 1. The remaining 7305 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.305   0.859   0.949   0.898   0.986   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.26e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Sample size for the above analysis:  7960"
## [1] ""
## [1] ""
## [1] "####################################"
## [1] "Analysis of AJSC 3207"
## [1] "####################################"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 1340 1352 1349 1351 1679 1775 1697 1459 1568 1603 1978 2214 2290 2784 3037
## 2011 2012
## 3132 3180
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 1168 1169 1103 1129 1367 1321 1479 1288 1352 1366 1687 1887 1956 2431 2603
## 2011 2012
```

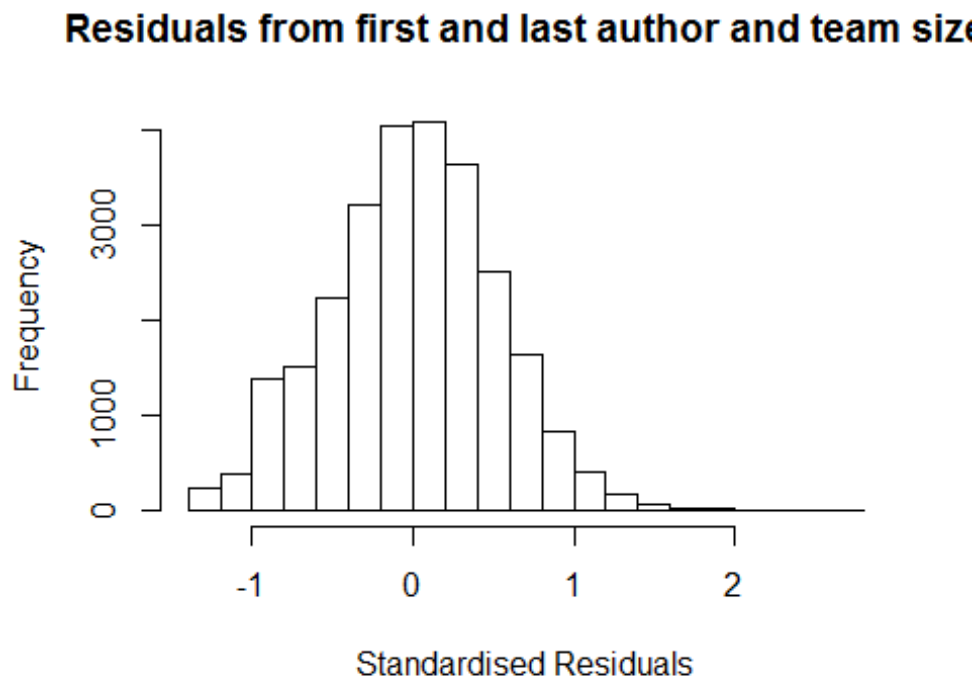
```
## 2708 2777
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 1097 1087 1027 1065 1278 1222 1360 1163 1226 1223 1519 1743 1766 2209 2348
## 2011 2012
## 2475 2520
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
##  Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 210, df = 16, p-value <2e-16
```




```
##
##  Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 80, df = 1, p-value <2e-16
```

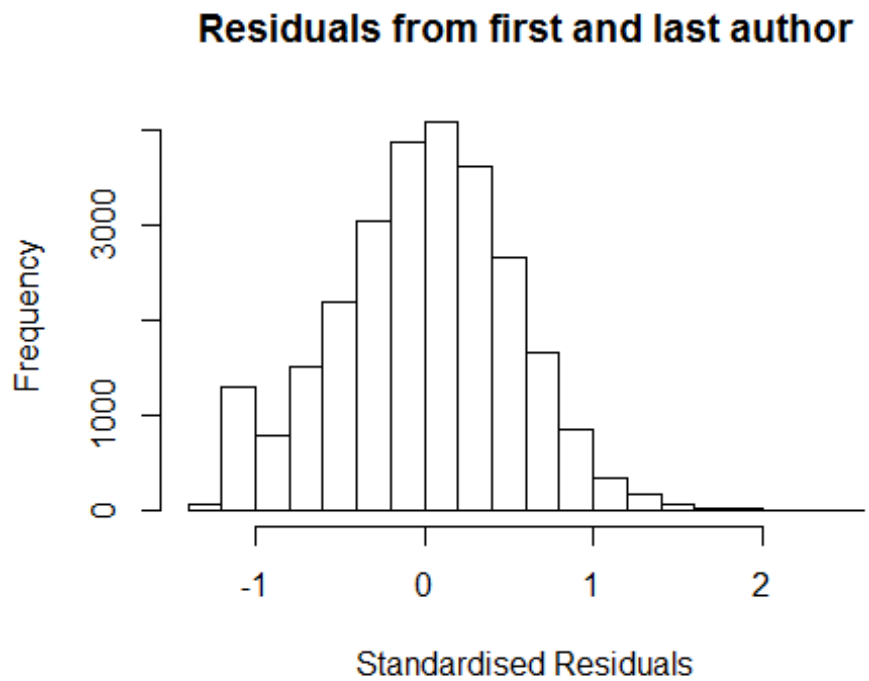



```
## [1] "Female first author team size 2018 geometric mean: 2.75911717066252"
## [1] "Male first author team size 2018 geometric mean: 2.48889186029935"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 860000, p-value = 6e-05
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 2.68017410933579"
## [1] "Male last author team size 2018 geometric mean: 2.59851364441364"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 830000, p-value = 0.2
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.124  1           1.060
## LastAuthorFemale  1.113  1           1.055
## UniqueAuthors     1.076  4           1.009
## Year              1.073 16           1.002
```



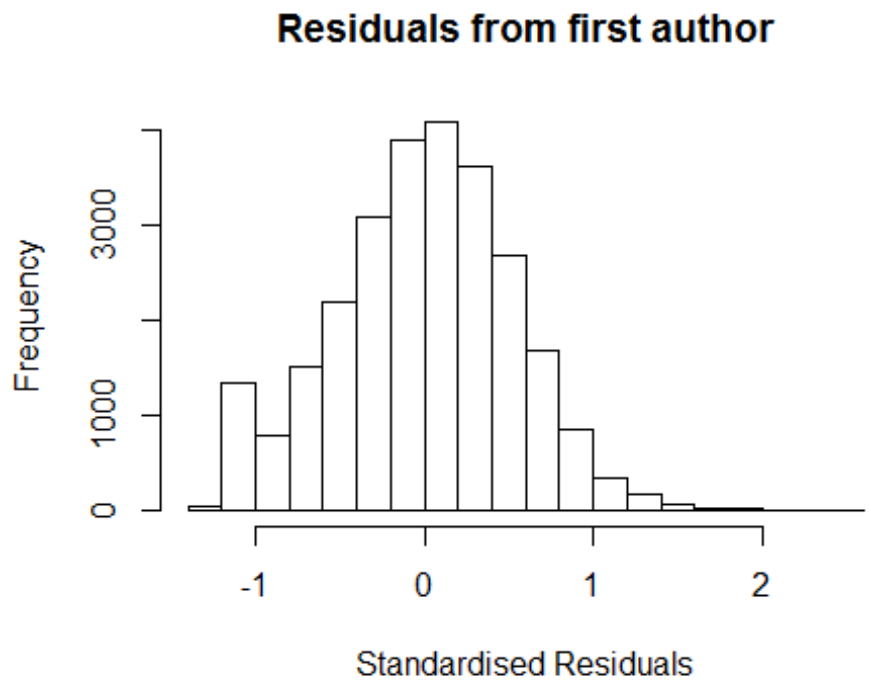

```
## [1] "List of  1 outliers with residuals above  2.5"
##          ScopusId NLCS Year OneField Fields residuals
## 39404 84858270024 3.72 2012     3207      3     2.785
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
UniqueAuthors +
##     Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##     k.max = 1000))
##  \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.35699 -0.33960  0.00943  0.34274  2.78530
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.02189    0.02038   50.14  < 2e-16 ***
## FirstAuthorFemale1  0.01670    0.00693    2.41  0.01593 *
## LastAuthorFemale1   0.00121    0.00691    0.18  0.86067
## UniqueAuthors2      0.21416    0.00898   23.84  < 2e-16 ***
## UniqueAuthors3      0.27348    0.00989   27.65  < 2e-16 ***
## UniqueAuthors4      0.31719    0.01176   26.98  < 2e-16 ***
## UniqueAuthors5      0.32129    0.01148   27.99  < 2e-16 ***
## Year1997           -0.13478    0.02659   -5.07  4.0e-07 ***
## Year1998           -0.11259    0.02712   -4.15  3.3e-05 ***
## Year1999           -0.13059    0.02544   -5.13  2.9e-07 ***
```

```
## Year2000           -0.11093    0.02413   -4.60  4.3e-06 ***
## Year2001           -0.10839    0.02509   -4.32  1.6e-05 ***
## Year2002           -0.03905    0.02451   -1.59  0.11117
## Year2003           -0.06898    0.02528   -2.73  0.00637 **
## Year2004           -0.03148    0.02454   -1.28  0.19965
## Year2005           -0.02533    0.02416   -1.05  0.29443
## Year2006           -0.06872    0.02311   -2.97  0.00295 **
## Year2007           -0.05353    0.02272   -2.36  0.01847 *
## Year2008           -0.05207    0.02266   -2.30  0.02160 *
## Year2009           -0.05692    0.02245   -2.53  0.01126 *
## Year2010           -0.05765    0.02222   -2.59  0.00949 **
## Year2011           -0.09599    0.02241   -4.28  1.8e-05 ***
## Year2012           -0.08719    0.02263   -3.85  0.00012 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.509
## Multiple R-squared:  0.0637, Adjusted R-squared:  0.063
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
##  observation 26184 is an outlier with |weight| = 0 ( < 3.8e-06);
##  2238 weights are ~= 1. The remaining 24089 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0408  0.8670  0.9510  0.9040  0.9860  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          3.80e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                    psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.109  1           1.053
## LastAuthorFemale  1.107  1           1.052
## Year              1.012 16           1.000
```



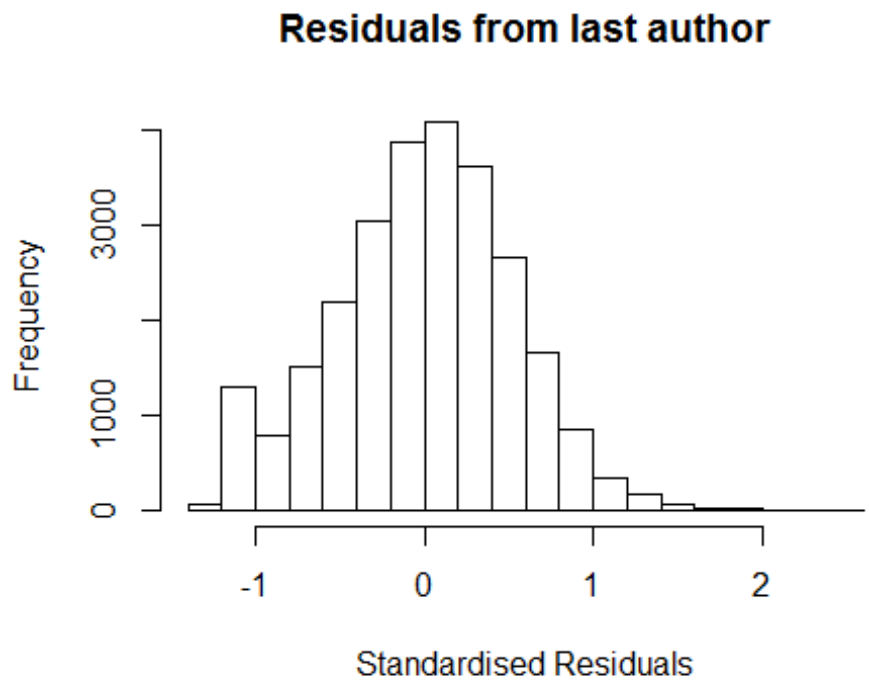

```
## [1] "List of  1 outliers with residuals above  2.5"
##          ScopusId NLCS Year OneField Fields residuals
## 39404 84858270024 3.72 2012     3207      3     2.591
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.2205 -0.3458  0.0192  0.3564  2.5911
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)          1.15818    0.02102   55.09  < 2e-16 ***
## FirstAuthorFemale1   0.03592    0.00707    5.08  3.8e-07 ***
## LastAuthorFemale1    0.00659    0.00707    0.93    0.351
## Year1997            -0.14852    0.02762   -5.38  7.6e-08 ***
## Year1998            -0.12437    0.02822   -4.41  1.0e-05 ***
## Year1999            -0.14352    0.02645   -5.43  5.8e-08 ***
## Year2000            -0.10223    0.02518   -4.06  4.9e-05 ***
## Year2001            -0.10404    0.02614   -3.98  6.9e-05 ***
## Year2002            -0.01214    0.02569   -0.47    0.636
## Year2003            -0.03754    0.02644   -1.42    0.156
## Year2004             0.01314    0.02560    0.51    0.608
## Year2005             0.01977    0.02517    0.79    0.432
```

```
## Year2006           -0.01697    0.02413   -0.70    0.482
## Year2007           -0.00768    0.02380   -0.32    0.747
## Year2008           -0.00579    0.02380   -0.24    0.808
## Year2009           -0.00212    0.02363   -0.09    0.928
## Year2010           -0.01496    0.02336   -0.64    0.522
## Year2011           -0.03982    0.02345   -1.70    0.089 .
## Year2012           -0.02928    0.02373   -1.23    0.217
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.52
## Multiple R-squared:  0.0099, Adjusted R-squared:  0.00922
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
##  observation 26184 is an outlier with |weight| = 0 ( < 3.8e-06);
##  2240 weights are ~= 1. The remaining 24087 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0674  0.8650  0.9500  0.9030  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          3.80e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                  psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.008  1           1.004
## Year              1.008 16           1.000
```




```
## [1] "List of  1 outliers with residuals above  2.5"
##          ScopusId NLCS Year OneField Fields residuals
## 39404 84858270024 3.72 2012     3207      3     2.591
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.2183 -0.3463  0.0194  0.3560  2.5892
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)          1.15999    0.02092   55.44  < 2e-16 ***
## FirstAuthorFemale1   0.03843    0.00676    5.68  1.4e-08 ***
## Year1997            -0.14896    0.02762   -5.39  7.0e-08 ***
## Year1998            -0.12446    0.02822   -4.41  1.0e-05 ***
## Year1999            -0.14346    0.02645   -5.42  5.9e-08 ***
## Year2000            -0.10219    0.02518   -4.06  5.0e-05 ***
## Year2001            -0.10412    0.02615   -3.98  6.9e-05 ***
## Year2002            -0.01205    0.02569   -0.47     0.64
## Year2003            -0.03754    0.02644   -1.42     0.16
## Year2004             0.01318    0.02561    0.51     0.61
## Year2005             0.01986    0.02517    0.79     0.43
## Year2006            -0.01683    0.02413   -0.70     0.49
```

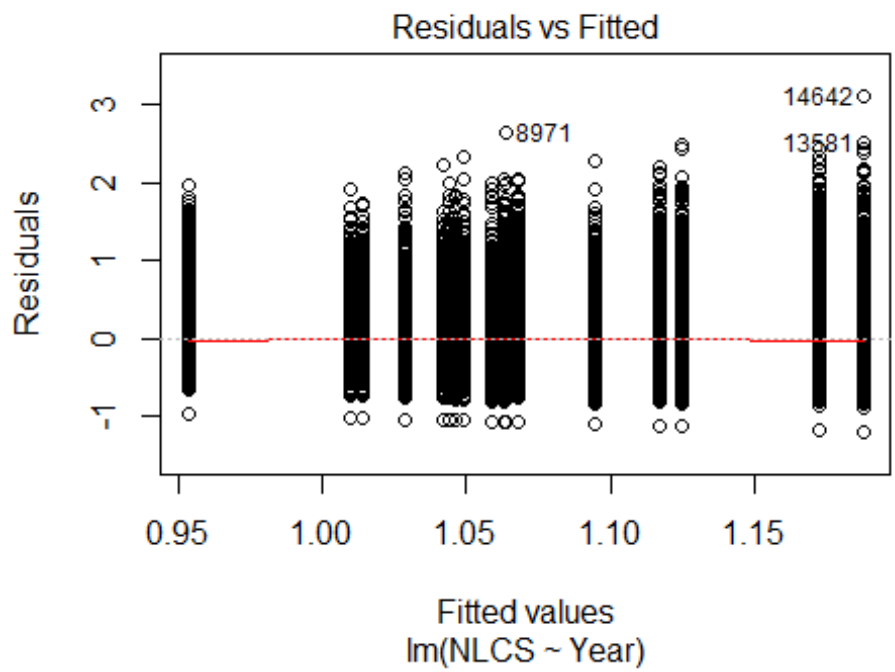
```
## Year2007           -0.00776    0.02380   -0.33     0.74
## Year2008           -0.00583    0.02380   -0.24     0.81
## Year2009           -0.00200    0.02363   -0.08     0.93
## Year2010           -0.01488    0.02336   -0.64     0.52
## Year2011           -0.03980    0.02345   -1.70     0.09 .
## Year2012           -0.02920    0.02374   -1.23     0.22
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.52
## Multiple R-squared:  0.00986,    Adjusted R-squared:  0.00922
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
##  observation 26184 is an outlier with |weight| = 0 ( < 3.8e-06);
##  2229 weights are ~= 1. The remaining 24098 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0681  0.8650  0.9500  0.9030  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          3.80e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##                   GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.006  1           1.003
## Year             1.006 16           1.000
```



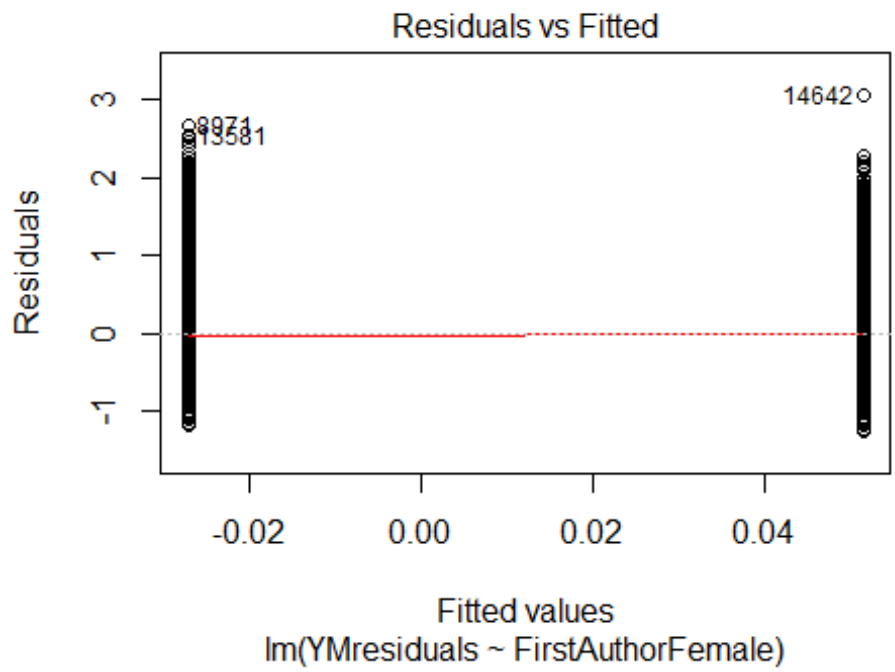

```
## [1] "List of  1 outliers with residuals above  2.5"
##          ScopusId NLCS Year OneField Fields residuals
## 39404 84858270024 3.72 2012     3207      3     2.591
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.2102 -0.3447  0.0187  0.3564  2.5783
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.16835    0.02084   56.05  < 2e-16 ***
## LastAuthorFemale1   0.02032    0.00676    3.00   0.0027 **
## Year1997           -0.14938    0.02761   -5.41  6.3e-08 ***
## Year1998           -0.12468    0.02823   -4.42  1.0e-05 ***
## Year1999           -0.14476    0.02645   -5.47  4.5e-08 ***
## Year2000           -0.10191    0.02516   -4.05  5.1e-05 ***
## Year2001           -0.10383    0.02613   -3.97  7.1e-05 ***
## Year2002           -0.01189    0.02566   -0.46   0.6432
## Year2003           -0.03607    0.02644   -1.36   0.1725
## Year2004            0.01432    0.02559    0.56   0.5756
## Year2005            0.02154    0.02516    0.86   0.3918
## Year2006           -0.01597    0.02411   -0.66   0.5078
```

```
## Year2007          -0.00565    0.02379   -0.24   0.8121
## Year2008          -0.00361    0.02379   -0.15   0.8793
## Year2009          -0.00107    0.02360   -0.05   0.9640
## Year2010          -0.01331    0.02334   -0.57   0.5684
## Year2011          -0.03774    0.02344   -1.61   0.1073
## Year2012          -0.02665    0.02371   -1.12   0.2609
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.52
## Multiple R-squared:  0.00894,    Adjusted R-squared:  0.0083
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
##  observation 26184 is an outlier with |weight| = 0 ( < 3.8e-06);
##  2255 weights are ~= 1. The remaining 24072 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0722  0.8650  0.9500  0.9030  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          3.80e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Sample size for the above analysis:  26328"
## [1] ""
## [1] ""
## [1] "####################################"
## [1] "Analysis of AJSC 3300"
## [1] "####################################"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  807  628  722  712  773  820  690  646  635  695  852  942 1091 1206 1174
## 2011 2012
## 1294 1357
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  703  531  620  607  609  606  585  558  543  585  740  816  922 1022 1012
```

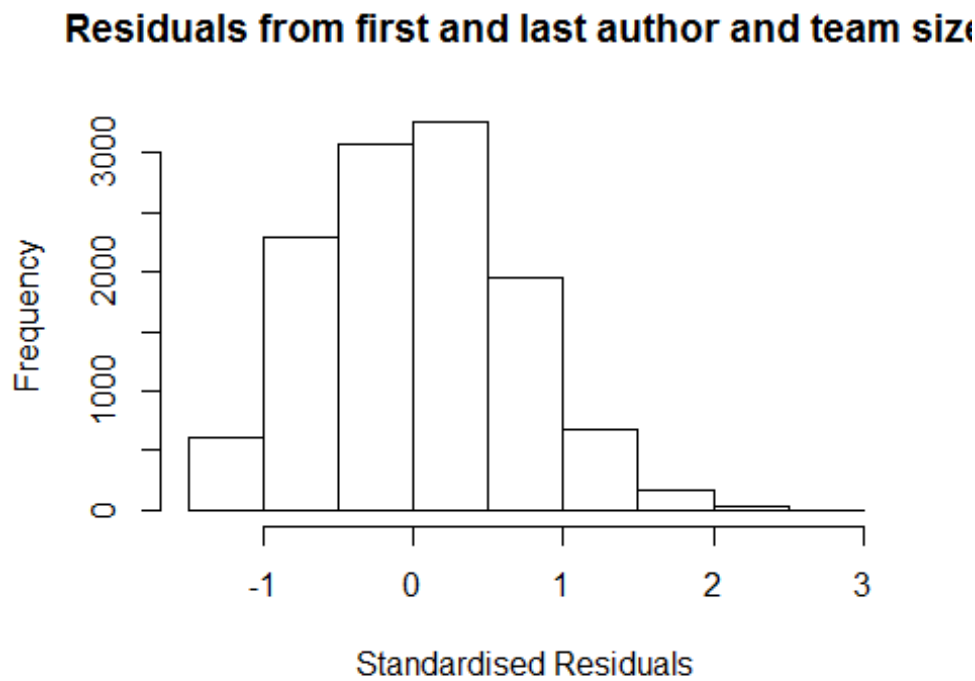
```
## 2011 2012
## 1088 1156
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  682  508  595  585  583  579  549  532  524  558  705  777  870  962  959
## 2011 2012
## 1022 1089
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
##  Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 210, df = 16, p-value <2e-16
```




```
##
##  Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 0.001, df = 1, p-value = 1
```

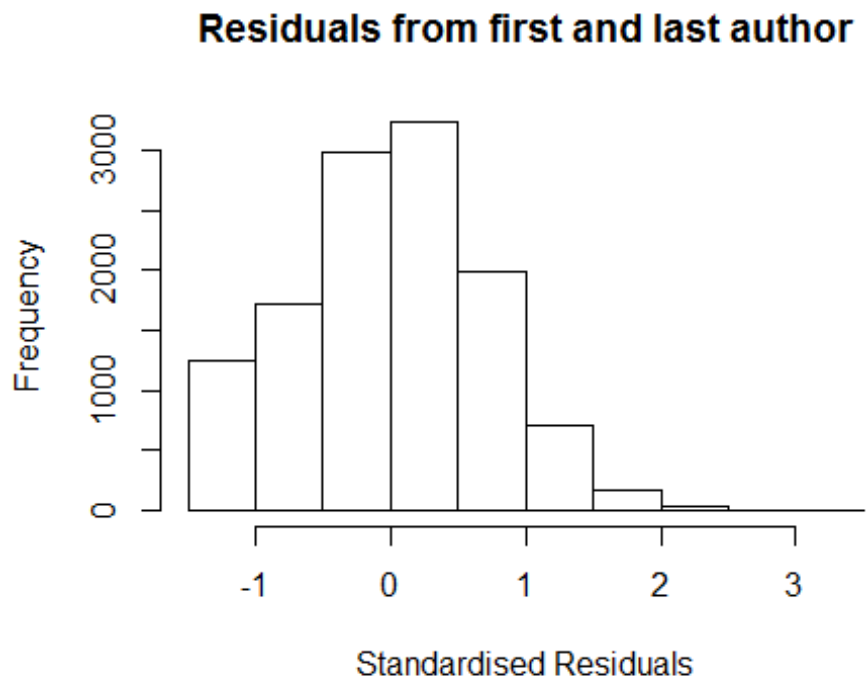



```
## [1] "Female first author team size 2018 geometric mean: 1.65415907983986"
## [1] "Male first author team size 2018 geometric mean: 1.47488952255926"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 150000, p-value = 0.001
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 1.60819390632637"
## [1] "Male last author team size 2018 geometric mean: 1.51014490690081"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 140000, p-value = 0.04
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##                     GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.912  1           1.383
## LastAuthorFemale  1.915  1           1.384
## UniqueAuthors     1.047  4           1.006
## Year              1.038 16           1.001
```



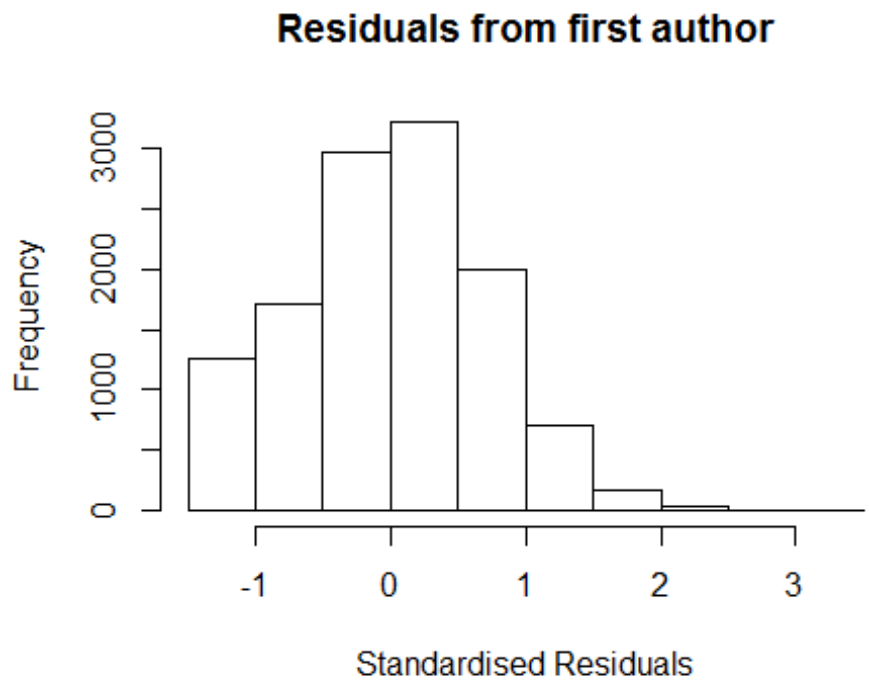

```
## [1] "List of  5 outliers with residuals above  2.5"
##          ScopusId  NLCS Year OneField Fields residuals
## 8971  34548599794 3.702 2007     2613      2     2.797
## 11504 64649093192 3.570 2009     3300      1     2.616
## 12054 70849093912 3.618 2009     3300      2     2.664
## 14642 78649757846 4.295 2011     3300      1     2.856
## 15389 84872107429 3.611 2012     1200      2     2.628
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
UniqueAuthors +
##     Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##     k.max = 1000))
##  \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.48660 -0.47587  0.00814  0.46741  2.85632
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)          0.8195     0.0273   29.98  < 2e-16 ***
## FirstAuthorFemale1   0.0521     0.0184    2.82  0.00475 **
## LastAuthorFemale1    0.0211     0.0183    1.15  0.24820
## UniqueAuthors2       0.2677     0.0155   17.26  < 2e-16 ***
## UniqueAuthors3       0.3810     0.0228   16.72  < 2e-16 ***
## UniqueAuthors4       0.3915     0.0345   11.34  < 2e-16 ***
```

```
## UniqueAuthors5       0.4078     0.0387   10.53  < 2e-16 ***
## Year1997             0.0527     0.0378    1.39  0.16305
## Year1998             0.0638     0.0369    1.73  0.08367 .
## Year1999             0.0992     0.0381    2.60  0.00925 **
## Year2000             0.0784     0.0363    2.16  0.03095 *
## Year2001             0.0942     0.0366    2.58  0.01000 *
## Year2002             0.0567     0.0375    1.51  0.13071
## Year2003             0.0820     0.0408    2.01  0.04470 *
## Year2004             0.0729     0.0388    1.88  0.06026 .
## Year2005             0.0467     0.0393    1.19  0.23502
## Year2006             0.1140     0.0354    3.22  0.00129 **
## Year2007             0.0856     0.0362    2.36  0.01812 *
## Year2008             0.0733     0.0361    2.03  0.04235 *
## Year2009             0.1346     0.0359    3.75  0.00018 ***
## Year2010             0.1406     0.0358    3.92  8.8e-05 ***
## Year2011             0.1861     0.0374    4.98  6.5e-07 ***
## Year2012             0.1635     0.0365    4.48  7.6e-06 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.696
## Multiple R-squared:  0.06,   Adjusted R-squared:  0.0583
## Convergence in 10 IRWLS iterations
##
## Robustness weights:
##  1060 weights are ~= 1. The remaining 11019 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0546  0.8610  0.9500  0.9140  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          8.28e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##                     GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.966  1           1.402
## LastAuthorFemale  1.967  1           1.403
## Year              1.015 16           1.000
```



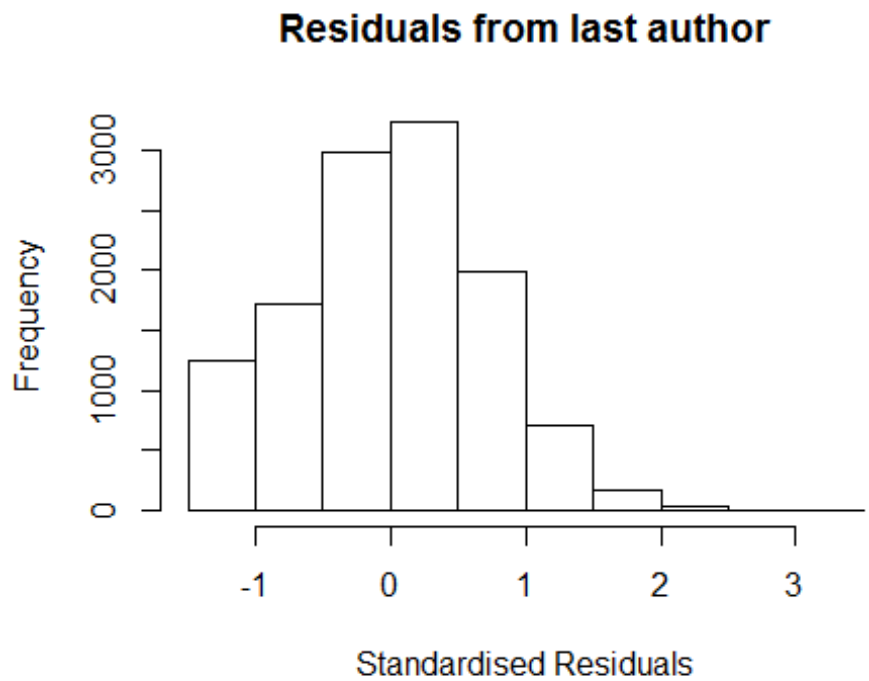

```
## [1] "List of  7 outliers with residuals above  2.5"
##          ScopusId  NLCS Year OneField Fields residuals
## 8971  34548599794 3.702 2007     2613      2     2.694
## 11504 64649093192 3.570 2009     3300      1     2.501
## 12054 70849093912 3.618 2009     3300      2     2.549
## 13581 80255137237 3.700 2011     1211      2     2.584
## 14619 79551514315 3.634 2011     1408      4     2.518
## 14642 78649757846 4.295 2011     3300      1     3.122
## 15389 84872107429 3.611 2012     1200      2     2.521
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.2248 -0.4891  0.0125  0.4788  3.1223
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)          0.8991     0.0276   32.53  < 2e-16 ***
## FirstAuthorFemale1   0.0568     0.0193    2.95   0.0032 **
## LastAuthorFemale1    0.0521     0.0191    2.73   0.0063 **
## Year1997             0.0634     0.0382    1.66   0.0973 .
## Year1998             0.0755     0.0375    2.02   0.0438 *
## Year1999             0.1087     0.0390    2.79   0.0053 **
```

```
## Year2000             0.0907     0.0371    2.44   0.0147 *
## Year2001             0.1130     0.0374    3.02   0.0025 **
## Year2002             0.0851     0.0384    2.22   0.0265 *
## Year2003             0.1019     0.0423    2.41   0.0160 *
## Year2004             0.1004     0.0401    2.50   0.0123 *
## Year2005             0.0743     0.0403    1.84   0.0652 .
## Year2006             0.1446     0.0363    3.98  6.9e-05 ***
## Year2007             0.1090     0.0374    2.92   0.0035 **
## Year2008             0.1092     0.0369    2.96   0.0031 **
## Year2009             0.1700     0.0368    4.62  3.8e-06 ***
## Year2010             0.1665     0.0367    4.53  5.8e-06 ***
## Year2011             0.2168     0.0382    5.67  1.4e-08 ***
## Year2012             0.1911     0.0376    5.08  3.7e-07 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.714
## Multiple R-squared:  0.011,  Adjusted R-squared:  0.00956
## Convergence in 10 IRWLS iterations
##
## Robustness weights:
##  997 weights are ~= 1. The remaining 11082 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0163  0.8610  0.9500  0.9140  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          8.28e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.008  1           1.004
## Year              1.008 16           1.000
```




```
## [1] "List of  7 outliers with residuals above  2.5"
##          ScopusId  NLCS Year OneField Fields residuals
## 8971  34548599794 3.702 2007     2613      2     2.694
## 11504 64649093192 3.570 2009     3300      1     2.501
## 12054 70849093912 3.618 2009     3300      2     2.549
## 13581 80255137237 3.700 2011     1211      2     2.584
## 14619 79551514315 3.634 2011     1408      4     2.518
## 14642 78649757846 4.295 2011     3300      1     3.122
## 15389 84872107429 3.611 2012     1200      2     2.521
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.2156 -0.4903  0.0122  0.4781  3.0794
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)          0.9053     0.0276   32.75  < 2e-16 ***
## FirstAuthorFemale1   0.0931     0.0138    6.75  1.5e-11 ***
## Year1997             0.0623     0.0383    1.63   0.1041
## Year1998             0.0747     0.0376    1.99   0.0467 *
## Year1999             0.1101     0.0391    2.81   0.0049 **
## Year2000             0.0895     0.0372    2.40   0.0162 *
```

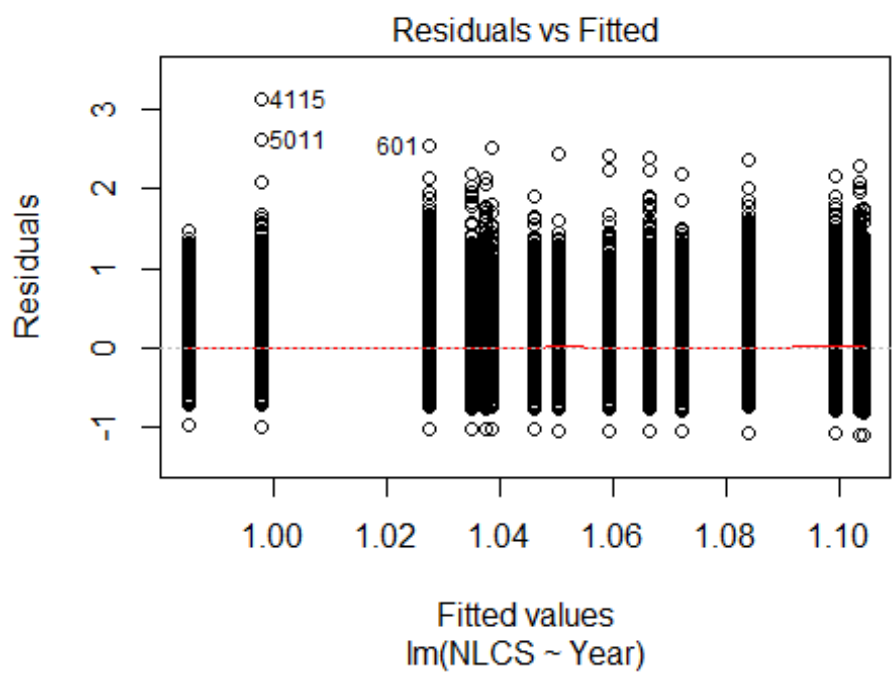
```
## Year2001             0.1115     0.0374    2.98   0.0029 **
## Year2002             0.0825     0.0384    2.15   0.0317 *
## Year2003             0.1019     0.0424    2.41   0.0162 *
## Year2004             0.0999     0.0402    2.49   0.0130 *
## Year2005             0.0734     0.0404    1.82   0.0693 .
## Year2006             0.1448     0.0364    3.98  7.1e-05 ***
## Year2007             0.1089     0.0374    2.91   0.0036 **
## Year2008             0.1085     0.0370    2.93   0.0034 **
## Year2009             0.1703     0.0369    4.62  3.9e-06 ***
## Year2010             0.1663     0.0368    4.52  6.3e-06 ***
## Year2011             0.2172     0.0383    5.67  1.4e-08 ***
## Year2012             0.1910     0.0377    5.07  4.0e-07 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.714
## Multiple R-squared:  0.0104, Adjusted R-squared:  0.00901
## Convergence in 10 IRWLS iterations
##
## Robustness weights:
##  992 weights are ~= 1. The remaining 11087 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0231  0.8590  0.9510  0.9140  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          8.28e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##                   GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.008  1           1.004
## Year             1.008 16           1.000
```



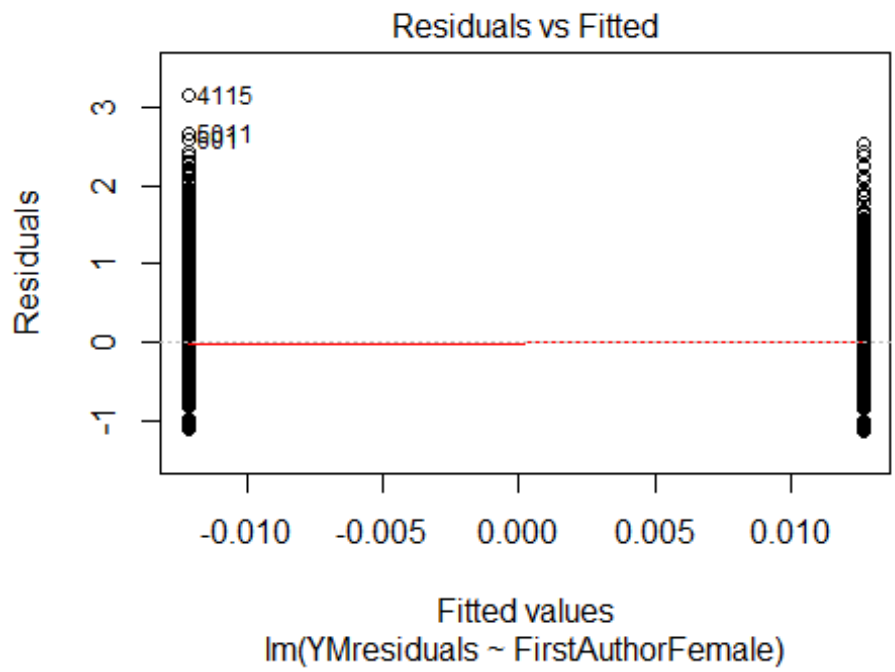

```
## [1] "List of  7 outliers with residuals above  2.5"
##          ScopusId  NLCS Year OneField Fields residuals
## 8971  34548599794 3.702 2007     2613      2     2.694
## 11504 64649093192 3.570 2009     3300      1     2.501
## 12054 70849093912 3.618 2009     3300      2     2.549
## 13581 80255137237 3.700 2011     1211      2     2.584
## 14619 79551514315 3.634 2011     1408      4     2.518
## 14642 78649757846 4.295 2011     3300      1     3.122
## 15389 84872107429 3.611 2012     1200      2     2.521
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.2134 -0.4909  0.0116  0.4780  3.1728
##
## Coefficients:
##                   Estimate Std. Error t value Pr(>|t|)
## (Intercept)         0.9033     0.0276   32.74  < 2e-16 ***
## LastAuthorFemale1   0.0912     0.0136    6.69  2.3e-11 ***
## Year1997            0.0642     0.0382    1.68   0.0927 .
## Year1998            0.0762     0.0375    2.03   0.0420 *
## Year1999            0.1088     0.0390    2.79   0.0053 **
## Year2000            0.0914     0.0372    2.46   0.0139 *
```

```
## Year2001            0.1149     0.0373    3.08   0.0021 **
## Year2002            0.0874     0.0383    2.28   0.0227 *
## Year2003            0.1031     0.0424    2.43   0.0149 *
## Year2004            0.1019     0.0401    2.54   0.0111 *
## Year2005            0.0762     0.0403    1.89   0.0584 .
## Year2006            0.1468     0.0363    4.04  5.3e-05 ***
## Year2007            0.1105     0.0374    2.96   0.0031 **
## Year2008            0.1117     0.0369    3.02   0.0025 **
## Year2009            0.1705     0.0367    4.64  3.5e-06 ***
## Year2010            0.1684     0.0367    4.59  4.5e-06 ***
## Year2011            0.2189     0.0382    5.73  1.0e-08 ***
## Year2012            0.1935     0.0375    5.15  2.6e-07 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.714
## Multiple R-squared:  0.0103, Adjusted R-squared:  0.00887
## Convergence in 10 IRWLS iterations
##
## Robustness weights:
##  977 weights are ~= 1. The remaining 11102 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0103  0.8600  0.9510  0.9150  0.9860  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          8.28e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Sample size for the above analysis:  12079"
## [1] ""
## [1] ""
## [1] "####################################"
## [1] "Analysis of AJSC 3301"
## [1] "####################################"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 1258 1130 1173 1242 1244 1371 1256 1138 1187 1229 1567 1790 1948 2117 2100
```

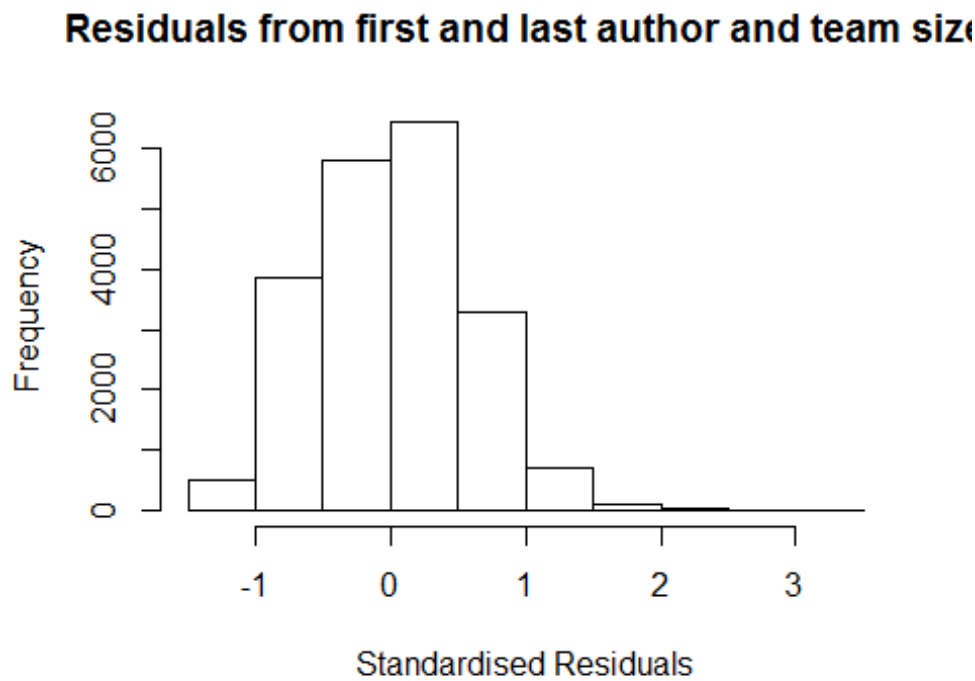
```
## 2011 2012
## 2211 2065
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 1092  976 1005 1088 1081 1145 1080  986 1037 1050 1351 1518 1673 1799 1816
## 2011 2012
## 1872 1728
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 1033  930  948 1035 1003 1083 1018  929  969  968 1245 1399 1544 1670 1686
## 2011 2012
## 1724 1589
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
##  Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 51, df = 16, p-value = 2e-05
```




```
##
##  Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 75, df = 1, p-value <2e-16
```



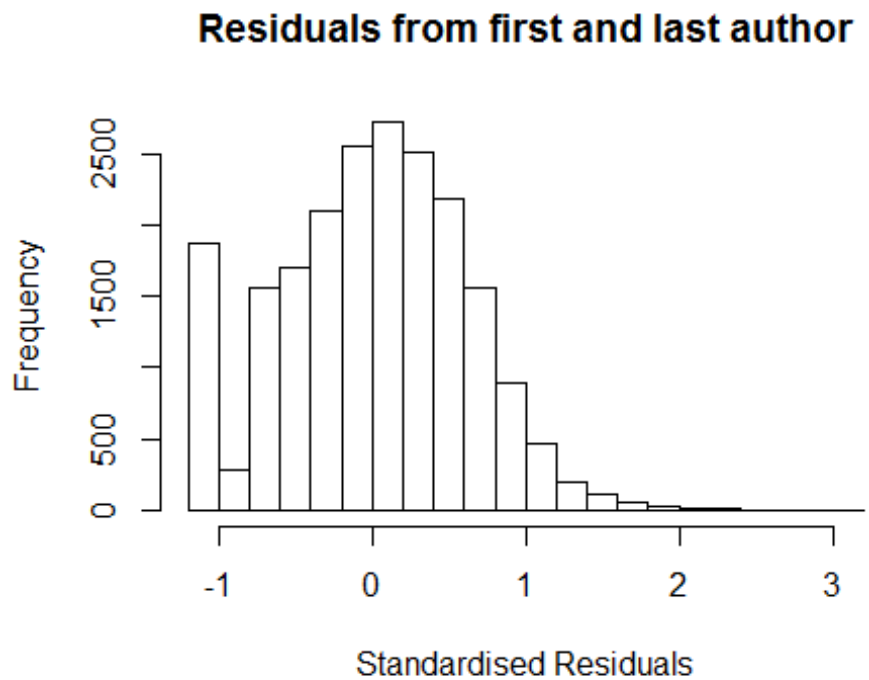

```
## [1] "Female first author team size 2018 geometric mean: 2.25574217560378"
## [1] "Male first author team size 2018 geometric mean: 1.92287985373344"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 450000, p-value = 4e-08
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 2.19878769167359"
## [1] "Male last author team size 2018 geometric mean: 2.00819228990165"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 430000, p-value = 0.002
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.384  1           1.176
## LastAuthorFemale  1.375  1           1.173
## UniqueAuthors     1.055  4           1.007
## Year              1.053 16           1.002
```




```
## [1] "List of  5 outliers with residuals above  2.5"
##          ScopusId  NLCS Year OneField Fields residuals
## 601    0001565086 3.590 1996     1212      2     2.719
## 3297   0032251946 3.572 1998     3301      2     2.674
## 4115   0000908020 4.133 1999     1401      2     3.277
## 5011  33749819408 3.636 1999     1212      2     2.780
## 21872 78049401607 3.451 2010     3301      2     2.569
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
UniqueAuthors +
##     Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##     k.max = 1000))
##  \--> method = "MM"
## Residuals:
##    Min     1Q Median     3Q    Max
## -1.308 -0.405  0.014  0.405  3.277
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)          0.87112    0.02164   40.26   <2e-16 ***
## FirstAuthorFemale1   0.01751    0.00989    1.77   0.0765 .
## LastAuthorFemale1   -0.01993    0.00983   -2.03   0.0427 *
## UniqueAuthors2       0.26744    0.01029   25.98   <2e-16 ***
## UniqueAuthors3       0.31632    0.01254   25.23   <2e-16 ***
## UniqueAuthors4       0.34724    0.01628   21.33   <2e-16 ***
```

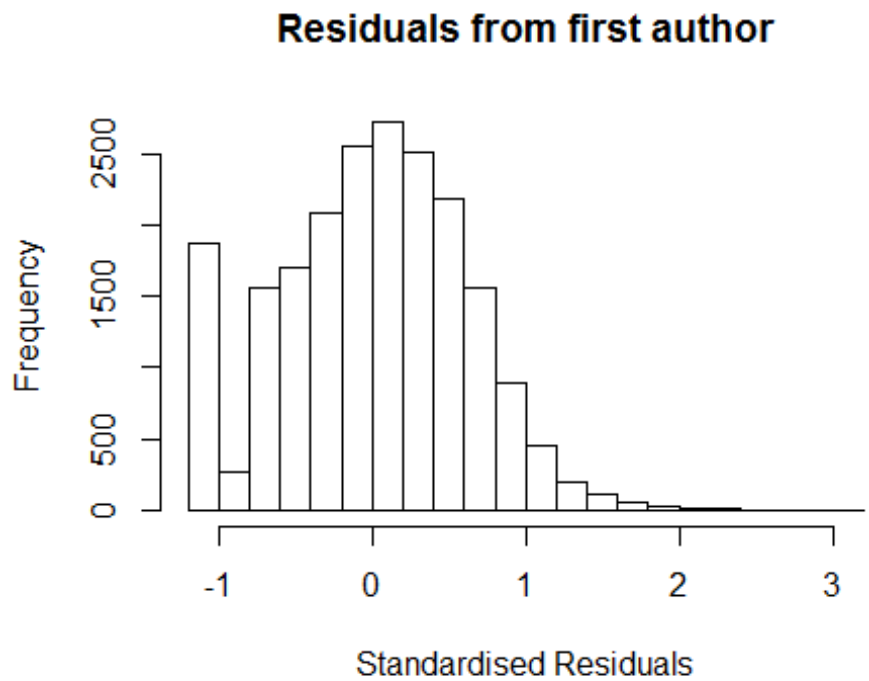
```
## UniqueAuthors5      0.34981    0.01700   20.58   <2e-16 ***
## Year1997            0.04004    0.02960    1.35   0.1762
## Year1998            0.02904    0.02924    0.99   0.3208
## Year1999           -0.01536    0.02802   -0.55   0.5835
## Year2000            0.06162    0.02808    2.19   0.0282 *
## Year2001            0.00876    0.02803    0.31   0.7546
## Year2002            0.00647    0.02803    0.23   0.8175
## Year2003           -0.02002    0.02903   -0.69   0.4903
## Year2004            0.01972    0.02806    0.70   0.4821
## Year2005            0.00488    0.02832    0.17   0.8632
## Year2006            0.01976    0.02635    0.75   0.4532
## Year2007            0.06914    0.02614    2.64   0.0082 **
## Year2008           -0.00916    0.02549   -0.36   0.7191
## Year2009            0.04584    0.02530    1.81   0.0701 .
## Year2010            0.01331    0.02580    0.52   0.6058
## Year2011            0.05094    0.02569    1.98   0.0474 *
## Year2012            0.01500    0.02601    0.58   0.5643
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.6
## Multiple R-squared:  0.0639, Adjusted R-squared:  0.0629
## Convergence in 10 IRWLS iterations
##
## Robustness weights:
##  observation 2942 is an outlier with |weight| = 0 ( < 4.8e-06);
##  1739 weights are ~= 1. The remaining 19033 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0005  0.8660  0.9500  0.9120  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          4.81e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.376  1           1.173
```

```
## LastAuthorFemale  1.377  1           1.174
## Year              1.016 16           1.001
```



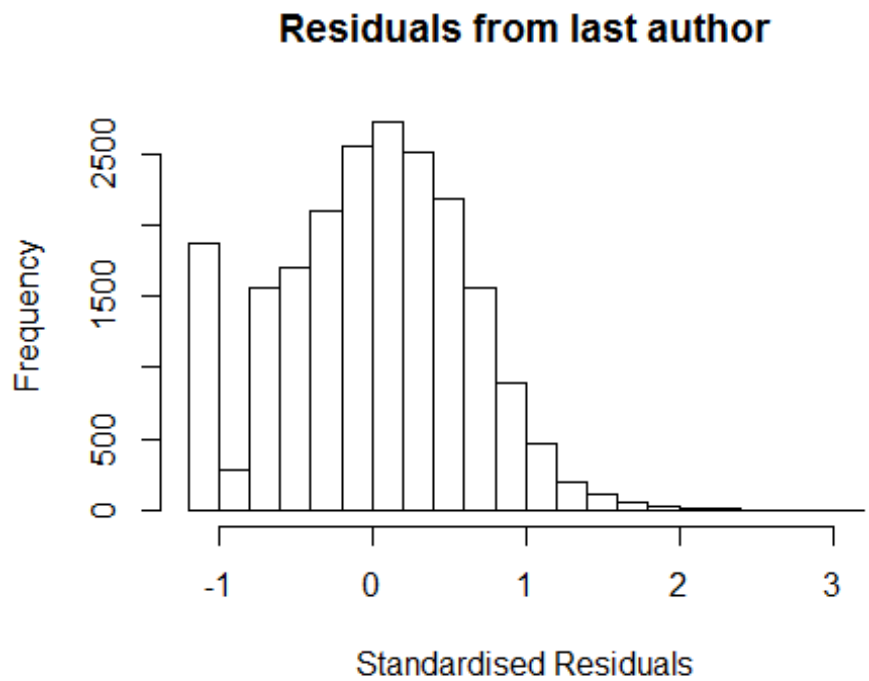

```
## [1] "List of  4 outliers with residuals above  2.5"
##         ScopusId  NLCS Year OneField Fields residuals
## 601   0001565086 3.590 1996     1212      2     2.605
## 3297  0032251946 3.572 1998     3301      2     2.525
## 4115  0000908020 4.133 1999     1401      2     3.174
## 5011 33749819408 3.636 1999     1212      2     2.677
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.1278 -0.4236  0.0256  0.4239  3.1738
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)           0.9852     0.0223   44.12  < 2e-16 ***
## FirstAuthorFemale1    0.0472     0.0102    4.62  3.9e-06 ***
## LastAuthorFemale1    -0.0135     0.0102   -1.33  0.18480
## Year1997              0.0376     0.0305    1.23  0.21711
## Year1998              0.0284     0.0299    0.95  0.34205
## Year1999             -0.0260     0.0292   -0.89  0.37273
```

```
## Year2000             0.0652     0.0290    2.25  0.02459 *
## Year2001             0.0173     0.0293    0.59  0.55463
## Year2002             0.0199     0.0293    0.68  0.49720
## Year2003            -0.0165     0.0304   -0.54  0.58755
## Year2004             0.0422     0.0292    1.45  0.14819
## Year2005             0.0242     0.0297    0.82  0.41450
## Year2006             0.0450     0.0275    1.64  0.10162
## Year2007             0.0954     0.0273    3.49  0.00048 ***
## Year2008             0.0258     0.0269    0.96  0.33701
## Year2009             0.0891     0.0265    3.36  0.00077 ***
## Year2010             0.0502     0.0270    1.86  0.06298 .
## Year2011             0.0902     0.0269    3.36  0.00079 ***
## Year2012             0.0718     0.0271    2.65  0.00799 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.618
## Multiple R-squared:  0.00445,    Adjusted R-squared:  0.00358
## Convergence in 10 IRWLS iterations
##
## Robustness weights:
##  observation 2942 is an outlier with |weight| = 0 ( < 4.8e-06);
##  1679 weights are ~= 1. The remaining 19093 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0214  0.8750  0.9490  0.9120  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          4.81e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                    psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##                   GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.01  1           1.005
## Year              1.01 16           1.000
```




```
## [1] "List of  4 outliers with residuals above  2.5"
##         ScopusId  NLCS Year OneField Fields residuals
## 601   0001565086 3.590 1996     1212      2     2.605
## 3297  0032251946 3.572 1998     3301      2     2.525
## 4115  0000908020 4.133 1999     1401      2     3.174
## 5011 33749819408 3.636 1999     1212      2     2.677
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.1174 -0.4241  0.0256  0.4248  3.1764
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         0.98277    0.02222   44.24  < 2e-16 ***
## FirstAuthorFemale1  0.03941    0.00878    4.49  7.3e-06 ***
## Year1997            0.03752    0.03048    1.23  0.21833
## Year1998            0.02846    0.02988    0.95  0.34085
## Year1999           -0.02615    0.02916   -0.90  0.36986
## Year2000            0.06485    0.02900    2.24  0.02537 *
## Year2001            0.01723    0.02926    0.59  0.55596
## Year2002            0.02004    0.02934    0.68  0.49455
## Year2003           -0.01657    0.03041   -0.54  0.58594
```

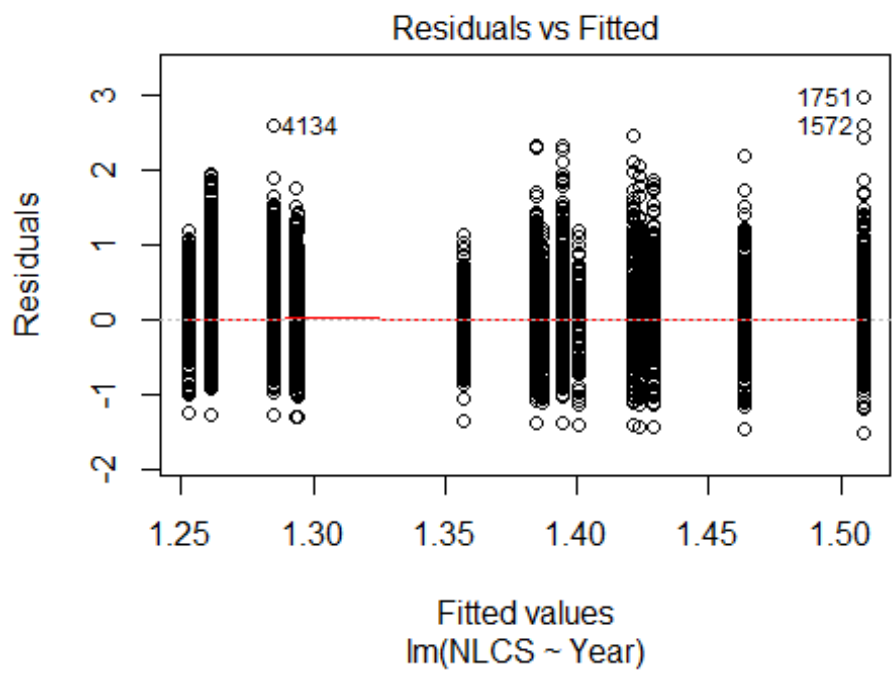
```
## Year2004            0.04222    0.02922    1.44  0.14858
## Year2005            0.02425    0.02970    0.82  0.41436
## Year2006            0.04475    0.02746    1.63  0.10318
## Year2007            0.09526    0.02730    3.49  0.00048 ***
## Year2008            0.02546    0.02691    0.95  0.34417
## Year2009            0.08897    0.02651    3.36  0.00079 ***
## Year2010            0.05037    0.02701    1.86  0.06223 .
## Year2011            0.08977    0.02687    3.34  0.00084 ***
## Year2012            0.07154    0.02708    2.64  0.00825 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.618
## Multiple R-squared:  0.00437,    Adjusted R-squared:  0.00356
## Convergence in 10 IRWLS iterations
##
## Robustness weights:
##  observation 2942 is an outlier with |weight| = 0 ( < 4.8e-06);
##  1683 weights are ~= 1. The remaining 19089 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0209  0.8750  0.9500  0.9120  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          4.81e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##                   GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.011  1           1.006
## Year             1.011 16           1.000
```



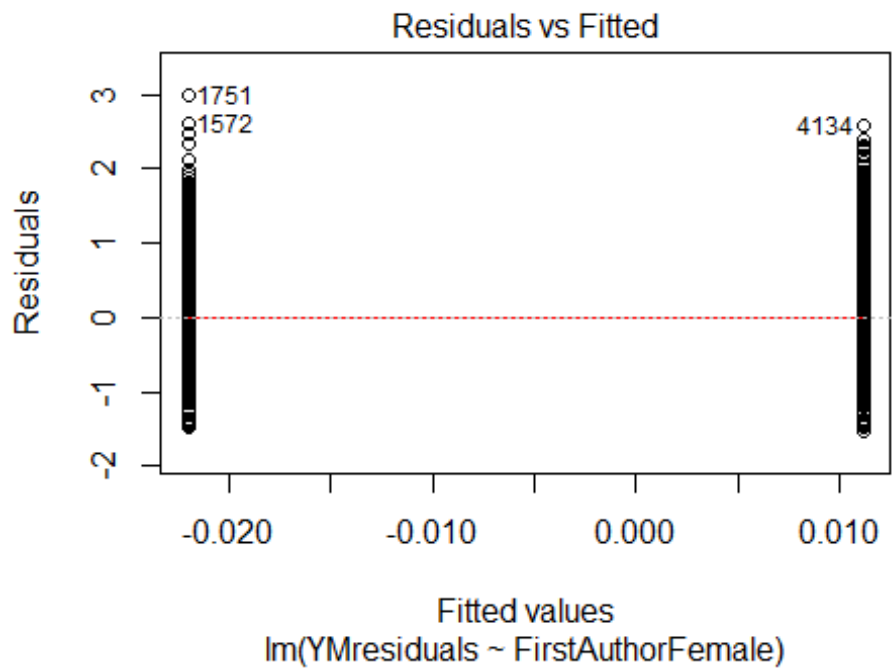

```
## [1] "List of  4 outliers with residuals above  2.5"
##         ScopusId  NLCS Year OneField Fields residuals
## 601   0001565086 3.590 1996     1212      2     2.605
## 3297  0032251946 3.572 1998     3301      2     2.525
## 4115  0000908020 4.133 1999     1401      2     3.174
## 5011 33749819408 3.636 1999     1212      2     2.677
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##    Min     1Q Median     3Q    Max
## -1.105 -0.424  0.025  0.426  3.164
##
## Coefficients:
##                   Estimate Std. Error t value Pr(>|t|)
## (Intercept)        0.99439    0.02223   44.74  < 2e-16 ***
## LastAuthorFemale1  0.01380    0.00877    1.57  0.11581
## Year1997           0.03837    0.03050    1.26  0.20835
## Year1998           0.02855    0.02990    0.95  0.33969
## Year1999          -0.02586    0.02916   -0.89  0.37514
## Year2000           0.06525    0.02901    2.25  0.02450 *
## Year2001           0.01822    0.02925    0.62  0.53344
## Year2002           0.02145    0.02936    0.73  0.46489
## Year2003          -0.01581    0.03042   -0.52  0.60321
```

```
## Year2004           0.04384    0.02925    1.50  0.13395
## Year2005           0.02588    0.02972    0.87  0.38379
## Year2006           0.04659    0.02746    1.70  0.08978 .
## Year2007           0.09706    0.02733    3.55  0.00038 ***
## Year2008           0.02741    0.02694    1.02  0.30904
## Year2009           0.09068    0.02653    3.42  0.00063 ***
## Year2010           0.05193    0.02703    1.92  0.05478 .
## Year2011           0.09166    0.02691    3.41  0.00066 ***
## Year2012           0.07360    0.02710    2.72  0.00661 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.619
## Multiple R-squared:  0.00348,    Adjusted R-squared:  0.00266
## Convergence in 10 IRWLS iterations
##
## Robustness weights:
##  observation 2942 is an outlier with |weight| = 0 ( < 4.8e-06);
##  1694 weights are ~= 1. The remaining 19078 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0241  0.8760  0.9490  0.9120  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          4.81e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Sample size for the above analysis:  20773"
## [1] ""
## [1] ""
## [1] "##################################"
## [1] "Analysis of AJSC 3302"
## [1] "##################################"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  168  197  214  198  208  273  344  355  370  342  392  393  429  383  398
## 2011 2012
##  425  478
```

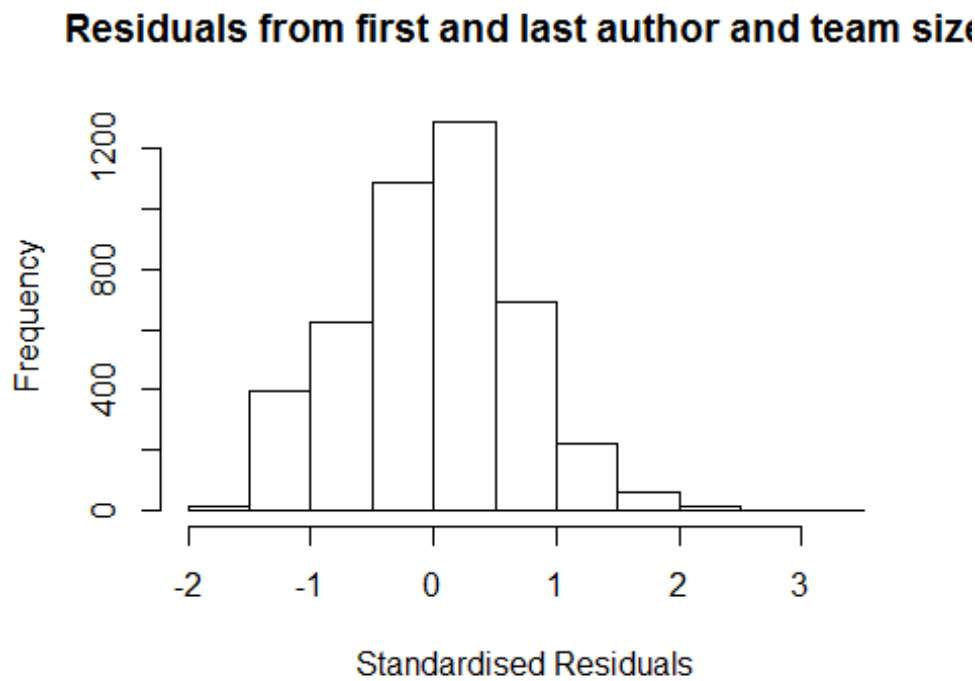
```
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  124  163  176  162  161  206  292  296  313  295  332  318  356  327  327
## 2011 2012
##  361  398
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  112  159  171  152  154  198  281  278  306  287  319  305  336  312  314
## 2011 2012
##  343  379
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
##  Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 250, df = 16, p-value <2e-16
```




```
##
##  Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 0.014, df = 1, p-value = 0.9
```

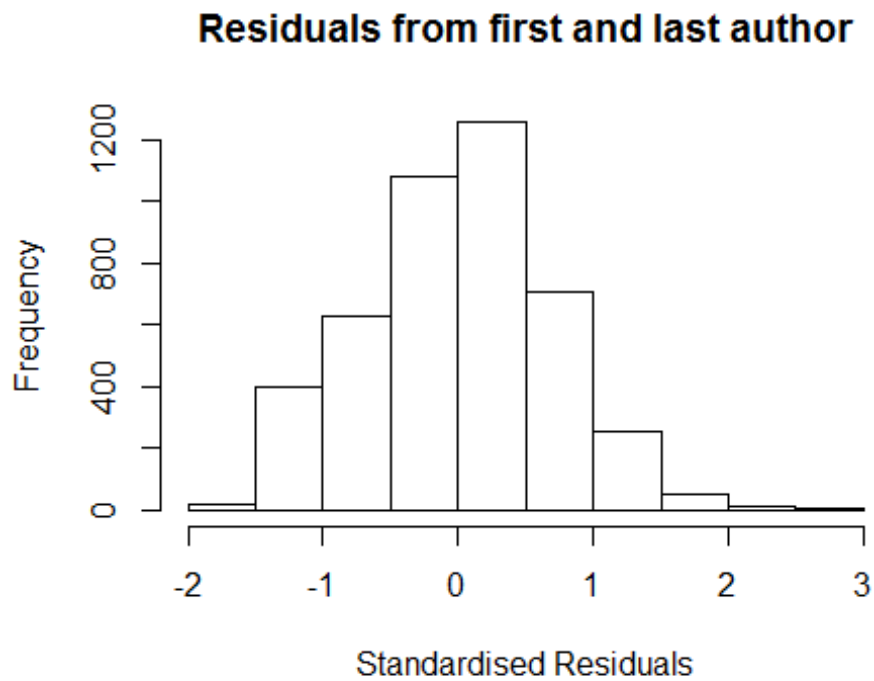



```
## [1] "Female first author team size 2018 geometric mean: 1.55820012753179"
## [1] "Male first author team size 2018 geometric mean: 1.69579960455079"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 26000, p-value = 0.2
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 1.45945662982525"
## [1] "Male last author team size 2018 geometric mean: 1.76463883297917"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 22000, p-value = 0.002
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.616  1           1.271
## LastAuthorFemale  1.617  1           1.272
## UniqueAuthors     1.115  4           1.014
## Year              1.137 16           1.004
```



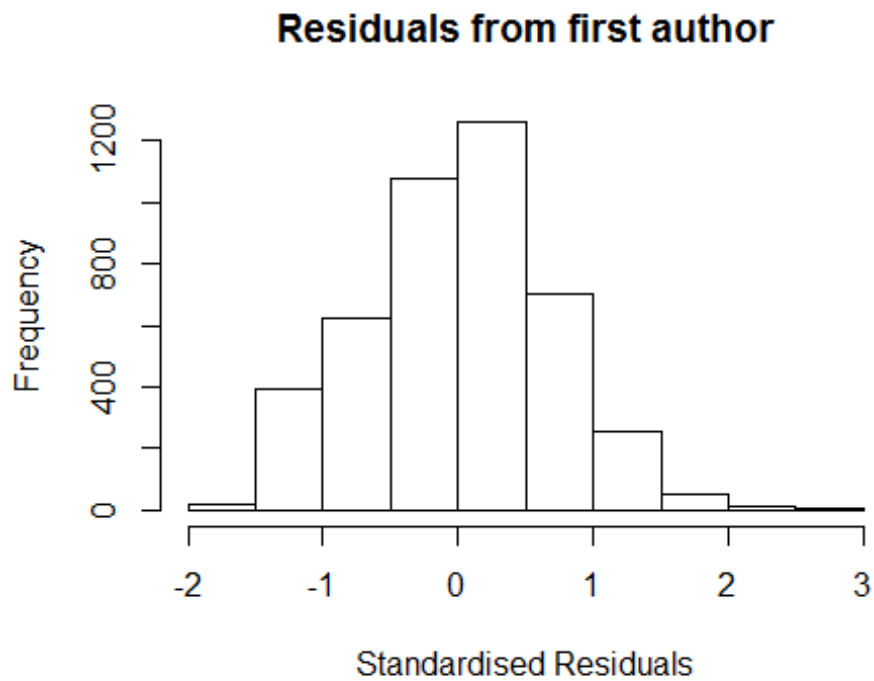

```
## [1] "List of  3 outliers with residuals above  2.5"
##         ScopusId  NLCS Year OneField Fields residuals
## 1751  0036596824 4.474 2002     1202      3     3.075
## 2108  0042885451 3.882 2003     1202      3     2.594
## 4134 46849103881 3.875 2008     1202      3     2.687
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
UniqueAuthors +
##     Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##     k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.6608 -0.4544  0.0311  0.4471  3.0750
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.32690    0.05207   25.48  < 2e-16 ***
## FirstAuthorFemale1 -0.03091    0.02864   -1.08   0.2805
## LastAuthorFemale1  -0.00309    0.02926   -0.11   0.9159
## UniqueAuthors2      0.22776    0.02702    8.43  < 2e-16 ***
## UniqueAuthors3      0.21089    0.03504    6.02  1.9e-09 ***
## UniqueAuthors4      0.28494    0.05173    5.51  3.8e-08 ***
## UniqueAuthors5      0.36129    0.04901    7.37  2.0e-13 ***
## Year1997           -0.01774    0.06737   -0.26   0.7923
```

```
## Year1998           -0.02715    0.06300   -0.43   0.6665
## Year1999           -0.12931    0.06712   -1.93   0.0541 .
## Year2000           -0.16691    0.06772   -2.46   0.0138 *
## Year2001           -0.10576    0.06258   -1.69   0.0911 .
## Year2002            0.10614    0.06794    1.56   0.1183
## Year2003           -0.00534    0.06773   -0.08   0.9372
## Year2004           -0.00604    0.06939   -0.09   0.9307
## Year2005            0.03488    0.06591    0.53   0.5967
## Year2006            0.06202    0.06366    0.97   0.3300
## Year2007            0.01699    0.06513    0.26   0.7942
## Year2008           -0.13854    0.06377   -2.17   0.0299 *
## Year2009           -0.11590    0.06647   -1.74   0.0813 .
## Year2010           -0.11369    0.06141   -1.85   0.0642 .
## Year2011           -0.21431    0.07377   -2.91   0.0037 **
## Year2012           -0.04068    0.07099   -0.57   0.5667
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.67
## Multiple R-squared:  0.0436, Adjusted R-squared:  0.0388
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
##  377 weights are ~= 1. The remaining 4029 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0016  0.8610  0.9500  0.9010  0.9870  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          2.27e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.638  1           1.280
## LastAuthorFemale  1.630  1           1.277
## Year              1.032 16           1.001
```



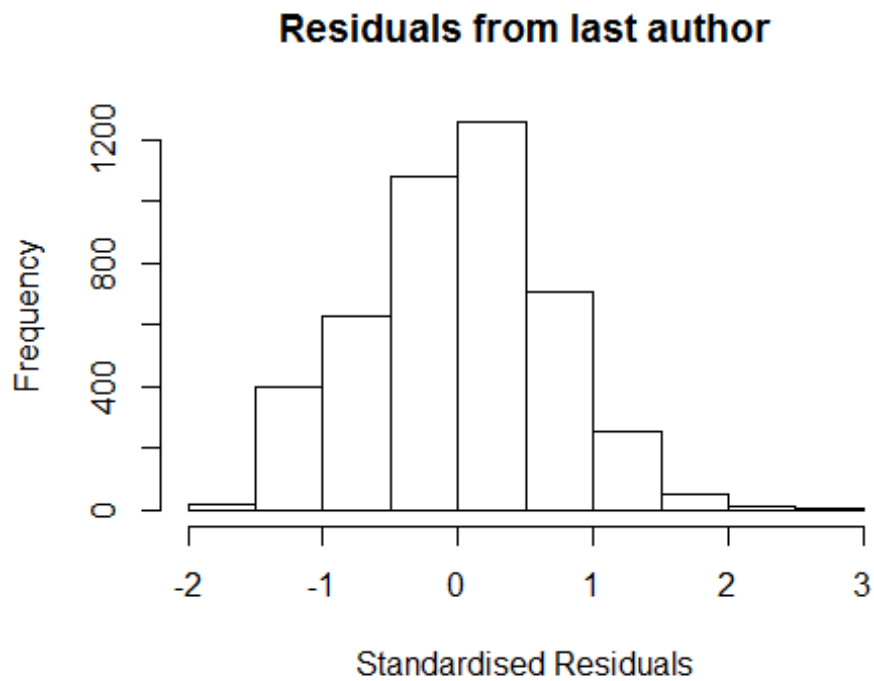

```
## [1] "List of  4 outliers with residuals above  2.5"
##         ScopusId  NLCS Year OneField Fields residuals
## 1572  0036622707 4.107 2002     1200      2     2.626
## 1751  0036596824 4.474 2002     1202      3     2.993
## 2108  0042885451 3.882 2003     1202      3     2.519
## 4134 46849103881 3.875 2008     1202      3     2.588
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.5270 -0.4682  0.0358  0.4613  2.9925
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.408047   0.051212   27.49   <2e-16 ***
## FirstAuthorFemale1 -0.027965   0.029175   -0.96    0.338
## LastAuthorFemale1  -0.017518   0.029768   -0.59    0.556
## Year1997           -0.007131   0.066968   -0.11    0.915
## Year1998           -0.032790   0.062485   -0.52    0.600
## Year1999           -0.115634   0.067110   -1.72    0.085 .
## Year2000           -0.146946   0.066676   -2.20    0.028 *
## Year2001           -0.101882   0.061915   -1.65    0.100 .
## Year2002            0.118907   0.068594    1.73    0.083 .
```

```
## Year2003            0.000478   0.068684    0.01    0.994
## Year2004           -0.006164   0.069820   -0.09    0.930
## Year2005            0.025672   0.066930    0.38    0.701
## Year2006            0.075000   0.063551    1.18    0.238
## Year2007            0.035273   0.065528    0.54    0.590
## Year2008           -0.121001   0.064469   -1.88    0.061 .
## Year2009           -0.088232   0.066529   -1.33    0.185
## Year2010           -0.098506   0.061148   -1.61    0.107
## Year2011           -0.187224   0.074669   -2.51    0.012 *
## Year2012           -0.003774   0.071250   -0.05    0.958
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.678
## Multiple R-squared:  0.0148, Adjusted R-squared:  0.0108
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
##  403 weights are ~= 1. The remaining 4003 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0126  0.8580  0.9480  0.9000  0.9860  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          2.27e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                    psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##                     GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.023  1           1.011
## Year              1.023 16           1.001
```




```
## [1] "List of  4 outliers with residuals above  2.5"
##         ScopusId  NLCS Year OneField Fields residuals
## 1572  0036622707 4.107 2002     1200      2     2.626
## 1751  0036596824 4.474 2002     1202      3     2.993
## 2108  0042885451 3.882 2003     1202      3     2.519
## 4134 46849103881 3.875 2008     1202      3     2.588
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.5247 -0.4690  0.0359  0.4651  2.9889
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.406351   0.051059   27.54   <2e-16 ***
## FirstAuthorFemale1 -0.039564   0.023133   -1.71    0.087 .
## Year1997           -0.006845   0.066909   -0.10    0.919
## Year1998           -0.032457   0.062391   -0.52    0.603
## Year1999           -0.116085   0.067122   -1.73    0.084 .
## Year2000           -0.147581   0.066651   -2.21    0.027 *
## Year2001           -0.101516   0.061851   -1.64    0.101
## Year2002            0.118334   0.068564    1.73    0.084 .
## Year2003            0.000355   0.068636    0.01    0.996
```

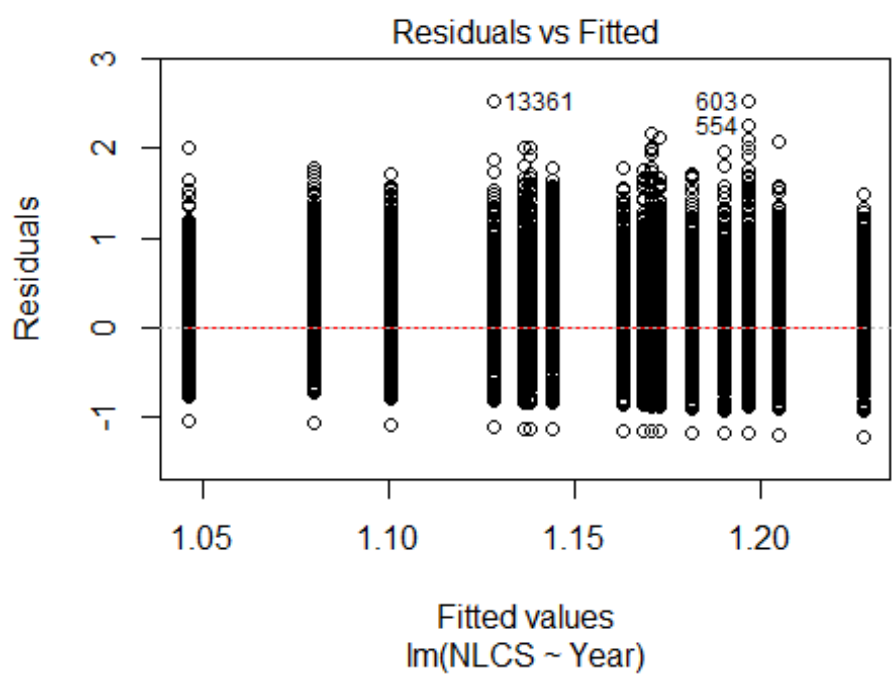
```
## Year2004           -0.006248   0.069853   -0.09    0.929
## Year2005            0.025460   0.066912    0.38    0.704
## Year2006            0.074673   0.063529    1.18    0.240
## Year2007            0.035103   0.065491    0.54    0.592
## Year2008           -0.120891   0.064436   -1.88    0.061 .
## Year2009           -0.088319   0.066513   -1.33    0.184
## Year2010           -0.098642   0.061121   -1.61    0.107
## Year2011           -0.187865   0.074683   -2.52    0.012 *
## Year2012           -0.003678   0.071267   -0.05    0.959
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.678
## Multiple R-squared:  0.0148, Adjusted R-squared:  0.0109
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
##  411 weights are ~= 1. The remaining 3995 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0131  0.8580  0.9480  0.9000  0.9860  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          2.27e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##                   GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.018  1           1.009
## Year             1.018 16           1.001
```



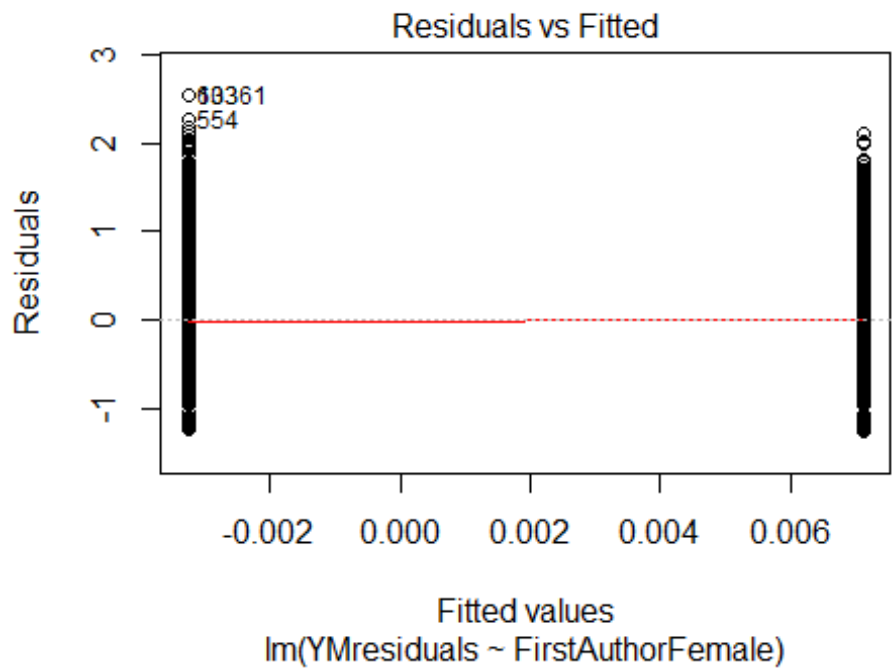

```
## [1] "List of  4 outliers with residuals above  2.5"
##         ScopusId  NLCS Year OneField Fields residuals
## 1572  0036622707 4.107 2002     1200      2     2.626
## 1751  0036596824 4.474 2002     1202      3     2.993
## 2108  0042885451 3.882 2003     1202      3     2.519
## 4134 46849103881 3.875 2008     1202      3     2.588
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.5243 -0.4703  0.0348  0.4637  2.9859
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)        1.405582   0.050998   27.56   <2e-16 ***
## LastAuthorFemale1 -0.036238   0.023577   -1.54    0.124
## Year1997          -0.006651   0.066971   -0.10    0.921
## Year1998          -0.033743   0.062470   -0.54    0.589
## Year1999          -0.116685   0.066992   -1.74    0.082 .
## Year2000          -0.146166   0.066634   -2.19    0.028 *
## Year2001          -0.103354   0.061895   -1.67    0.095 .
## Year2002           0.118751   0.068534    1.73    0.083 .
## Year2003           0.000306   0.068654    0.00    0.996
```

```
## Year2004          -0.006369   0.069711   -0.09    0.927
## Year2005           0.024224   0.066835    0.36    0.717
## Year2006           0.074638   0.063486    1.18    0.240
## Year2007           0.034149   0.065504    0.52    0.602
## Year2008          -0.122674   0.064415   -1.90    0.057 .
## Year2009          -0.089948   0.066471   -1.35    0.176
## Year2010          -0.099388   0.061157   -1.63    0.104
## Year2011          -0.188269   0.074581   -2.52    0.012 *
## Year2012          -0.005353   0.071171   -0.08    0.940
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.678
## Multiple R-squared:  0.0146, Adjusted R-squared:  0.0108
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
##  398 weights are ~= 1. The remaining 4008 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0138  0.8590  0.9480  0.9000  0.9860  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          2.27e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Sample size for the above analysis:  4406"
## [1] ""
## [1] ""
## [1] "####################################"
## [1] "Analysis of AJSC 3303"
## [1] "####################################"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  695  703  718  698  850  883  882  897  871  730  832  785  824  931 1057
## 2011 2012
## 1064 1003
##
```

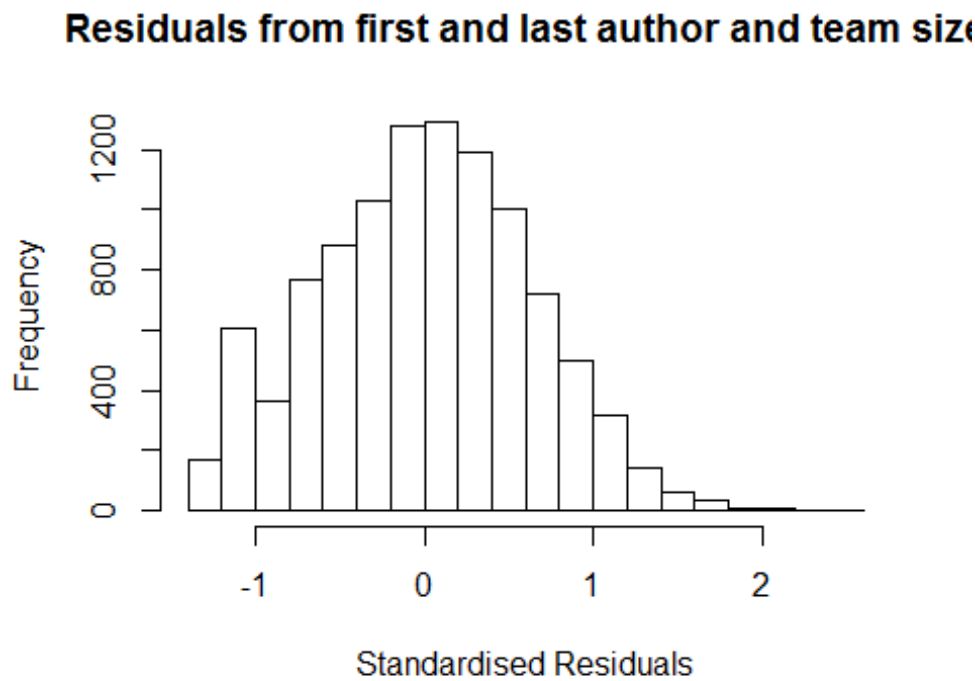
```
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  537  559  568  508  551  506  697  710  720  600  663  615  654  739  840
## 2011 2012
##  852  798
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  514  522  538  478  516  479  666  672  675  554  616  561  605  692  773
## 2011 2012
##  774  725
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
##  Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 43, df = 16, p-value = 3e-04
```




```
##
##  Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 11, df = 1, p-value = 0.001
```

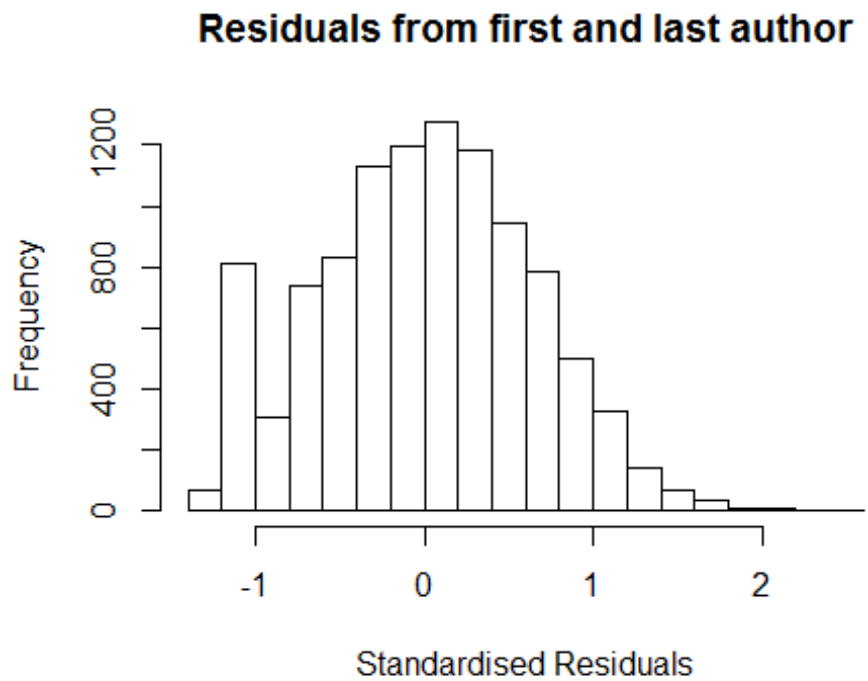



```
## [1] "Female first author team size 2018 geometric mean: 1.6803163345385"
## [1] "Male first author team size 2018 geometric mean: 1.65907002053174"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 76000, p-value = 0.8
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 1.67211837022242"
## [1] "Male last author team size 2018 geometric mean: 1.66526788069487"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 74000, p-value = 1
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.804  1           1.343
## LastAuthorFemale  1.806  1           1.344
## UniqueAuthors     1.054  4           1.007
## Year              1.071 16           1.002
```



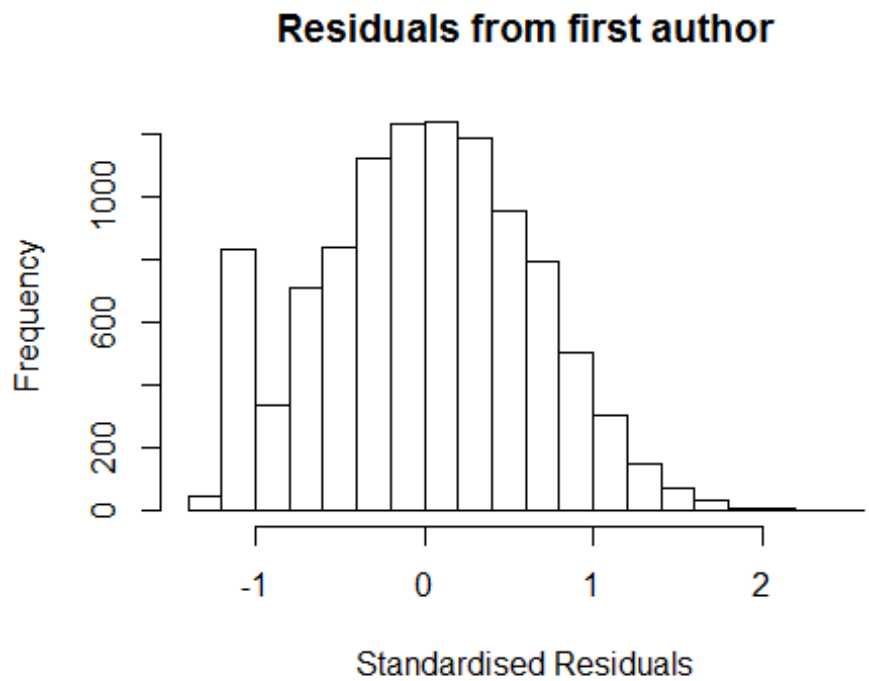

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
UniqueAuthors +
##     Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##     k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.3797 -0.4393  0.0146  0.4368  2.4386
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)          1.1107     0.0310   35.82  < 2e-16 ***
## FirstAuthorFemale1  -0.0266     0.0183   -1.45  0.14602
## LastAuthorFemale1    0.0564     0.0184    3.07  0.00212 **
## UniqueAuthors2       0.1991     0.0152   13.10  < 2e-16 ***
## UniqueAuthors3       0.2090     0.0226    9.24  < 2e-16 ***
## UniqueAuthors4       0.2468     0.0411    6.01  1.9e-09 ***
## UniqueAuthors5       0.2081     0.0469    4.44  9.2e-06 ***
## Year1997            -0.0193     0.0442   -0.44  0.66199
## Year1998            -0.0524     0.0436   -1.20  0.22938
## Year1999            -0.0402     0.0435   -0.92  0.35514
```

```
## Year2000            -0.0680     0.0422   -1.61  0.10719
## Year2001            -0.1678     0.0434   -3.87  0.00011 ***
## Year2002            -0.0177     0.0396   -0.45  0.65524
## Year2003             0.0180     0.0389    0.46  0.64373
## Year2004            -0.0129     0.0394   -0.33  0.74246
## Year2005            -0.0353     0.0407   -0.87  0.38602
## Year2006            -0.0343     0.0394   -0.87  0.38410
## Year2007             0.0236     0.0393    0.60  0.54705
## Year2008            -0.0466     0.0397   -1.17  0.24130
## Year2009            -0.1196     0.0399   -3.00  0.00273 **
## Year2010            -0.0965     0.0381   -2.53  0.01141 *
## Year2011            -0.0790     0.0390   -2.03  0.04263 *
## Year2012            -0.1396     0.0396   -3.52  0.00043 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.645
## Multiple R-squared:  0.0295, Adjusted R-squared:  0.0274
## Convergence in 10 IRWLS iterations
##
## Robustness weights:
##  884 weights are ~= 1. The remaining 9476 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.121   0.873   0.950   0.912   0.985   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          9.65e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.827  1           1.352
## LastAuthorFemale  1.827  1           1.351
## Year              1.024 16           1.001
```



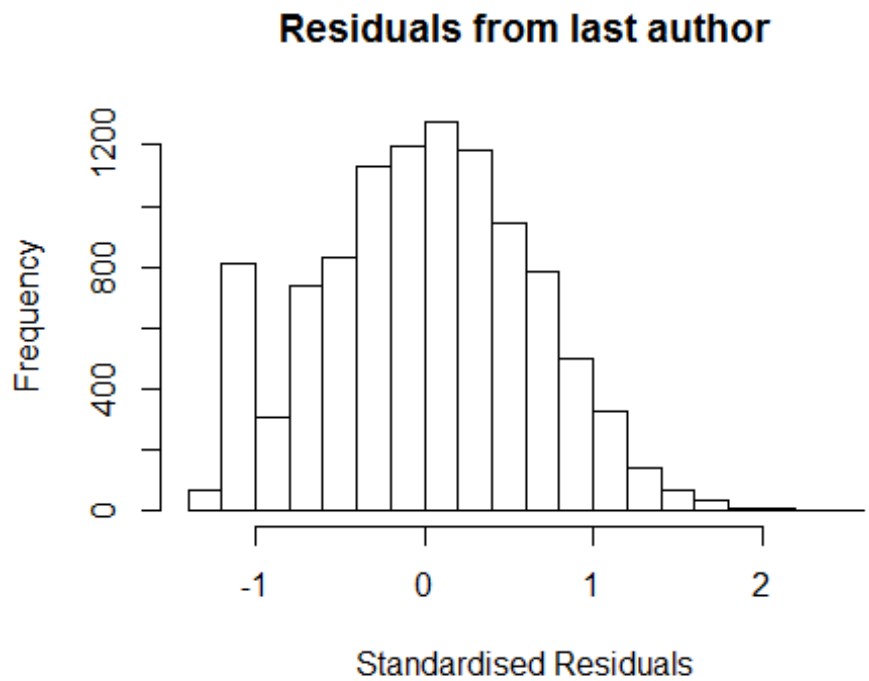

```
## [1] "List of  1 outliers with residuals above  2.5"
##          ScopusId  NLCS Year OneField Fields residuals
## 13361 77954069125 3.652 2010     3303      3     2.541
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.2460 -0.4406  0.0146  0.4403  2.5410
##
## Coefficients:
##                      Estimate Std. Error t value Pr(>|t|)
## (Intercept)          1.173378   0.030711   38.21  < 2e-16 ***
## FirstAuthorFemale1 -0.023467   0.018667   -1.26  0.20872
## LastAuthorFemale1   0.055958   0.018707    2.99  0.00278 **
## Year1997           -0.012663   0.044008   -0.29  0.77355
## Year1998           -0.058921   0.043688   -1.35  0.17747
## Year1999           -0.033925   0.043727   -0.78  0.43787
## Year2000           -0.054245   0.042391   -1.28  0.20070
## Year2001           -0.148519   0.043751   -3.39  0.00069 ***
## Year2002           -0.013958   0.039780   -0.35  0.72570
## Year2003            0.020218   0.039008    0.52  0.60426
## Year2004           -0.000056   0.039430    0.00  0.99887
## Year2005           -0.020640   0.041151   -0.50  0.61598
```

```
## Year2006           -0.020546   0.039610   -0.52  0.60397
## Year2007            0.040176   0.039267    1.02  0.30626
## Year2008           -0.028413   0.040096   -0.71  0.47858
## Year2009           -0.101005   0.040386   -2.50  0.01240 *
## Year2010           -0.062365   0.038460   -1.62  0.10493
## Year2011           -0.049854   0.039128   -1.27  0.20265
## Year2012           -0.114101   0.040016   -2.85  0.00436 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.654
## Multiple R-squared:  0.00602,    Adjusted R-squared:  0.00429
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  888 weights are ~= 1. The remaining 9472 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0979  0.8720  0.9490  0.9130  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          9.65e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.017  1           1.008
## Year              1.017 16           1.001
```




```
## [1] "List of  1 outliers with residuals above  2.5"
##          ScopusId  NLCS Year OneField Fields residuals
## 13361 77954069125 3.652 2010     3303      3     2.541
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.2334 -0.4402  0.0107  0.4399  2.5341
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)          1.17755    0.03064   38.43   <2e-16 ***
## FirstAuthorFemale1  0.01423    0.01390    1.02   0.3063
## Year1997           -0.01144    0.04393   -0.26   0.7946
## Year1998           -0.05843    0.04365   -1.34   0.1807
## Year1999           -0.03318    0.04370   -0.76   0.4478
## Year2000           -0.05050    0.04230   -1.19   0.2326
## Year2001           -0.14678    0.04376   -3.35   0.0008 ***
## Year2002           -0.01423    0.03973   -0.36   0.7203
## Year2003            0.02188    0.03896    0.56   0.5744
## Year2004            0.00144    0.03939    0.04   0.9708
## Year2005           -0.01919    0.04107   -0.47   0.6402
## Year2006           -0.02009    0.03956   -0.51   0.6115
```

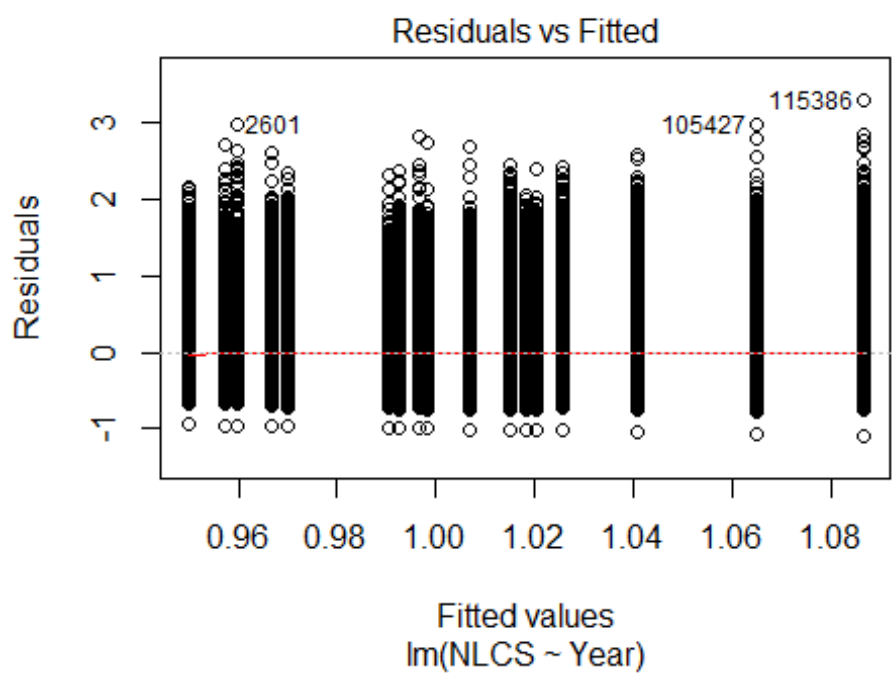
```
## Year2007            0.04159    0.03920    1.06   0.2886
## Year2008           -0.02544    0.03999   -0.64   0.5247
## Year2009           -0.09922    0.04035   -2.46   0.0139 *
## Year2010           -0.05967    0.03835   -1.56   0.1198
## Year2011           -0.04677    0.03909   -1.20   0.2315
## Year2012           -0.11332    0.04001   -2.83   0.0046 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.654
## Multiple R-squared:  0.00515,    Adjusted R-squared:  0.00351
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  883 weights are ~= 1. The remaining 9477 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.100   0.872   0.949   0.913   0.985   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          9.65e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##                   GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.016  1           1.008
## Year             1.016 16           1.001
```



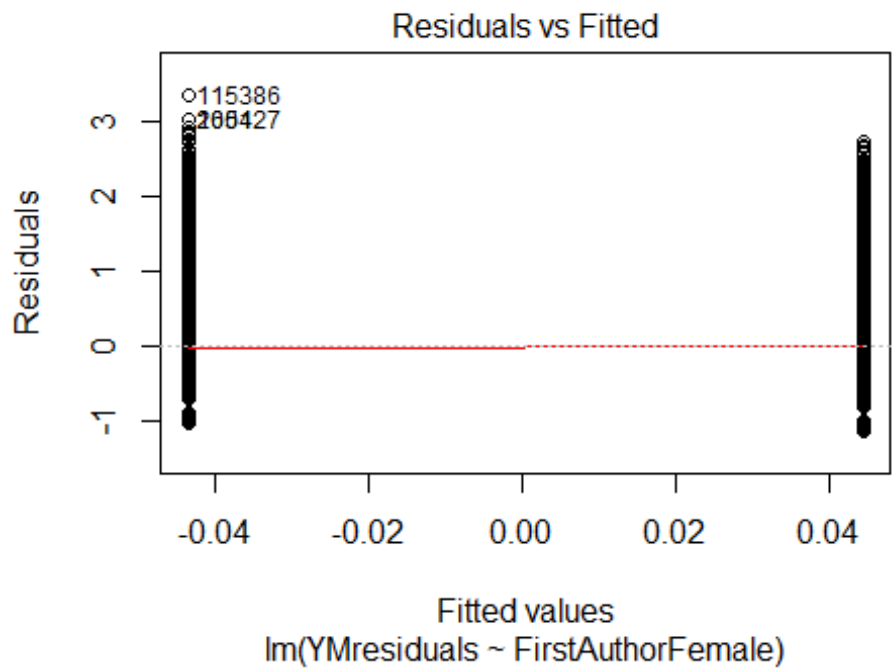

```
## [1] "List of  1 outliers with residuals above  2.5"
##          ScopusId  NLCS Year OneField Fields residuals
## 13361 77954069125 3.652 2010     3303      3     2.541
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.2511 -0.4389  0.0128  0.4396  2.5441
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.171411   0.030650   38.22  < 2e-16 ***
## LastAuthorFemale1   0.040104   0.013968    2.87  0.00410 **
## Year1997           -0.012576   0.043989   -0.29  0.77497
## Year1998           -0.058766   0.043683   -1.35  0.17856
## Year1999           -0.033569   0.043727   -0.77  0.44268
## Year2000           -0.053931   0.042415   -1.27  0.20358
## Year2001           -0.148933   0.043780   -3.40  0.00067 ***
## Year2002           -0.014708   0.039778   -0.37  0.71157
## Year2003            0.019842   0.039020    0.51  0.61111
## Year2004           -0.000573   0.039431   -0.01  0.98841
## Year2005           -0.020995   0.041138   -0.51  0.60981
## Year2006           -0.021320   0.039623   -0.54  0.59053
```

```
## Year2007           0.039624   0.039268    1.01  0.31297
## Year2008          -0.028938   0.040093   -0.72  0.47045
## Year2009          -0.101813   0.040392   -2.52  0.01173 *
## Year2010          -0.063464   0.038443   -1.65  0.09880 .
## Year2011          -0.050503   0.039139   -1.29  0.19695
## Year2012          -0.115655   0.039998   -2.89  0.00384 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.654
## Multiple R-squared:  0.00587,    Adjusted R-squared:  0.00423
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  894 weights are ~= 1. The remaining 9466 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0969  0.8730  0.9490  0.9120  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          9.65e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Sample size for the above analysis:  10360"
## [1] ""
## [1] ""
## [1] "####################################"
## [1] "Analysis of AJSC 3304"
## [1] "####################################"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 4042 4673 4880 4871 5343 5489 5567 4651 4854 5560 6181 6635 7740 8862 9598
## 2011 2012
## 9594 9780
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 3464 3997 4127 4133 4431 4510 4870 4018 4203 4760 5302 5716 6656 7584 8123
## 2011 2012
```

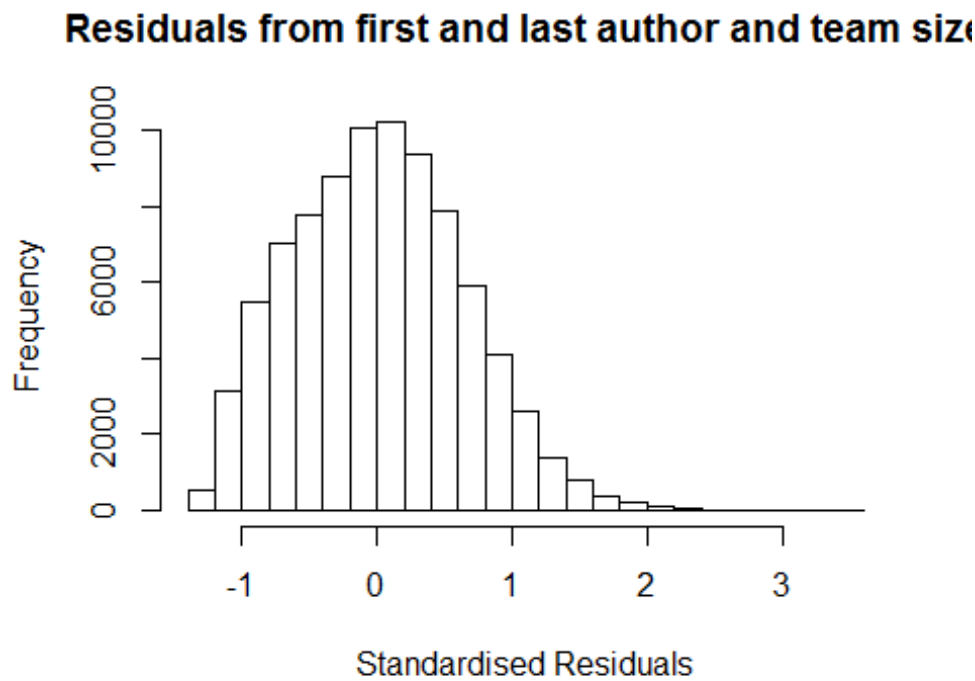
```
## 8178 8336
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 3291 3799 3891 3898 4144 4196 4536 3749 3945 4438 4947 5291 6161 7013 7470
## 2011 2012
## 7556 7607
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
##  Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 360, df = 16, p-value <2e-16
```




```
##
##  Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 28, df = 1, p-value = 1e-07
```



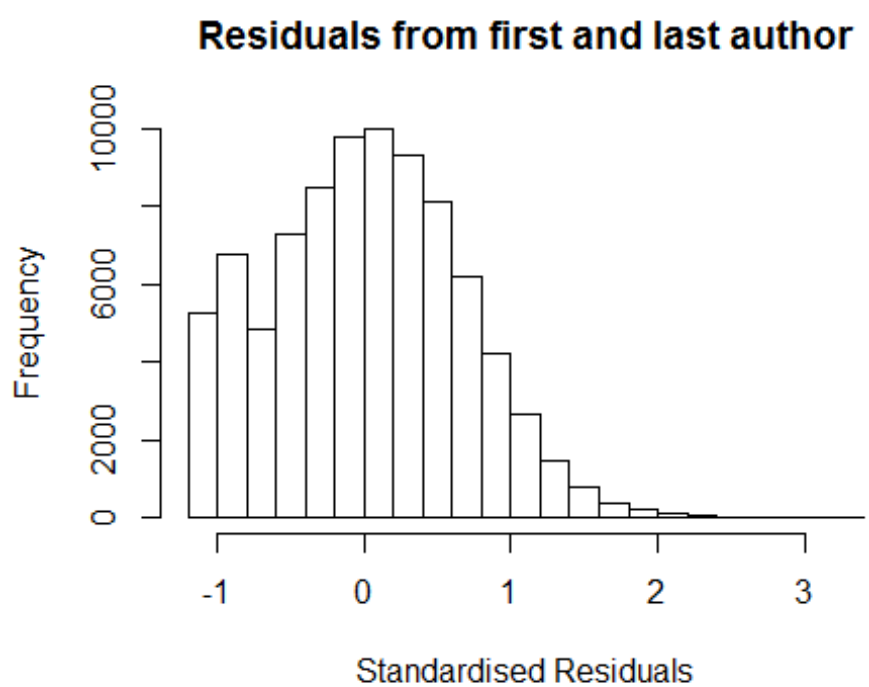

```
## [1] "Female first author team size 2018 geometric mean: 2.28291705523884"
## [1] "Male first author team size 2018 geometric mean: 2.08111614587102"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 8500000, p-value = 4e-11
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 2.23403879023806"
## [1] "Male last author team size 2018 geometric mean: 2.1510716133355"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 8200000, p-value = 0.006
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##                     GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.373  1           1.172
## LastAuthorFemale  1.363  1           1.168
## UniqueAuthors     1.033  4           1.004
## Year              1.031 16           1.001
```




```
## [1] "List of  26 outliers with residuals above  2.5"
##           ScopusId  NLCS Year OneField Fields residuals
## 2601    0030352186 3.933 1996     3304      1     3.171
## 2653    0030532180 3.339 1996     3304      1     2.577
## 3042   85030536309 3.393 1996     3304      1     2.631
## 3922   21344474293 3.593 1996     2601      2     2.535
## 6896    0031514189 3.674 1997     3304      1     2.613
## 13051   0032251946 3.572 1998     3301      2     2.708
## 27335  84993073584 3.324 2001     3304      1     2.529
## 27676  84993110618 3.814 2001     3304      1     3.019
## 48719  84880609979 3.757 2005     3304      1     2.681
## 54453  33845334226 3.311 2006     1202      4     2.532
## 59749  33646724062 3.707 2006     2200      2     2.697
## 62900  38849137287 3.410 2007     2700      2     2.530
## 65860  34147141318 3.409 2007     1208      2     2.620
## 73170  40049092173 3.457 2008     3304      1     2.583
## 73245  43449127813 3.335 2008     3304      1     2.552
## 76310  85014946308 3.363 2008     1203      4     2.580
## 81781  68349103080 3.569 2009     3304      1     2.678
## 84815  70849093912 3.618 2009     3300      2     2.818
## 93783  73949128991 3.447 2010     3304      1     2.662
## 105288 79952183161 3.858 2011     2700      2     2.606
## 105427 79958822104 4.047 2011     2700      2     2.770
## 109525 84857713208 3.859 2012     1700      2     2.771
## 112538 84857205310 3.786 2012     1203      3     2.633
## 112810 84859606223 3.758 2012     3304      1     2.631
## 114239 84856817391 3.466 2012     1213      2     2.634
```

```
## 115386 80052530633 4.385 2012     1700      2     3.553
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
UniqueAuthors +
##     Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##     k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.3603 -0.4552  0.0026  0.4470  3.5533
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         0.76223    0.01249   61.03  < 2e-16 ***
## FirstAuthorFemale1  0.06550    0.00528   12.40  < 2e-16 ***
## LastAuthorFemale1   0.02543    0.00526    4.83  1.4e-06 ***
## UniqueAuthors2      0.23057    0.00559   41.24  < 2e-16 ***
## UniqueAuthors3      0.29529    0.00671   43.99  < 2e-16 ***
## UniqueAuthors4      0.37261    0.00883   42.21  < 2e-16 ***
## UniqueAuthors5      0.43766    0.00889   49.23  < 2e-16 ***
## Year1997            0.00310    0.01650    0.19  0.85105
## Year1998            0.01088    0.01612    0.67  0.49981
## Year1999            0.00388    0.01591    0.24  0.80737
## Year2000            0.01942    0.01555    1.25  0.21170
## Year2001            0.03297    0.01569    2.10  0.03561 *
## Year2002            0.04467    0.01541    2.90  0.00375 **
## Year2003            0.02106    0.01608    1.31  0.19045
## Year2004            0.01399    0.01577    0.89  0.37500
## Year2005            0.01721    0.01538    1.12  0.26336
## Year2006            0.01698    0.01489    1.14  0.25406
## Year2007            0.02656    0.01481    1.79  0.07297 .
## Year2008            0.02117    0.01459    1.45  0.14670
## Year2009            0.03747    0.01457    2.57  0.01011 *
## Year2010            0.02302    0.01436    1.60  0.10904
## Year2011            0.05191    0.01440    3.61  0.00031 ***
## Year2012            0.06944    0.01478    4.70  2.6e-06 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.665
## Multiple R-squared:  0.0619, Adjusted R-squared:  0.0617
## Convergence in 10 IRWLS iterations
##
## Robustness weights:
##  2 observations c(1937,84778) are outliers with |weight| = 0 ( < 1.2e-06);
##  7168 weights are ~= 1. The remaining 78762 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0038  0.8750  0.9510  0.9150  0.9850  0.9990
```

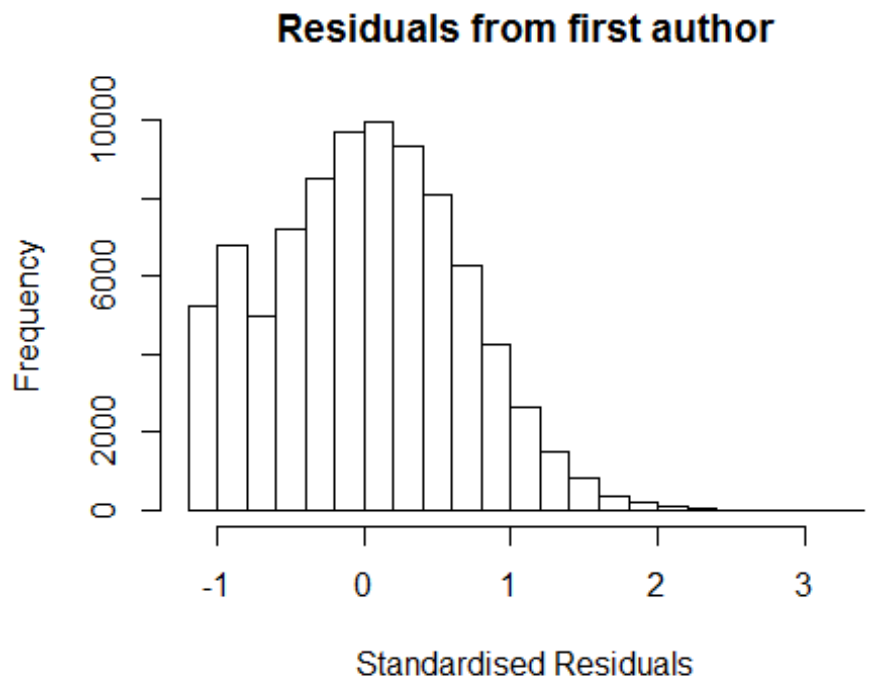
```
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.16e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.377  1           1.174
## LastAuthorFemale  1.373  1           1.172
## Year              1.010 16           1.000
```




```
## [1] "List of  18 outliers with residuals above  2.5"
##           ScopusId  NLCS Year OneField Fields residuals
## 35      0002004285 3.441 1996     3304      1     2.557
## 2601    0030352186 3.933 1996     3304      1     3.049
```

```
## 3042   85030536309 3.393 1996     3304      1     2.509
## 3922   21344474293 3.593 1996     2601      2     2.623
## 6896    0031514189 3.674 1997     3304      1     2.707
## 13051   0032251946 3.572 1998     3301      2     2.563
## 27676  84993110618 3.814 2001     3304      1     2.890
## 48719  84880609979 3.757 2005     3304      1     2.750
## 54825  33845474311 3.471 2006     3301      2     2.536
## 59749  33646724062 3.707 2006     2200      2     2.772
## 84815  70849093912 3.618 2009     3300      2     2.658
## 93783  73949128991 3.447 2010     3304      1     2.501
## 105288 79952183161 3.858 2011     2700      2     2.873
## 105427 79958822104 4.047 2011     2700      2     3.033
## 109525 84857713208 3.859 2012     1700      2     2.826
## 112538 84857205310 3.786 2012     1203      3     2.667
## 112810 84859606223 3.758 2012     3304      1     2.754
## 115386 80052530633 4.385 2012     1700      2     3.381
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.1189 -0.4671  0.0112  0.4624  3.3808
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         0.88440    0.01259   70.24  < 2e-16 ***
## FirstAuthorFemale1  0.08589    0.00546   15.73  < 2e-16 ***
## LastAuthorFemale1   0.02877    0.00545    5.27  1.3e-07 ***
## Year1997           -0.00331    0.01685   -0.20  0.84406
## Year1998            0.01036    0.01640    0.63  0.52761
## Year1999           -0.00277    0.01622   -0.17  0.86440
## Year2000            0.02049    0.01595    1.29  0.19878
## Year2001            0.03984    0.01602    2.49  0.01290 *
## Year2002            0.06240    0.01586    3.94  8.3e-05 ***
## Year2003            0.03490    0.01649    2.12  0.03431 *
## Year2004            0.03755    0.01627    2.31  0.02100 *
## Year2005            0.03645    0.01582    2.30  0.02124 *
## Year2006            0.05071    0.01531    3.31  0.00093 ***
## Year2007            0.06338    0.01522    4.16  3.1e-05 ***
## Year2008            0.05563    0.01502    3.70  0.00021 ***
## Year2009            0.07519    0.01498    5.02  5.2e-07 ***
## Year2010            0.06210    0.01479    4.20  2.7e-05 ***
## Year2011            0.10056    0.01476    6.81  9.6e-12 ***
## Year2012            0.11984    0.01517    7.90  2.8e-15 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
```

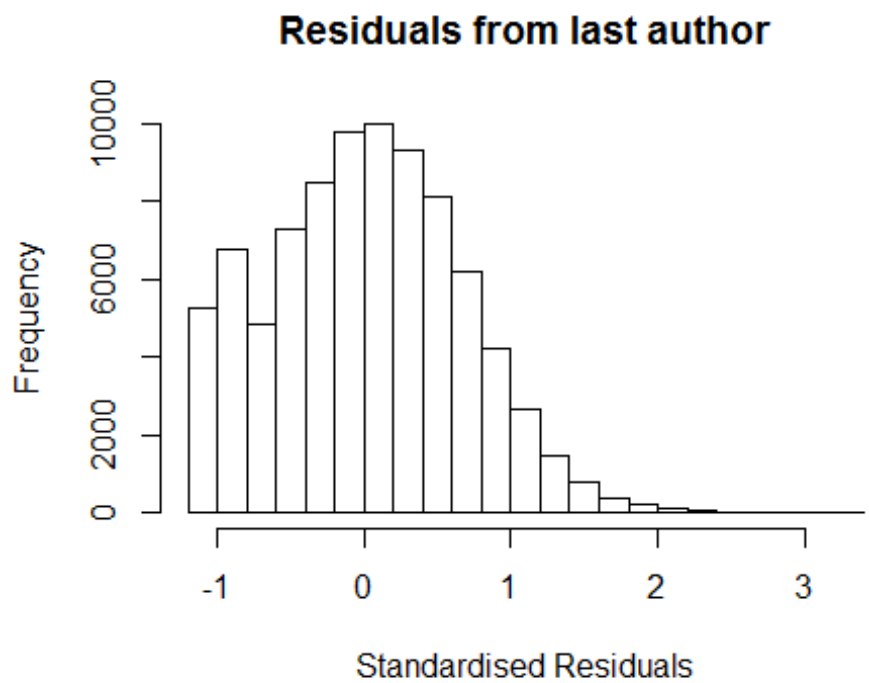
```
## Robust residual standard error: 0.688
## Multiple R-squared:  0.0094, Adjusted R-squared:  0.00919
## Convergence in 10 IRWLS iterations
##
## Robustness weights:
##  observation 84778 is an outlier with |weight| = 0 ( < 1.2e-06);
##  7330 weights are ~= 1. The remaining 78601 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0111  0.8560  0.9500  0.9160  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.16e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.009  1           1.004
## Year              1.009 16           1.000
```




```
## [1] "List of  18 outliers with residuals above  2.5"
##           ScopusId  NLCS Year OneField Fields residuals
## 35      0002004285 3.441 1996     3304      1     2.557
## 2601    0030352186 3.933 1996     3304      1     3.049
## 3042   85030536309 3.393 1996     3304      1     2.509
## 3922   21344474293 3.593 1996     2601      2     2.623
## 6896    0031514189 3.674 1997     3304      1     2.707
## 13051   0032251946 3.572 1998     3301      2     2.563
## 27676  84993110618 3.814 2001     3304      1     2.890
## 48719  84880609979 3.757 2005     3304      1     2.750
## 54825  33845474311 3.471 2006     3301      2     2.536
## 59749  33646724062 3.707 2006     2200      2     2.772
## 84815  70849093912 3.618 2009     3300      2     2.658
## 93783  73949128991 3.447 2010     3304      1     2.501
## 105288 79952183161 3.858 2011     2700      2     2.873
## 105427 79958822104 4.047 2011     2700      2     3.033
## 109525 84857713208 3.859 2012     1700      2     2.826
## 112538 84857205310 3.786 2012     1203      3     2.667
## 112810 84859606223 3.758 2012     3304      1     2.754
## 115386 80052530633 4.385 2012     1700      2     3.381
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
```

```
##     Min      1Q  Median      3Q     Max
## -1.1125 -0.4648  0.0128  0.4637  3.3742
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         0.88988    0.01254   70.98  < 2e-16 ***
## FirstAuthorFemale1  0.10170    0.00468   21.74  < 2e-16 ***
## Year1997           -0.00353    0.01685   -0.21  0.83398
## Year1998            0.01058    0.01641    0.64  0.51909
## Year1999           -0.00252    0.01622   -0.16  0.87637
## Year2000            0.02050    0.01594    1.29  0.19867
## Year2001            0.04034    0.01602    2.52  0.01181 *
## Year2002            0.06257    0.01585    3.95  7.9e-05 ***
## Year2003            0.03537    0.01649    2.15  0.03195 *
## Year2004            0.03803    0.01627    2.34  0.01943 *
## Year2005            0.03696    0.01582    2.34  0.01945 *
## Year2006            0.05163    0.01531    3.37  0.00075 ***
## Year2007            0.06434    0.01522    4.23  2.4e-05 ***
## Year2008            0.05633    0.01501    3.75  0.00018 ***
## Year2009            0.07616    0.01498    5.08  3.7e-07 ***
## Year2010            0.06286    0.01479    4.25  2.1e-05 ***
## Year2011            0.10124    0.01476    6.86  6.9e-12 ***
## Year2012            0.12094    0.01517    7.97  1.6e-15 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.688
## Multiple R-squared:  0.00906,    Adjusted R-squared:  0.00887
## Convergence in 10 IRWLS iterations
##
## Robustness weights:
##  observation 84778 is an outlier with |weight| = 0 ( < 1.2e-06);
##  7255 weights are ~= 1. The remaining 78676 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0104  0.8550  0.9490  0.9160  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.16e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
```

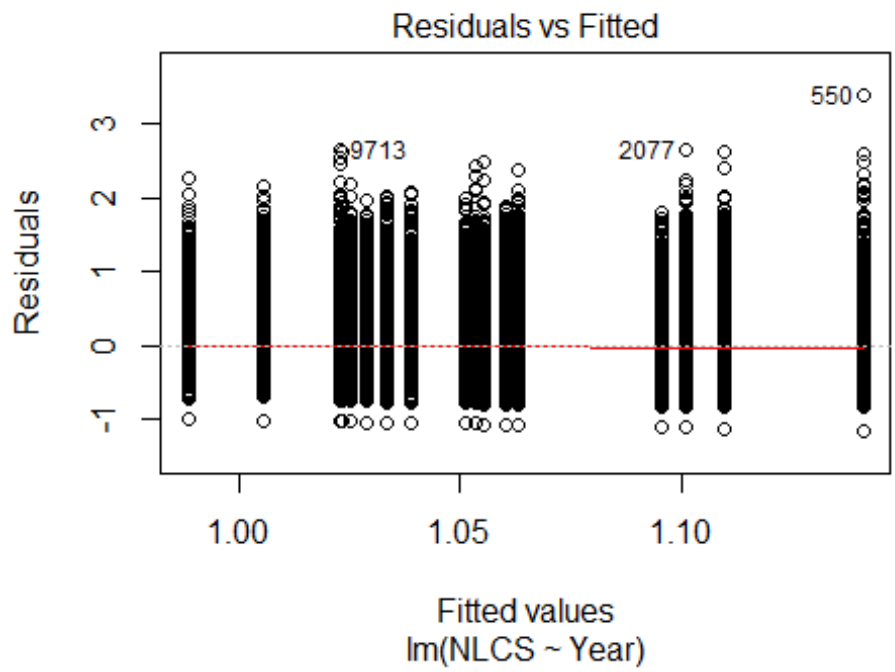
```
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##                   GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.006  1           1.003
## Year             1.006 16           1.000
```



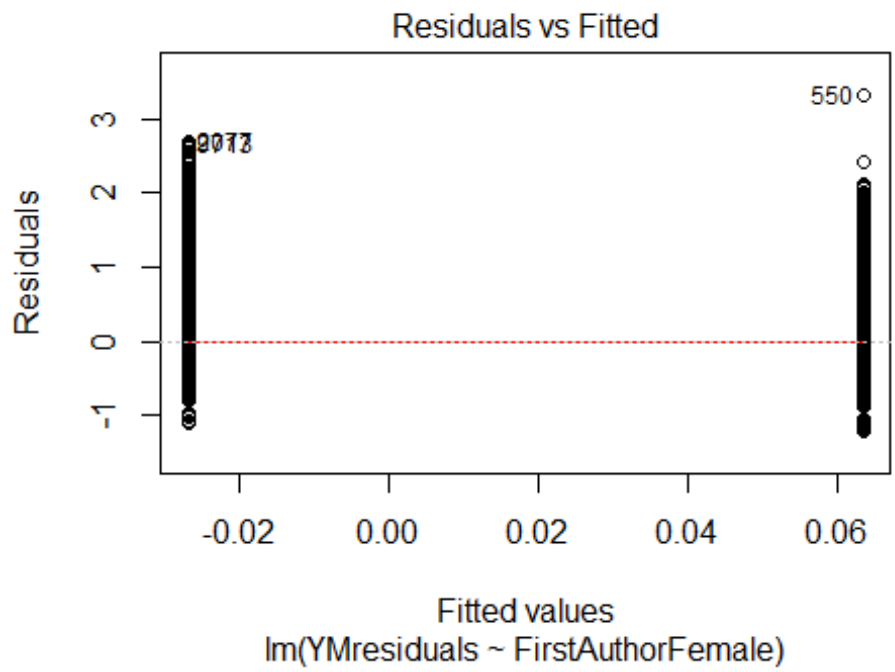

```
## [1] "List of  18 outliers with residuals above  2.5"
##           ScopusId  NLCS Year OneField Fields residuals
## 35      0002004285 3.441 1996     3304      1     2.557
## 2601    0030352186 3.933 1996     3304      1     3.049
## 3042   85030536309 3.393 1996     3304      1     2.509
## 3922   21344474293 3.593 1996     2601      2     2.623
## 6896    0031514189 3.674 1997     3304      1     2.707
## 13051   0032251946 3.572 1998     3301      2     2.563
## 27676  84993110618 3.814 2001     3304      1     2.890
## 48719  84880609979 3.757 2005     3304      1     2.750
## 54825  33845474311 3.471 2006     3301      2     2.536
## 59749  33646724062 3.707 2006     2200      2     2.772
## 84815  70849093912 3.618 2009     3300      2     2.658
## 93783  73949128991 3.447 2010     3304      1     2.501
## 105288 79952183161 3.858 2011     2700      2     2.873
## 105427 79958822104 4.047 2011     2700      2     3.033
## 109525 84857713208 3.859 2012     1700      2     2.826
## 112538 84857205310 3.786 2012     1203      3     2.667
## 112810 84859606223 3.758 2012     3304      1     2.754
## 115386 80052530633 4.385 2012     1700      2     3.381
##
```

```
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.1017 -0.4681  0.0118  0.4637  3.3593
##
## Coefficients:
##                   Estimate Std. Error t value Pr(>|t|)
## (Intercept)        0.90116    0.01256   71.77  < 2e-16 ***
## LastAuthorFemale1  0.07593    0.00468   16.24  < 2e-16 ***
## Year1997          -0.00309    0.01687   -0.18  0.85481
## Year1998           0.01009    0.01642    0.61  0.53888
## Year1999          -0.00331    0.01624   -0.20  0.83830
## Year2000           0.02243    0.01597    1.40  0.16028
## Year2001           0.04165    0.01603    2.60  0.00938 **
## Year2002           0.06513    0.01588    4.10  4.1e-05 ***
## Year2003           0.03698    0.01651    2.24  0.02506 *
## Year2004           0.04111    0.01629    2.52  0.01160 *
## Year2005           0.03974    0.01584    2.51  0.01213 *
## Year2006           0.05447    0.01533    3.55  0.00038 ***
## Year2007           0.06695    0.01524    4.39  1.1e-05 ***
## Year2008           0.05971    0.01505    3.97  7.2e-05 ***
## Year2009           0.07899    0.01500    5.26  1.4e-07 ***
## Year2010           0.06646    0.01481    4.49  7.2e-06 ***
## Year2011           0.10602    0.01477    7.18  7.2e-13 ***
## Year2012           0.12457    0.01519    8.20  2.4e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.69
## Multiple R-squared:  0.00653,    Adjusted R-squared:  0.00633
## Convergence in 10 IRWLS iterations
##
## Robustness weights:
##  observation 84778 is an outlier with |weight| = 0 ( < 1.2e-06);
##  7306 weights are ~= 1. The remaining 78625 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0143  0.8520  0.9500  0.9170  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.16e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
```

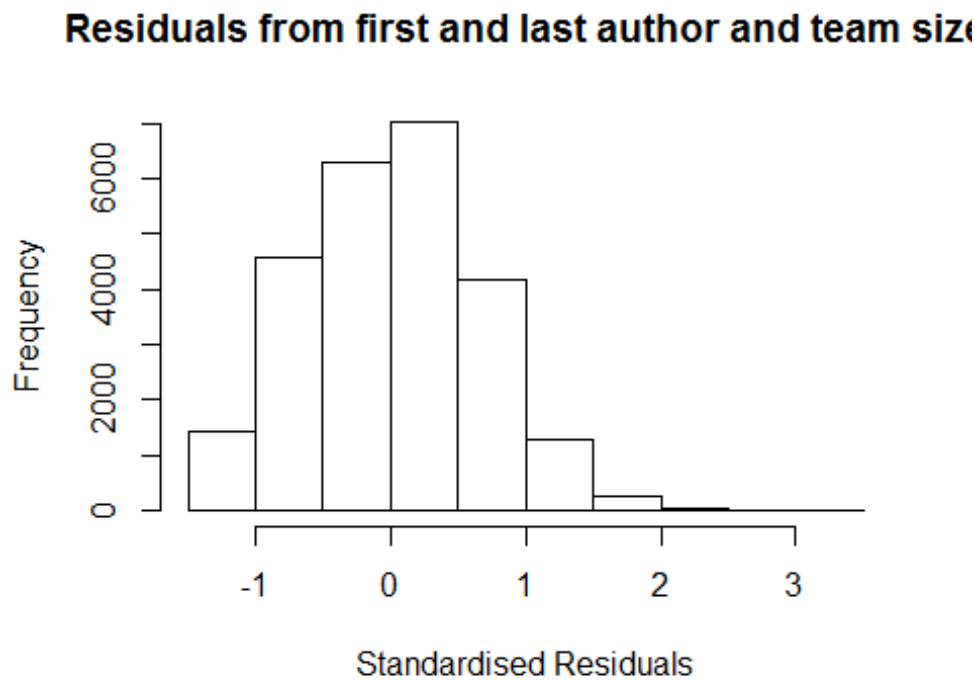
```
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Sample size for the above analysis:  85932"
## [1] ""
## [1] ""
## [1] "####################################"
## [1] "Analysis of AJSC 3305"
## [1] "####################################"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 1353 1406 1447 1648 1810 1921 1999 2103 1959 1863 2023 2289 2235 2356 2660
## 2011 2012
## 2532 2665
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 1032 1062 1045 1080 1068 1071 1642 1723 1635 1533 1691 1875 1798 1903 2130
## 2011 2012
## 2049 2164
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  984 1004  992 1027 1001 1013 1574 1648 1560 1458 1596 1796 1710 1786 2006
## 2011 2012
## 1896 1998
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
##  Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 88, df = 16, p-value = 6e-12
```




```
##
##  Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 7.9, df = 1, p-value = 0.005
```



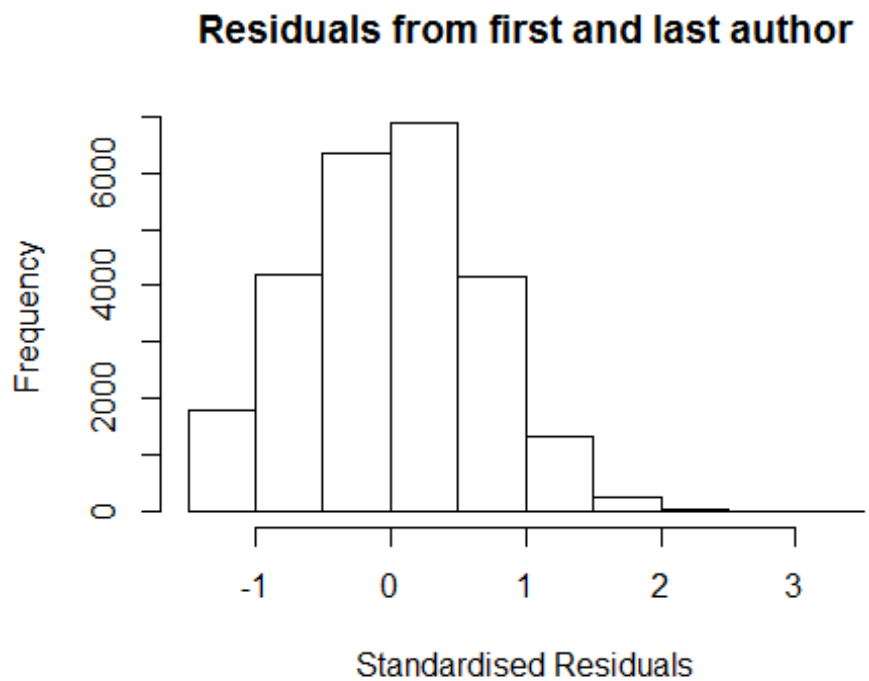

```
## [1] "Female first author team size 2018 geometric mean: 1.89287455945294"
## [1] "Male first author team size 2018 geometric mean: 1.7713540863654"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 5e+05, p-value = 0.03
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 1.82352930368327"
## [1] "Male last author team size 2018 geometric mean: 1.81439407623594"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 460000, p-value = 0.9
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.937  1           1.392
## LastAuthorFemale  1.934  1           1.391
## UniqueAuthors     1.034  4           1.004
## Year              1.041 16           1.001
```




```
## [1] "List of  7 outliers with residuals above  2.5"
##          ScopusId  NLCS Year OneField Fields residuals
## 550    0001019986 4.524 1996     3305      1     3.369
## 2077   0030813443 3.760 1997     3305      1     2.753
## 4385   0032447095 3.733 1998     1710      3     2.578
## 8378   0035556741 3.567 2001     3305      1     2.654
## 9222   0035101649 3.482 2001     1202      3     2.569
## 9713   0035560936 3.672 2001     3305      1     2.549
## 30478 77954069125 3.652 2010     3303      3     2.608
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
UniqueAuthors +
##     Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##     k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.4071 -0.4725  0.0158  0.4573  3.3689
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)           1.0485     0.0243   43.07  < 2e-16 ***
## FirstAuthorFemale1    0.0692     0.0128    5.41  6.5e-08 ***
## LastAuthorFemale1     0.0374     0.0130    2.89   0.0039 **
## UniqueAuthors2        0.1410     0.0110   12.87  < 2e-16 ***
```

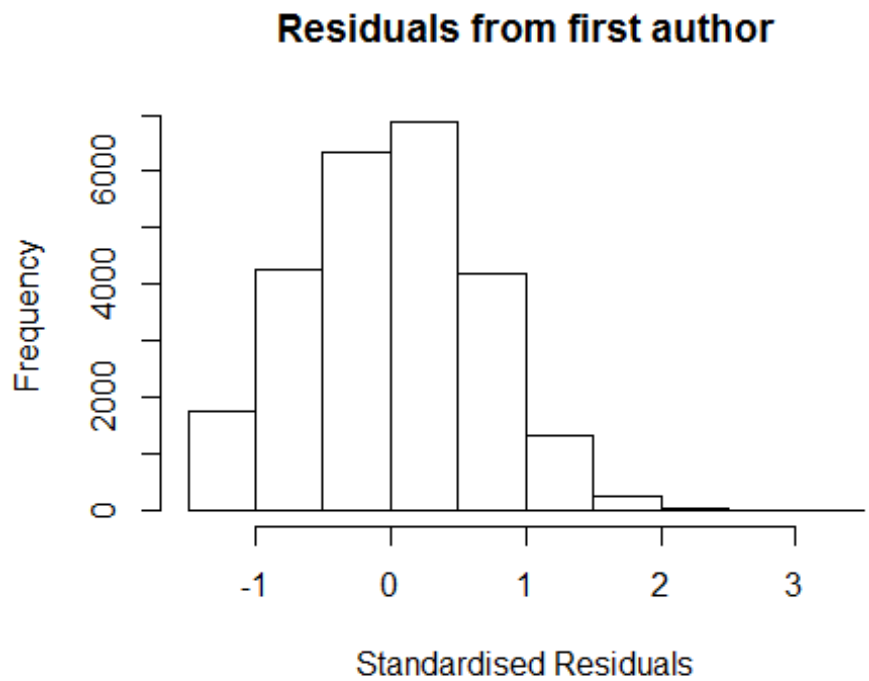
```
## UniqueAuthors3       0.2092     0.0153   13.67  < 2e-16 ***
## UniqueAuthors4       0.2120     0.0217    9.76  < 2e-16 ***
## UniqueAuthors5       0.2893     0.0247   11.71  < 2e-16 ***
## Year1997            -0.0417     0.0335   -1.25   0.2122
## Year1998            -0.0344     0.0330   -1.04   0.2966
## Year1999            -0.0988     0.0322   -3.07   0.0021 **
## Year2000            -0.0454     0.0323   -1.40   0.1600
## Year2001            -0.1351     0.0331   -4.08  4.5e-05 ***
## Year2002            -0.0794     0.0300   -2.65   0.0081 **
## Year2003            -0.0863     0.0294   -2.93   0.0033 **
## Year2004            -0.0790     0.0297   -2.66   0.0078 **
## Year2005            -0.0935     0.0296   -3.16   0.0016 **
## Year2006            -0.1143     0.0291   -3.92  8.7e-05 ***
## Year2007            -0.1584     0.0291   -5.45  5.2e-08 ***
## Year2008            -0.1288     0.0291   -4.43  9.5e-06 ***
## Year2009            -0.1359     0.0285   -4.76  2.0e-06 ***
## Year2010            -0.1459     0.0279   -5.22  1.8e-07 ***
## Year2011            -0.1548     0.0286   -5.42  6.1e-08 ***
## Year2012            -0.1727     0.0282   -6.13  8.8e-10 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.682
## Multiple R-squared:  0.0245, Adjusted R-squared:  0.0236
## Convergence in 10 IRWLS iterations
##
## Robustness weights:
##  observation 378 is an outlier with |weight| = 0 ( < 4e-06);
##  2085 weights are ~= 1. The remaining 22963 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0663  0.8660  0.9500  0.9170  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          3.99e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                    psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##                     GVIF Df GVIF^(1/(2*Df))
```

```
## FirstAuthorFemale 1.961  1           1.400
## LastAuthorFemale  1.958  1           1.399
## Year              1.010 16           1.000
```



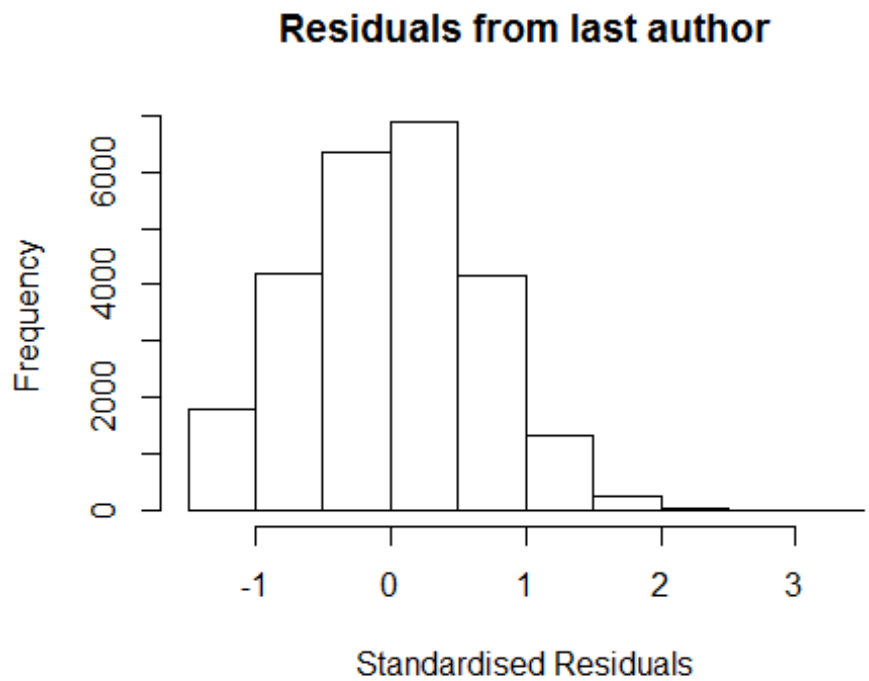

```
## [1] "List of  7 outliers with residuals above  2.5"
##          ScopusId  NLCS Year OneField Fields residuals
## 550    0001019986 4.524 1996     3305      1     3.315
## 2077   0030813443 3.760 1997     3305      1     2.705
## 4385   0032447095 3.733 1998     1710      3     2.668
## 8378   0035556741 3.567 2001     3305      1     2.595
## 9222   0035101649 3.482 2001     1202      3     2.510
## 9713   0035560936 3.672 2001     3305      1     2.700
## 30478 77954069125 3.652 2010     3303      3     2.680
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.2087 -0.4751  0.0114  0.4634  3.3153
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)          1.1003     0.0240   45.79  < 2e-16 ***
## FirstAuthorFemale1   0.0773     0.0130    5.94  2.9e-09 ***
```

```
## LastAuthorFemale1    0.0311     0.0131    2.37  0.01786 *
## Year1997            -0.0457     0.0334   -1.37  0.17122
## Year1998            -0.0348     0.0329   -1.06  0.28938
## Year1999            -0.1000     0.0321   -3.11  0.00185 **
## Year2000            -0.0369     0.0323   -1.14  0.25342
## Year2001            -0.1287     0.0331   -3.89  0.00010 ***
## Year2002            -0.0816     0.0299   -2.73  0.00628 **
## Year2003            -0.0902     0.0293   -3.08  0.00210 **
## Year2004            -0.0807     0.0296   -2.73  0.00634 **
## Year2005            -0.0894     0.0296   -3.02  0.00249 **
## Year2006            -0.1082     0.0290   -3.73  0.00019 ***
## Year2007            -0.1538     0.0292   -5.28  1.3e-07 ***
## Year2008            -0.1164     0.0290   -4.01  6.1e-05 ***
## Year2009            -0.1221     0.0286   -4.27  1.9e-05 ***
## Year2010            -0.1283     0.0280   -4.59  4.5e-06 ***
## Year2011            -0.1306     0.0285   -4.58  4.7e-06 ***
## Year2012            -0.1457     0.0282   -5.17  2.3e-07 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.689
## Multiple R-squared:  0.0076, Adjusted R-squared:  0.00689
## Convergence in 10 IRWLS iterations
##
## Robustness weights:
##  observation 378 is an outlier with |weight| = 0 ( < 4e-06);
##  2090 weights are ~= 1. The remaining 22958 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.088   0.873   0.949   0.917   0.985   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          3.99e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.008  1           1.004
## Year              1.008 16           1.000
```




```
## [1] "List of  7 outliers with residuals above  2.5"
##          ScopusId  NLCS Year OneField Fields residuals
## 550    0001019986 4.524 1996     3305      1     3.315
## 2077   0030813443 3.760 1997     3305      1     2.705
## 4385   0032447095 3.733 1998     1710      3     2.668
## 8378   0035556741 3.567 2001     3305      1     2.595
## 9222   0035101649 3.482 2001     1202      3     2.510
## 9713   0035560936 3.672 2001     3305      1     2.700
## 30478 77954069125 3.652 2010     3303      3     2.680
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.2014 -0.4743  0.0118  0.4632  3.3226
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.10229    0.02400   45.93  < 2e-16 ***
## FirstAuthorFemale1  0.09908    0.00932   10.63  < 2e-16 ***
## Year1997           -0.04571    0.03341   -1.37  0.17125
## Year1998           -0.03486    0.03286   -1.06  0.28871
## Year1999           -0.09994    0.03209   -3.11  0.00184 **
## Year2000           -0.03638    0.03232   -1.13  0.26030
```

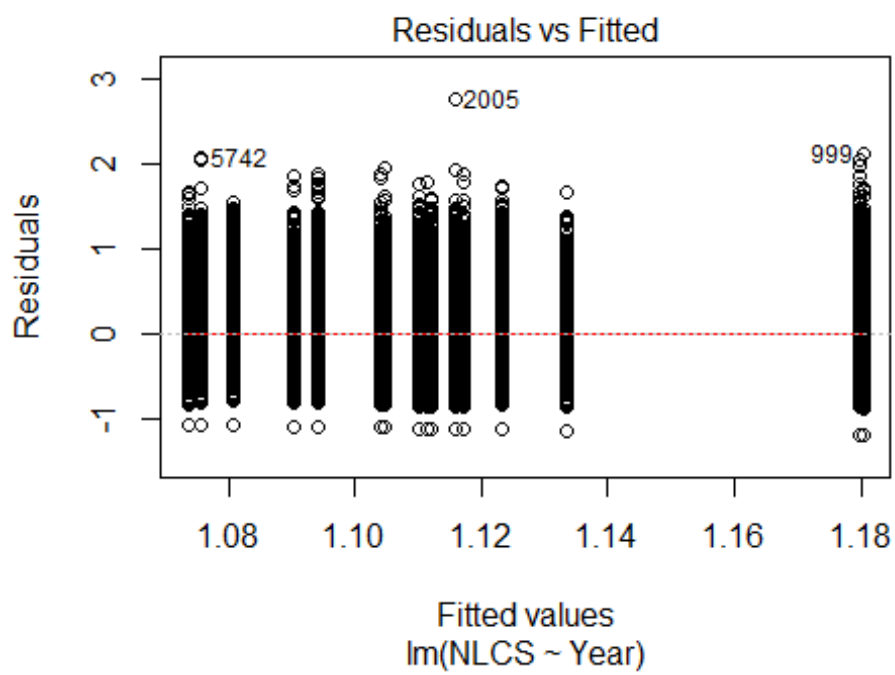
```
## Year2001           -0.12808    0.03306   -3.87  0.00011 ***
## Year2002           -0.08106    0.02985   -2.72  0.00662 **
## Year2003           -0.08959    0.02931   -3.06  0.00224 **
## Year2004           -0.07990    0.02954   -2.70  0.00684 **
## Year2005           -0.08923    0.02955   -3.02  0.00253 **
## Year2006           -0.10739    0.02902   -3.70  0.00022 ***
## Year2007           -0.15307    0.02915   -5.25  1.5e-07 ***
## Year2008           -0.11557    0.02902   -3.98  6.8e-05 ***
## Year2009           -0.12126    0.02855   -4.25  2.2e-05 ***
## Year2010           -0.12819    0.02795   -4.59  4.5e-06 ***
## Year2011           -0.13005    0.02849   -4.57  5.0e-06 ***
## Year2012           -0.14465    0.02816   -5.14  2.8e-07 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.689
## Multiple R-squared:  0.00738,    Adjusted R-squared:  0.0067
## Convergence in 10 IRWLS iterations
##
## Robustness weights:
##  observation 378 is an outlier with |weight| = 0 ( < 4e-06);
##  2100 weights are ~= 1. The remaining 22948 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0886  0.8720  0.9490  0.9170  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          3.99e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##                   GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.007  1           1.003
## Year             1.007 16           1.000
```



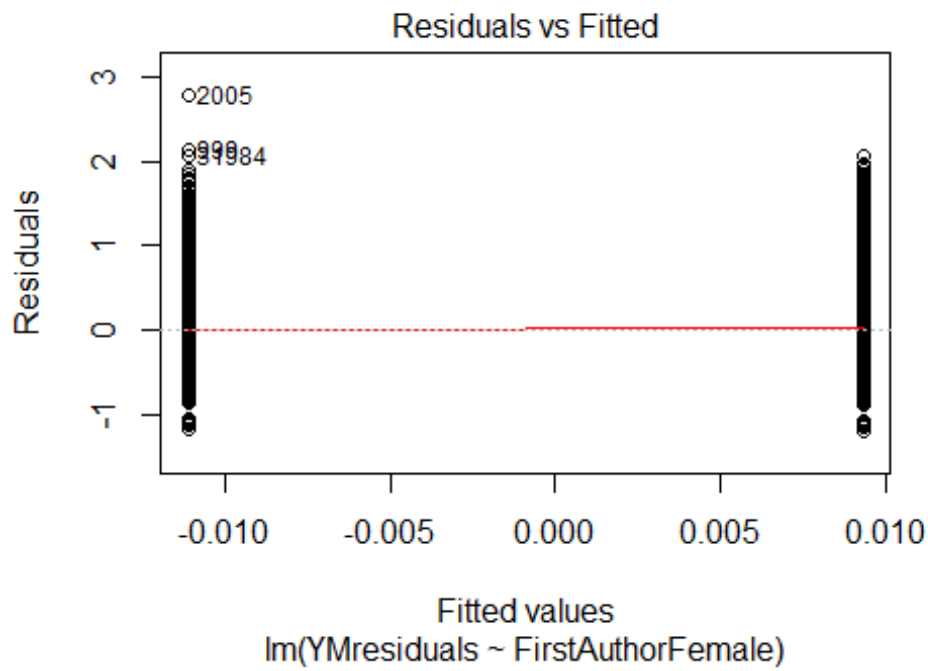

```
## [1] "List of  7 outliers with residuals above  2.5"
##          ScopusId  NLCS Year OneField Fields residuals
## 550    0001019986 4.524 1996     3305      1     3.315
## 2077   0030813443 3.760 1997     3305      1     2.705
## 4385   0032447095 3.733 1998     1710      3     2.668
## 8378   0035556741 3.567 2001     3305      1     2.595
## 9222   0035101649 3.482 2001     1202      3     2.510
## 9713   0035560936 3.672 2001     3305      1     2.700
## 30478 77954069125 3.652 2010     3303      3     2.680
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.1908 -0.4762  0.0134  0.4659  3.3332
##
## Coefficients:
##                   Estimate Std. Error t value Pr(>|t|)
## (Intercept)        1.10490    0.02401   46.02  < 2e-16 ***
## LastAuthorFemale1  0.08591    0.00941    9.13  < 2e-16 ***
## Year1997          -0.04401    0.03345   -1.32  0.18821
## Year1998          -0.03452    0.03289   -1.05  0.29388
## Year1999          -0.09786    0.03213   -3.05  0.00232 **
## Year2000          -0.03504    0.03234   -1.08  0.27868
```

```
## Year2001          -0.12692    0.03310   -3.83  0.00013 ***
## Year2002          -0.07984    0.02986   -2.67  0.00750 **
## Year2003          -0.08781    0.02932   -2.99  0.00275 **
## Year2004          -0.07931    0.02956   -2.68  0.00730 **
## Year2005          -0.08664    0.02957   -2.93  0.00339 **
## Year2006          -0.10577    0.02901   -3.65  0.00027 ***
## Year2007          -0.15147    0.02917   -5.19  2.1e-07 ***
## Year2008          -0.11408    0.02903   -3.93  8.5e-05 ***
## Year2009          -0.11966    0.02856   -4.19  2.8e-05 ***
## Year2010          -0.12325    0.02796   -4.41  1.1e-05 ***
## Year2011          -0.12595    0.02849   -4.42  9.9e-06 ***
## Year2012          -0.14235    0.02815   -5.06  4.3e-07 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.689
## Multiple R-squared:  0.0062, Adjusted R-squared:  0.00553
## Convergence in 10 IRWLS iterations
##
## Robustness weights:
##  observation 378 is an outlier with |weight| = 0 ( < 4e-06);
##  2108 weights are ~= 1. The remaining 22940 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0905  0.8720  0.9490  0.9170  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          3.99e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Sample size for the above analysis:  25049"
## [1] ""
## [1] ""
## [1] "####################################"
## [1] "Analysis of AJSC 3306"
## [1] "####################################"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
```

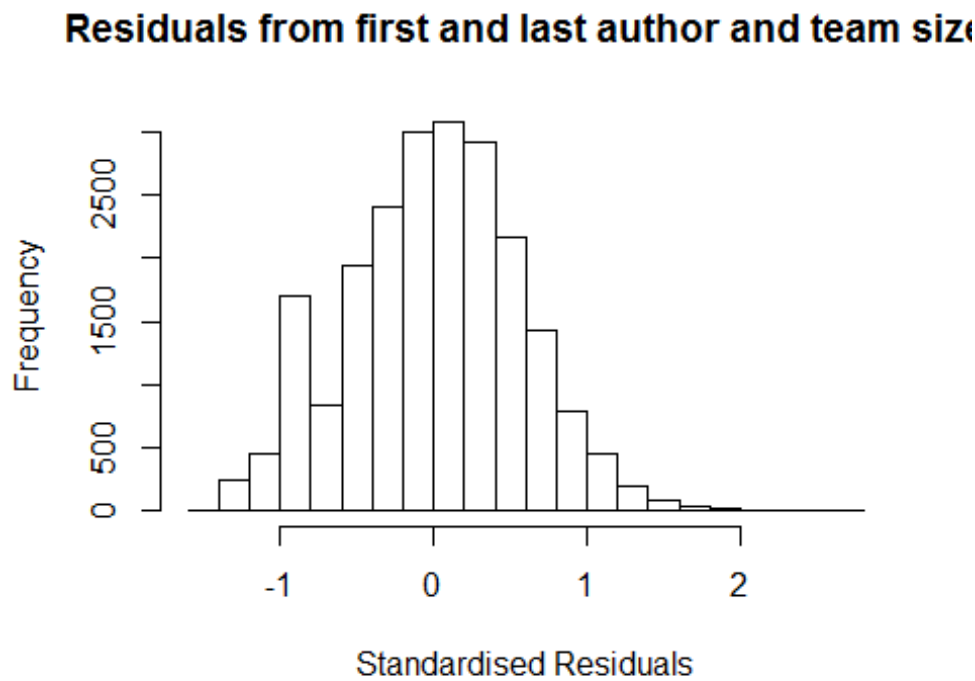
```
## 1176 1224 1183 1162 1354 1502 1510 1367 1522 1584 1775 1976 2154 2336 2530
## 2011 2012
## 2621 2721
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  856  889  803  867  859  877 1244 1188 1287 1334 1484 1651 1809 1991 2135
## 2011 2012
## 2211 2312
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  798  824  741  799  794  818 1156 1100 1164 1197 1330 1497 1653 1813 1966
## 2011 2012
## 1997 2073
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
##  Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 240, df = 16, p-value <2e-16
```




```
##
##  Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 0.033, df = 1, p-value = 0.9
```

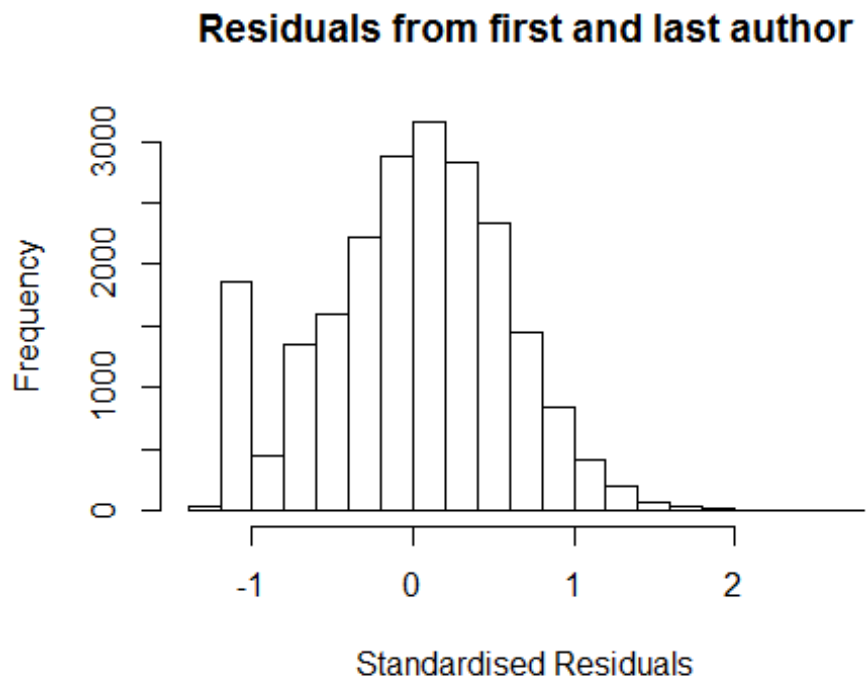



```
## [1] "Female first author team size 2018 geometric mean: 3.22963403487477"
## [1] "Male first author team size 2018 geometric mean: 2.82607175306154"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 880000, p-value = 1e-05
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 3.06029923159863"
## [1] "Male last author team size 2018 geometric mean: 3.09967228083774"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 840000, p-value = 0.6
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##                     GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.116  1           1.057
## LastAuthorFemale  1.109  1           1.053
## UniqueAuthors     1.058  4           1.007
## Year              1.059 16           1.002
```



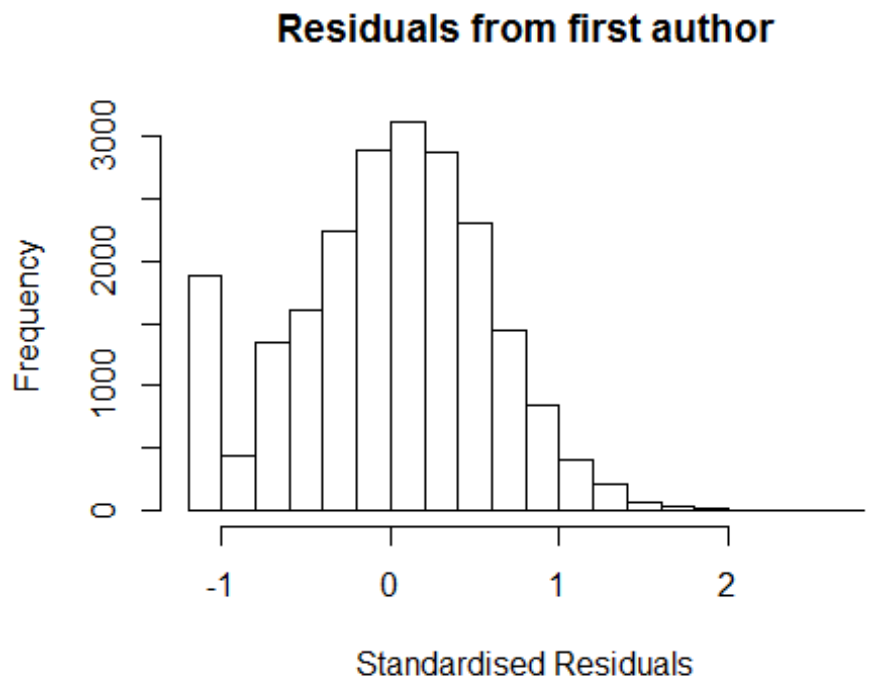

```
## [1] "List of  1 outliers with residuals above  2.5"
##        ScopusId  NLCS Year OneField Fields residuals
## 2005 0030881741 3.884 1997     2739      2     2.705
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
UniqueAuthors +
##     Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##     k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.4396 -0.3769  0.0159  0.3770  2.7048
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         0.98240    0.02777   35.38  < 2e-16 ***
## FirstAuthorFemale1  0.01232    0.00833    1.48  0.13891
## LastAuthorFemale1  -0.02676    0.00827   -3.24  0.00121 **
## UniqueAuthors2      0.25062    0.01193   21.02  < 2e-16 ***
## UniqueAuthors3      0.34186    0.01212   28.20  < 2e-16 ***
## UniqueAuthors4      0.37303    0.01321   28.24  < 2e-16 ***
## UniqueAuthors5      0.44485    0.01136   39.16  < 2e-16 ***
## Year1997           -0.05381    0.03544   -1.52  0.12891
## Year1998           -0.08461    0.03530   -2.40  0.01654 *
## Year1999           -0.07698    0.03418   -2.25  0.02434 *
```

```
## Year2000           -0.11997    0.03408   -3.52  0.00043 ***
## Year2001           -0.05731    0.03444   -1.66  0.09614 .
## Year2002           -0.09806    0.03164   -3.10  0.00194 **
## Year2003           -0.13016    0.03225   -4.04  5.5e-05 ***
## Year2004           -0.09904    0.03250   -3.05  0.00231 **
## Year2005           -0.10583    0.03165   -3.34  0.00083 ***
## Year2006           -0.11846    0.03024   -3.92  9.0e-05 ***
## Year2007           -0.09613    0.03028   -3.18  0.00150 **
## Year2008           -0.12494    0.02953   -4.23  2.3e-05 ***
## Year2009           -0.15273    0.02944   -5.19  2.2e-07 ***
## Year2010           -0.13451    0.02949   -4.56  5.1e-06 ***
## Year2011           -0.11333    0.02948   -3.84  0.00012 ***
## Year2012           -0.06809    0.03006   -2.27  0.02351 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.557
## Multiple R-squared:  0.088,  Adjusted R-squared:  0.0871
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
##  observation 1214 is an outlier with |weight| = 0 ( < 4.6e-06);
##  1844 weights are ~= 1. The remaining 19875 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0601  0.8620  0.9490  0.9070  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          4.60e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.100  1           1.049
## LastAuthorFemale  1.093  1           1.046
## Year              1.014 16           1.000
```



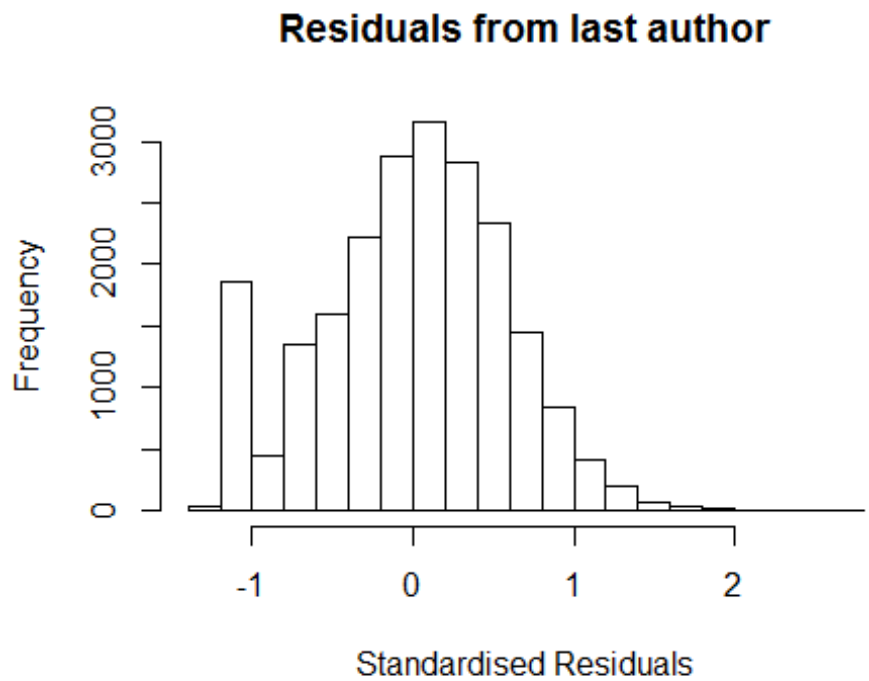

```
## [1] "List of  1 outliers with residuals above  2.5"
##        ScopusId  NLCS Year OneField Fields residuals
## 2005 0030881741 3.884 1997     2739      2     2.767
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##    Min     1Q Median     3Q    Max
## -1.222 -0.384  0.031  0.394  2.767
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)          1.17632    0.02850   41.27  < 2e-16 ***
## FirstAuthorFemale1  0.03849    0.00866    4.44  8.9e-06 ***
## LastAuthorFemale1  -0.04391    0.00860   -5.11  3.3e-07 ***
## Year1997           -0.05934    0.03711   -1.60   0.1099
## Year1998           -0.08886    0.03705   -2.40   0.0165 *
## Year1999           -0.07607    0.03600   -2.11   0.0346 *
## Year2000           -0.10565    0.03584   -2.95   0.0032 **
## Year2001           -0.05548    0.03586   -1.55   0.1219
## Year2002           -0.06379    0.03353   -1.90   0.0572 .
## Year2003           -0.10561    0.03403   -3.10   0.0019 **
## Year2004           -0.05669    0.03449   -1.64   0.1002
## Year2005           -0.05899    0.03354   -1.76   0.0786 .
```

```
## Year2006           -0.06099    0.03183   -1.92   0.0553 .
## Year2007           -0.03901    0.03169   -1.23   0.2184
## Year2008           -0.05913    0.03095   -1.91   0.0561 .
## Year2009           -0.09642    0.03100   -3.11   0.0019 **
## Year2010           -0.08374    0.03092   -2.71   0.0068 **
## Year2011           -0.05061    0.03083   -1.64   0.1007
## Year2012            0.00721    0.03148    0.23   0.8188
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.571
## Multiple R-squared:  0.00432,    Adjusted R-squared:  0.00349
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
##  observation 1214 is an outlier with |weight| = 0 ( < 4.6e-06);
##  1878 weights are ~= 1. The remaining 19841 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.137   0.855   0.949   0.903   0.985   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          4.60e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.012  1           1.006
## Year              1.012 16           1.000
```




```
## [1] "List of  1 outliers with residuals above  2.5"
##        ScopusId  NLCS Year OneField Fields residuals
## 2005 0030881741 3.884 1997     2739      2     2.767
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.1918 -0.3828  0.0304  0.3916  2.7783
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)          1.16487    0.02827   41.20   <2e-16 ***
## FirstAuthorFemale1  0.02161    0.00834    2.59   0.0096 **
## Year1997           -0.05912    0.03696   -1.60   0.1097
## Year1998           -0.08842    0.03698   -2.39   0.0168 *
## Year1999           -0.07667    0.03595   -2.13   0.0329 *
## Year2000           -0.10672    0.03575   -2.99   0.0028 **
## Year2001           -0.05546    0.03580   -1.55   0.1214
## Year2002           -0.06448    0.03345   -1.93   0.0540 .
## Year2003           -0.10790    0.03399   -3.17   0.0015 **
## Year2004           -0.05703    0.03446   -1.66   0.0979 .
## Year2005           -0.06010    0.03349   -1.79   0.0727 .
## Year2006           -0.06247    0.03175   -1.97   0.0491 *
```

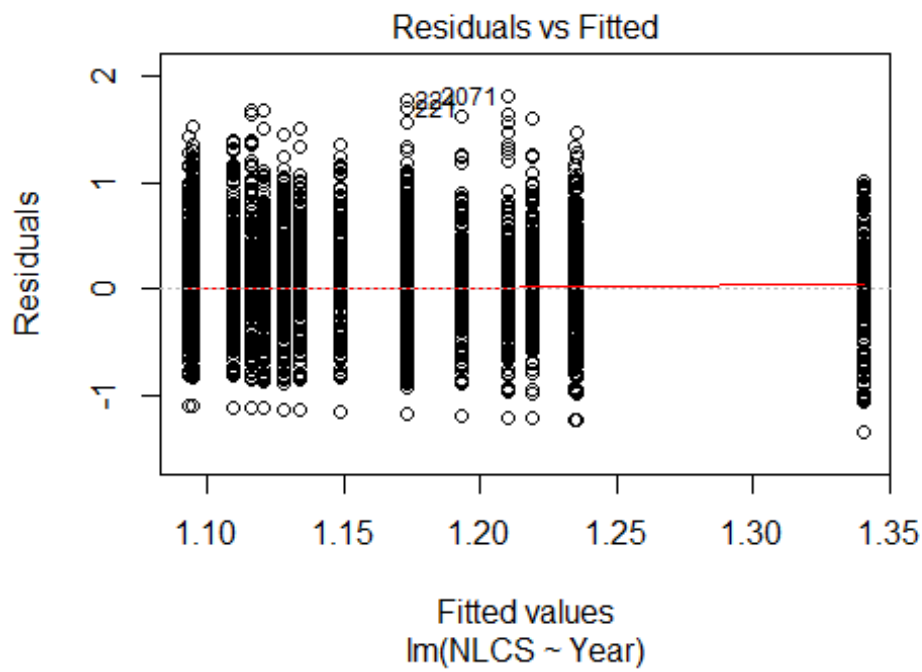
```
## Year2007           -0.03969    0.03161   -1.26   0.2092
## Year2008           -0.05976    0.03088   -1.93   0.0530 .
## Year2009           -0.09722    0.03093   -3.14   0.0017 **
## Year2010           -0.08545    0.03085   -2.77   0.0056 **
## Year2011           -0.05192    0.03073   -1.69   0.0911 .
## Year2012            0.00531    0.03139    0.17   0.8657
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.571
## Multiple R-squared:  0.00318,    Adjusted R-squared:  0.0024
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
##  observation 1214 is an outlier with |weight| = 0 ( < 4.6e-06);
##  1833 weights are ~= 1. The remaining 19886 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.132   0.856   0.949   0.903   0.985   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          4.60e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##                   GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.005  1           1.002
## Year             1.005 16           1.000
```



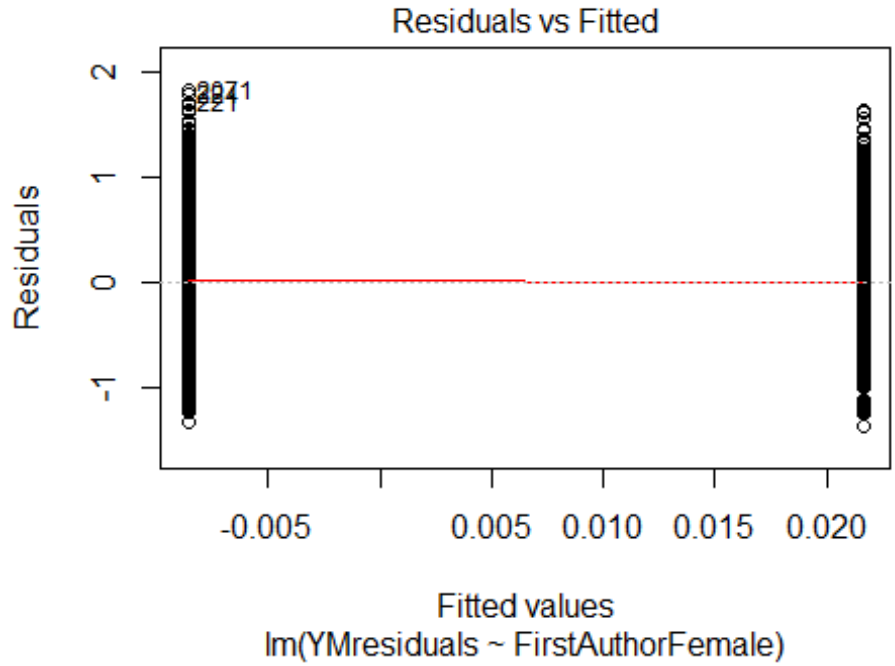

```
## [1] "List of  1 outliers with residuals above  2.5"
##        ScopusId  NLCS Year OneField Fields residuals
## 2005 0030881741 3.884 1997     2739      2     2.767
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.1986 -0.3858  0.0297  0.3940  2.7562
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.18812    0.02842   41.81  < 2e-16 ***
## LastAuthorFemale1 -0.02930    0.00829   -3.54  0.00041 ***
## Year1997          -0.06031    0.03715   -1.62  0.10450
## Year1998          -0.08864    0.03715   -2.39  0.01705 *
## Year1999          -0.07656    0.03607   -2.12  0.03380 *
## Year2000          -0.10534    0.03589   -2.94  0.00333 **
## Year2001          -0.05545    0.03590   -1.54  0.12253
## Year2002          -0.06435    0.03358   -1.92  0.05537 .
## Year2003          -0.10380    0.03409   -3.04  0.00233 **
## Year2004          -0.05508    0.03455   -1.59  0.11094
## Year2005          -0.05775    0.03361   -1.72  0.08578 .
## Year2006          -0.05820    0.03187   -1.83  0.06785 .
```

```
## Year2007          -0.03657    0.03175   -1.15  0.24933
## Year2008          -0.05679    0.03103   -1.83  0.06726 .
## Year2009          -0.09359    0.03106   -3.01  0.00259 **
## Year2010          -0.08059    0.03098   -2.60  0.00930 **
## Year2011          -0.04682    0.03088   -1.52  0.12941
## Year2012           0.01050    0.03154    0.33  0.73929
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.571
## Multiple R-squared:  0.00347,    Adjusted R-squared:  0.00268
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
##  observation 1214 is an outlier with |weight| = 0 ( < 4.6e-06);
##  1853 weights are ~= 1. The remaining 19866 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.143   0.856   0.949   0.903   0.985   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          4.60e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Sample size for the above analysis:  21720"
## [1] ""
## [1] ""
## [1] "####################################"
## [1] "Analysis of AJSC 3307"
## [1] "####################################"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  215  219  228  252  265  249  233  248  270  281  347  357  388  334  350
## 2011 2012
##  334  303
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  170  177  174  180  181  173  175  194  218  214  250  261  290  233  256
```

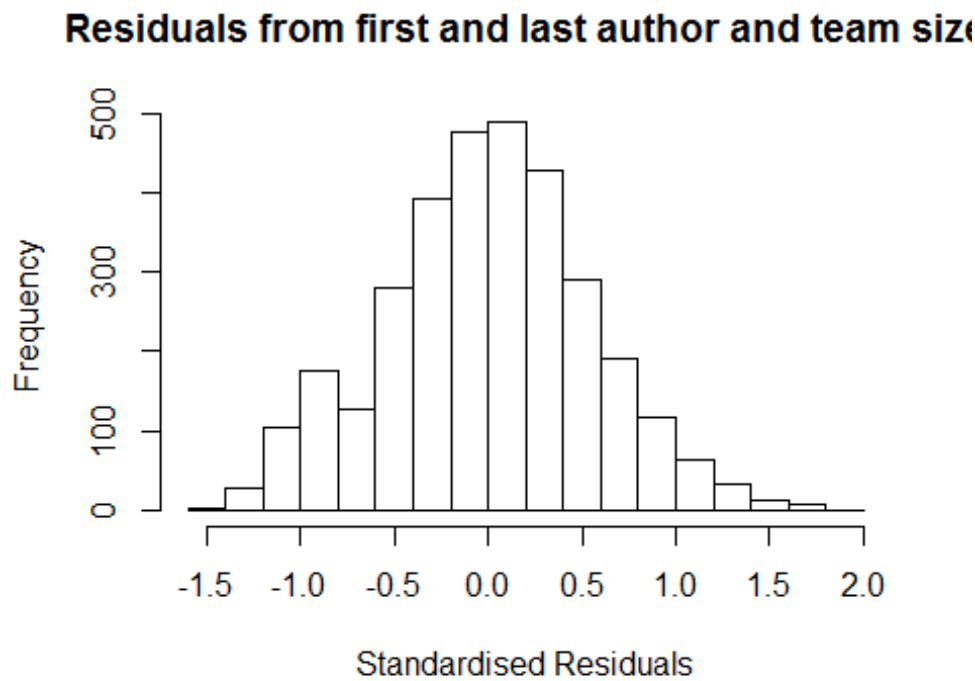
```
## 2011 2012
##  243  220
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  156  164  163  159  170  161  157  173  199  181  221  227  257  202  220
## 2011 2012
##  209  198
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
##  Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 35, df = 16, p-value = 0.004
```




```
##
##  Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 0.029, df = 1, p-value = 0.9
```

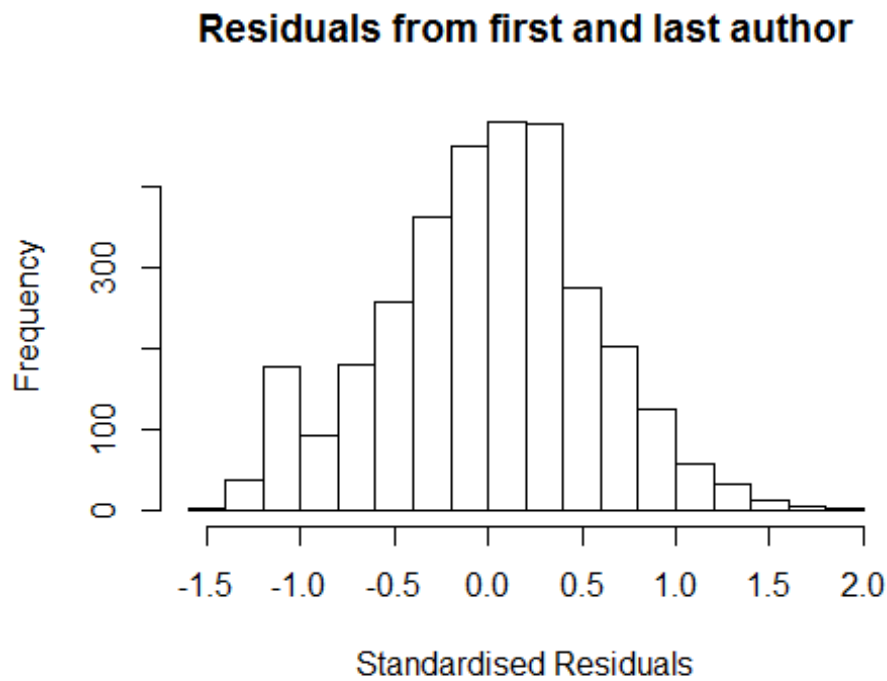



```
## [1] "Female first author team size 2018 geometric mean: 2.79586371478864"
## [1] "Male first author team size 2018 geometric mean: 2.46345041770524"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 7000, p-value = 0.1
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 2.78975913085681"
## [1] "Male last author team size 2018 geometric mean: 2.49232748561726"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 6300, p-value = 0.2
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.093  1           1.046
## LastAuthorFemale  1.090  1           1.044
## UniqueAuthors     1.116  4           1.014
## Year              1.162 16           1.005
```



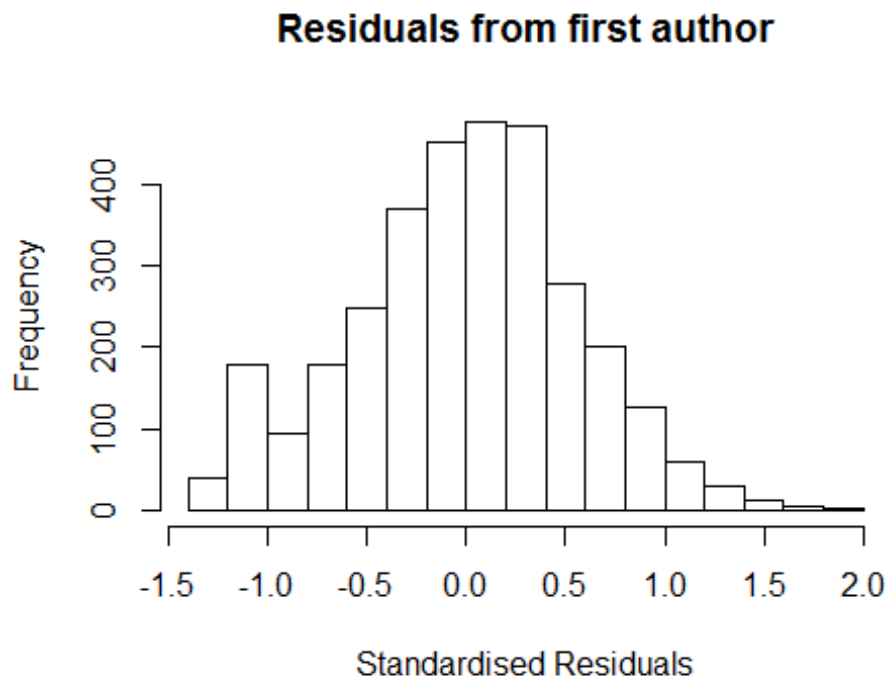

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
UniqueAuthors +
##     Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##     k.max = 1000))
##  \--> method = "MM"
## Residuals:
##    Min     1Q Median     3Q    Max
## -1.560 -0.346  0.014  0.355  1.902
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         0.95596    0.05796   16.49   <2e-16 ***
## FirstAuthorFemale1  0.00642    0.02300    0.28    0.780
## LastAuthorFemale1  -0.02171    0.02348   -0.92    0.355
## UniqueAuthors2      0.25617    0.02884    8.88   <2e-16 ***
## UniqueAuthors3      0.27123    0.03055    8.88   <2e-16 ***
## UniqueAuthors4      0.28502    0.03292    8.66   <2e-16 ***
## UniqueAuthors5      0.36544    0.03350   10.91   <2e-16 ***
## Year1997            0.08542    0.07168    1.19    0.233
## Year1998            0.06186    0.07190    0.86    0.390
## Year1999            0.04848    0.06823    0.71    0.477
```

```
## Year2000            0.11982    0.06957    1.72    0.085 .
## Year2001            0.23217    0.07068    3.28    0.001 **
## Year2002            0.09567    0.07211    1.33    0.185
## Year2003            0.01731    0.06760    0.26    0.798
## Year2004           -0.06319    0.06872   -0.92    0.358
## Year2005           -0.06457    0.07106   -0.91    0.364
## Year2006           -0.03259    0.06564   -0.50    0.620
## Year2007           -0.06105    0.06498   -0.94    0.348
## Year2008           -0.06789    0.06271   -1.08    0.279
## Year2009           -0.03045    0.06627   -0.46    0.646
## Year2010           -0.09500    0.06621   -1.43    0.151
## Year2011           -0.13834    0.06936   -1.99    0.046 *
## Year2012           -0.05009    0.06993   -0.72    0.474
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.529
## Multiple R-squared:  0.0729, Adjusted R-squared:  0.0665
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  238 weights are ~= 1. The remaining 2979 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.169   0.864   0.953   0.902   0.987   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          3.11e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.078  1           1.038
## LastAuthorFemale  1.082  1           1.040
## Year              1.057 16           1.002
```



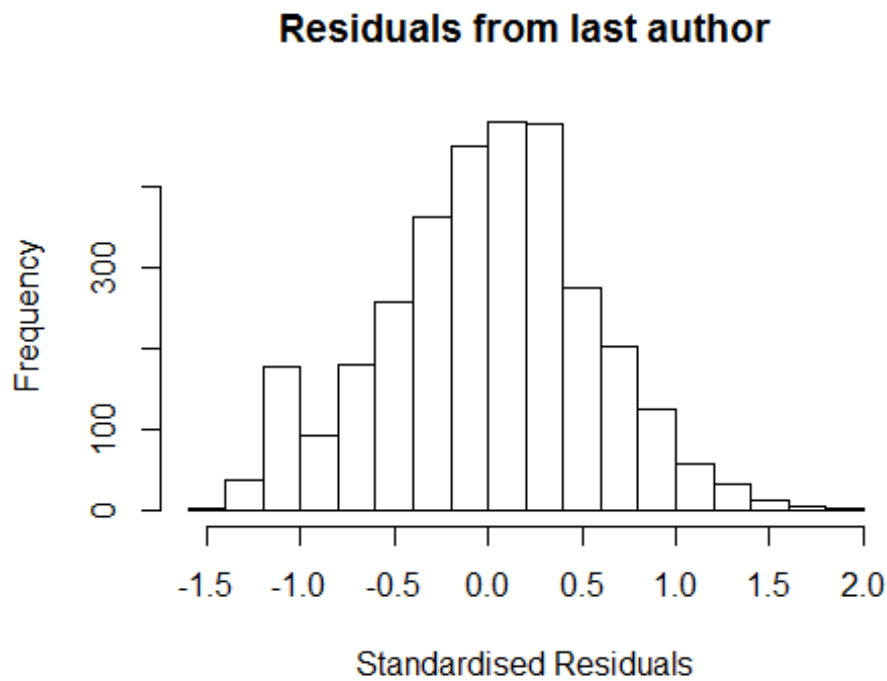

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.4011 -0.3631  0.0212  0.3522  1.8586
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)          1.13791    0.05814   19.57   <2e-16 ***
## FirstAuthorFemale1   0.02842    0.02334    1.22   0.2236
## LastAuthorFemale1   -0.01423    0.02385   -0.60   0.5509
## Year1997             0.10998    0.07518    1.46   0.1436
## Year1998             0.04495    0.07430    0.60   0.5453
## Year1999             0.05143    0.07123    0.72   0.4703
## Year2000             0.11985    0.07285    1.65   0.1000
## Year2001             0.23479    0.07295    3.22   0.0013 **
## Year2002             0.09917    0.07436    1.33   0.1824
## Year2003             0.02945    0.07176    0.41   0.6815
## Year2004            -0.04852    0.07244   -0.67   0.5030
## Year2005            -0.03710    0.07447   -0.50   0.6184
```

```
## Year2006           -0.01878    0.06820   -0.28   0.7830
## Year2007           -0.03355    0.06818   -0.49   0.6226
## Year2008           -0.05121    0.06548   -0.78   0.4342
## Year2009           -0.00666    0.06941   -0.10   0.9235
## Year2010           -0.06925    0.07020   -0.99   0.3239
## Year2011           -0.10378    0.07308   -1.42   0.1557
## Year2012            0.00523    0.07368    0.07   0.9434
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.538
## Multiple R-squared:  0.0201, Adjusted R-squared:  0.0146
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
##  259 weights are ~= 1. The remaining 2958 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.207   0.858   0.953   0.899   0.986   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          3.11e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.029  1           1.014
## Year              1.029 16           1.001
```




```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.3937 -0.3595  0.0223  0.3505  1.8610
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)          1.13604    0.05807   19.56   <2e-16 ***
## FirstAuthorFemale1   0.02434    0.02289    1.06   0.2876
## Year1997             0.10949    0.07517    1.46   0.1454
## Year1998             0.04552    0.07424    0.61   0.5398
## Year1999             0.05000    0.07111    0.70   0.4821
## Year2000             0.12004    0.07284    1.65   0.0994 .
## Year2001             0.23336    0.07283    3.20   0.0014 **
## Year2002             0.09853    0.07431    1.33   0.1850
## Year2003             0.02900    0.07176    0.40   0.6862
## Year2004            -0.04881    0.07245   -0.67   0.5005
## Year2005            -0.03773    0.07446   -0.51   0.6124
## Year2006            -0.01898    0.06819   -0.28   0.7808
```

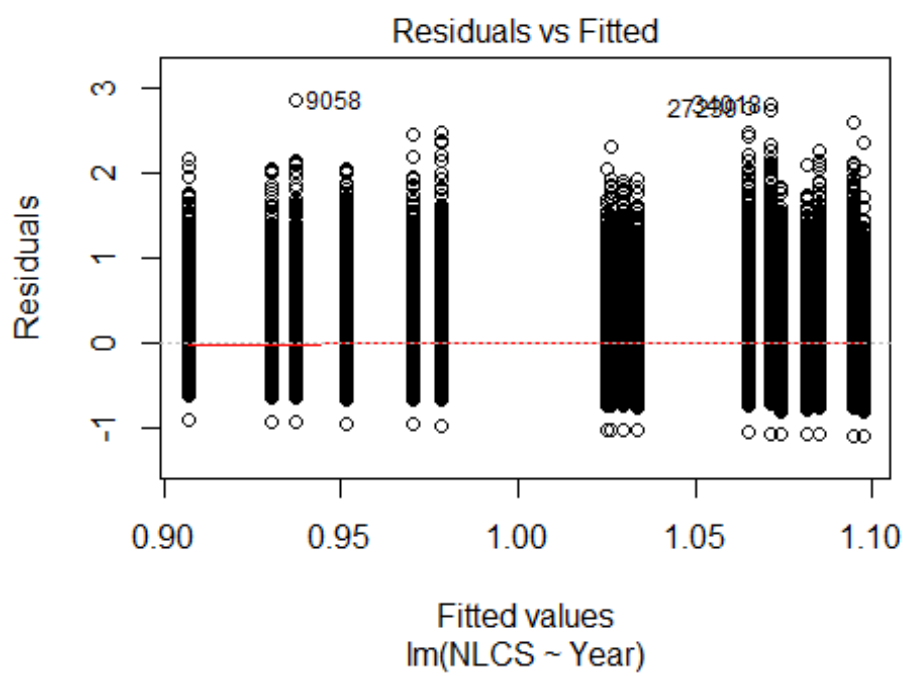
```
## Year2007           -0.03459    0.06815   -0.51   0.6118
## Year2008           -0.05240    0.06542   -0.80   0.4232
## Year2009           -0.00757    0.06939   -0.11   0.9131
## Year2010           -0.07052    0.07010   -1.01   0.3145
## Year2011           -0.10423    0.07305   -1.43   0.1537
## Year2012            0.00406    0.07365    0.06   0.9560
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.538
## Multiple R-squared:  0.02,   Adjusted R-squared:  0.0148
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
##  261 weights are ~= 1. The remaining 2956 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.207   0.859   0.953   0.900   0.986   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          3.11e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##                   GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.032  1           1.016
## Year             1.032 16           1.001
```



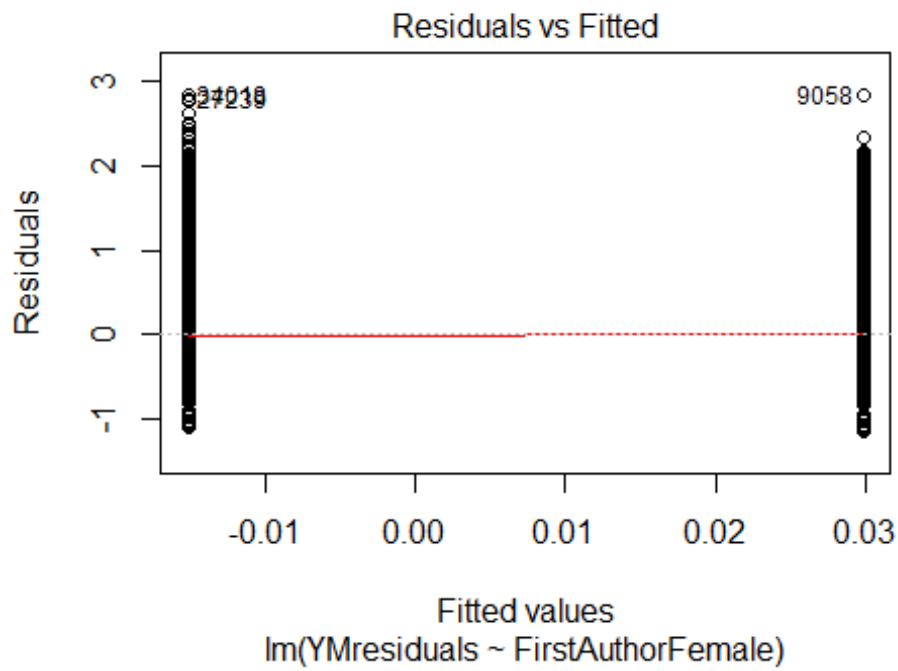

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.3782 -0.3588  0.0237  0.3519  1.8551
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.14254    0.05828   19.60   <2e-16 ***
## LastAuthorFemale1 -0.00556    0.02336   -0.24   0.8120
## Year1997            0.10956    0.07539    1.45   0.1463
## Year1998            0.04507    0.07444    0.61   0.5449
## Year1999            0.05159    0.07146    0.72   0.4704
## Year2000            0.12071    0.07314    1.65   0.0990 .
## Year2001            0.23565    0.07324    3.22   0.0013 **
## Year2002            0.09912    0.07457    1.33   0.1839
## Year2003            0.02834    0.07199    0.39   0.6939
## Year2004           -0.04792    0.07262   -0.66   0.5095
## Year2005           -0.03389    0.07442   -0.46   0.6488
## Year2006           -0.01773    0.06839   -0.26   0.7954
```

```
## Year2007          -0.03190    0.06830   -0.47   0.6405
## Year2008          -0.04942    0.06567   -0.75   0.4518
## Year2009          -0.00435    0.06967   -0.06   0.9503
## Year2010          -0.06726    0.07044   -0.95   0.3397
## Year2011          -0.10072    0.07317   -1.38   0.1688
## Year2012           0.00811    0.07375    0.11   0.9124
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.537
## Multiple R-squared:  0.0197, Adjusted R-squared:  0.0145
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
##  268 weights are ~= 1. The remaining 2949 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.209   0.857   0.952   0.899   0.985   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          3.11e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Sample size for the above analysis:  3217"
## [1] ""
## [1] ""
## [1] "####################################"
## [1] "Analysis of AJSC 3308"
## [1] "####################################"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 1572 1590 1535 1453 1729 1675 1539 1290 1278 1524 1605 1924 2271 3116 3654
## 2011 2012
## 3377 3276
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 1371 1390 1199 1100 1337 1380 1172 1165 1138 1327 1415 1681 1981 2733 3202
## 2011 2012
```

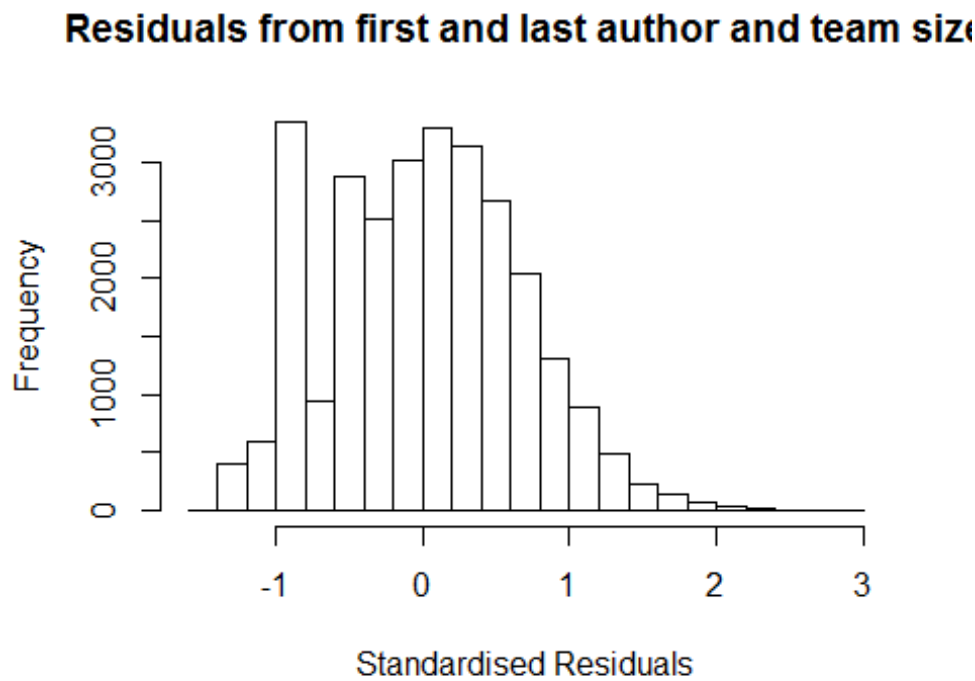
```
## 2979 2878
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 1329 1329 1151 1057 1287 1317 1123 1109 1063 1263 1333 1586 1862 2570 3070
## 2011 2012
## 2822 2712
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
##  Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 140, df = 16, p-value <2e-16
```




```
##
##  Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 18, df = 1, p-value = 3e-05
```



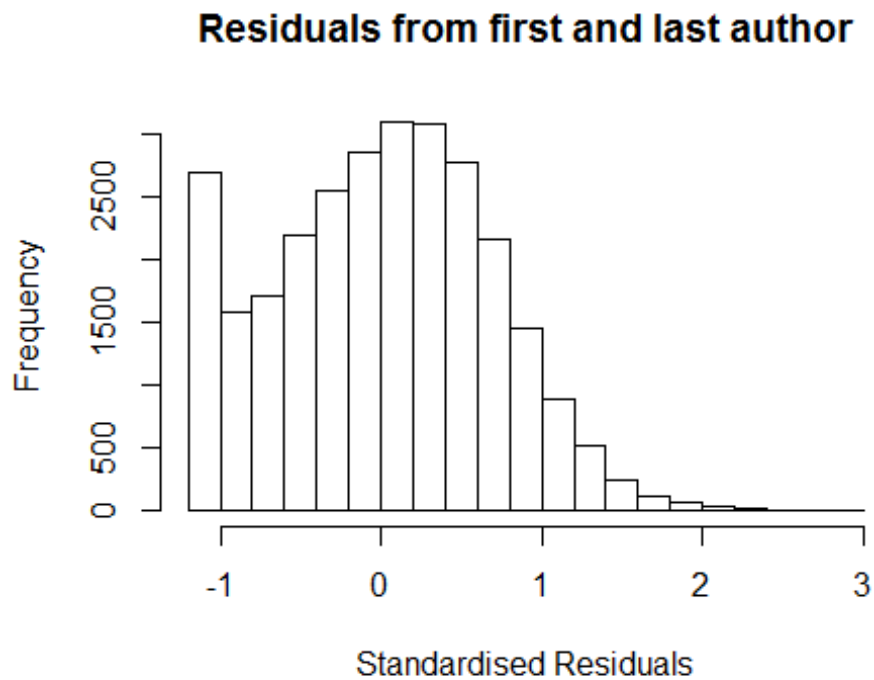

```
## [1] "Female first author team size 2018 geometric mean: 1.88166637782127"
## [1] "Male first author team size 2018 geometric mean: 1.61542686266426"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 830000, p-value = 7e-09
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 1.822712691572"
## [1] "Male last author team size 2018 geometric mean: 1.65838425717462"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 780000, p-value = 3e-04
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##                     GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.679  1           1.296
## LastAuthorFemale  1.667  1           1.291
## UniqueAuthors     1.053  4           1.006
## Year              1.044 16           1.001
```




```
## [1] "List of  6 outliers with residuals above  2.5"
##          ScopusId  NLCS Year OneField Fields residuals
## 1432  21344445311 3.461 1996     3308      3     2.611
## 19476 33845535808 3.450 2007     1200      2     2.507
## 27045 77957594024 3.536 2010     3308      1     2.628
## 27239 77953568642 3.825 2010     3308      3     2.528
## 34018 84866279731 3.877 2012     3308      3     2.979
## 36238 83055186560 3.802 2012     3308      3     2.516
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
UniqueAuthors +
##     Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##     k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.4326 -0.4803  0.0202  0.4568  2.9788
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         0.84984    0.01985   42.81  < 2e-16 ***
## FirstAuthorFemale1  0.04542    0.01100    4.13  3.6e-05 ***
## LastAuthorFemale1  -0.03521    0.01107   -3.18  0.00147 **
## UniqueAuthors2      0.27507    0.01050   26.20  < 2e-16 ***
## UniqueAuthors3      0.38824    0.01362   28.50  < 2e-16 ***
```

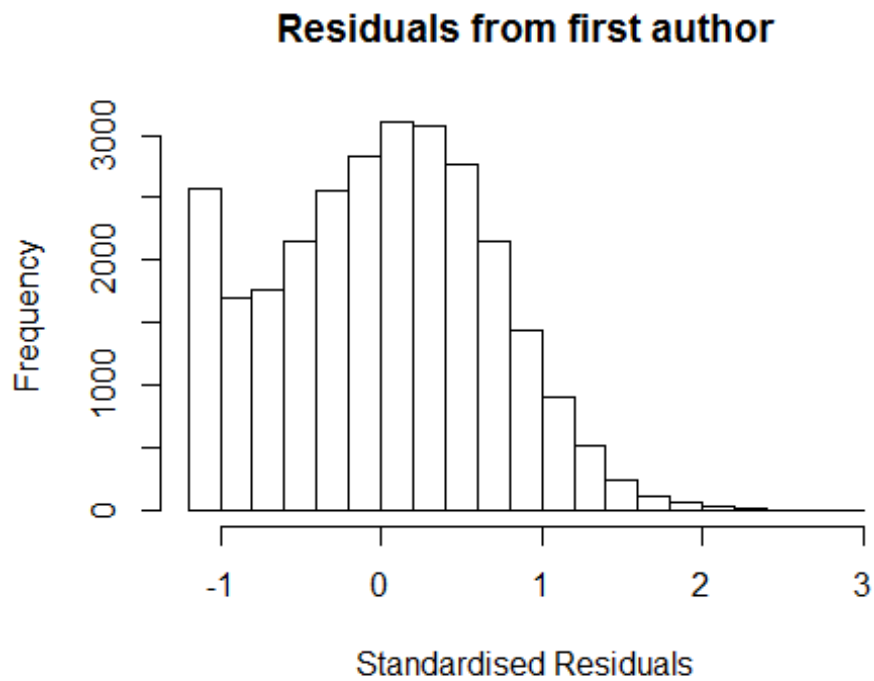
```
## UniqueAuthors4      0.42870    0.01739   24.66  < 2e-16 ***
## UniqueAuthors5      0.44841    0.01951   22.98  < 2e-16 ***
## Year1997           -0.00966    0.02728   -0.35  0.72335
## Year1998           -0.03936    0.02734   -1.44  0.14992
## Year1999           -0.06081    0.02807   -2.17  0.03028 *
## Year2000           -0.02161    0.02672   -0.81  0.41852
## Year2001           -0.05037    0.02760   -1.83  0.06797 .
## Year2002            0.03494    0.02782    1.26  0.20910
## Year2003            0.01379    0.02821    0.49  0.62506
## Year2004            0.01734    0.02747    0.63  0.52803
## Year2005            0.02827    0.02728    1.04  0.30006
## Year2006            0.05817    0.02547    2.28  0.02237 *
## Year2007            0.08298    0.02489    3.33  0.00086 ***
## Year2008            0.06273    0.02475    2.53  0.01126 *
## Year2009            0.07325    0.02400    3.05  0.00228 **
## Year2010            0.05844    0.02358    2.48  0.01321 *
## Year2011            0.08898    0.02402    3.70  0.00021 ***
## Year2012            0.04833    0.02419    2.00  0.04572 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.689
## Multiple R-squared:  0.0682, Adjusted R-squared:  0.0674
## Convergence in 10 IRWLS iterations
##
## Robustness weights:
##  2321 weights are ~= 1. The remaining 25662 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0219  0.8660  0.9520  0.9160  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          3.57e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.707  1           1.306
```

```
## LastAuthorFemale  1.705  1           1.306
## Year              1.009 16           1.000
```



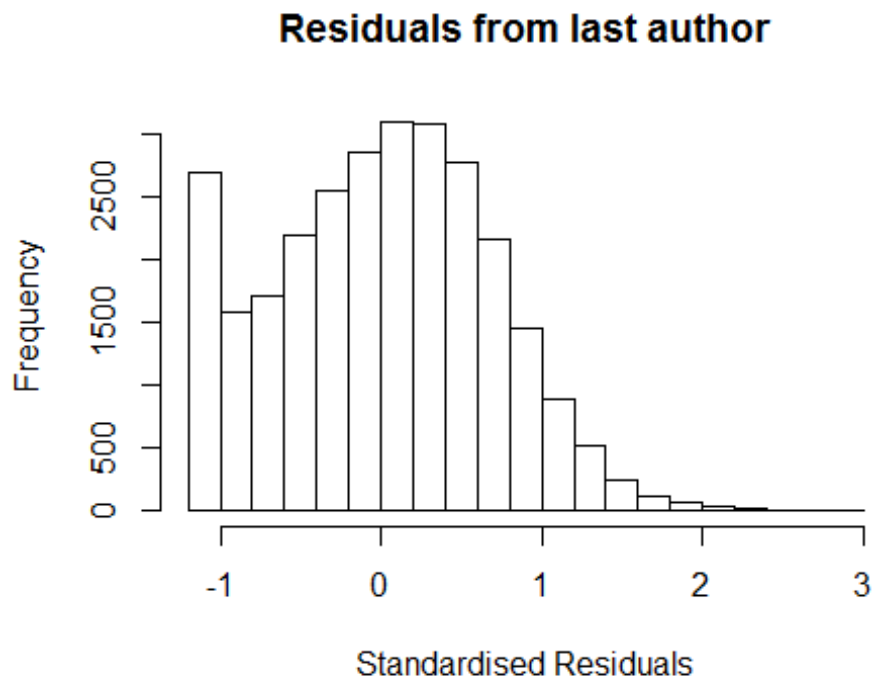

```
## [1] "List of  6 outliers with residuals above  2.5"
##          ScopusId  NLCS Year OneField Fields residuals
## 1432  21344445311 3.461 1996     3308      3     2.532
## 27045 77957594024 3.536 2010     3308      1     2.505
## 27239 77953568642 3.825 2010     3308      3     2.794
## 32584 84857663622 3.693 2011     3308      1     2.628
## 34018 84866279731 3.877 2012     3308      3     2.845
## 36238 83055186560 3.802 2012     3308      3     2.770
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.1461 -0.5185  0.0251  0.4845  2.8452
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)          0.92945    0.02028   45.83  < 2e-16 ***
## FirstAuthorFemale1   0.06747    0.01156    5.84  5.3e-09 ***
## LastAuthorFemale1   -0.02348    0.01165   -2.01   0.0440 *
## Year1997             0.00353    0.02826    0.12   0.9006
```

```
## Year1998           -0.03195    0.02837   -1.13   0.2600
## Year1999           -0.05820    0.02869   -2.03   0.0425 *
## Year2000           -0.01421    0.02748   -0.52   0.6050
## Year2001           -0.03654    0.02863   -1.28   0.2019
## Year2002            0.06560    0.02848    2.30   0.0213 *
## Year2003            0.06946    0.02944    2.36   0.0183 *
## Year2004            0.07543    0.02860    2.64   0.0084 **
## Year2005            0.07221    0.02818    2.56   0.0104 *
## Year2006            0.12383    0.02629    4.71  2.5e-06 ***
## Year2007            0.14918    0.02574    5.80  6.8e-09 ***
## Year2008            0.11971    0.02564    4.67  3.0e-06 ***
## Year2009            0.12741    0.02466    5.17  2.4e-07 ***
## Year2010            0.10135    0.02421    4.19  2.8e-05 ***
## Year2011            0.13554    0.02462    5.51  3.7e-08 ***
## Year2012            0.10231    0.02486    4.11  3.9e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.721
## Multiple R-squared:  0.00977,    Adjusted R-squared:  0.00913
## Convergence in 10 IRWLS iterations
##
## Robustness weights:
##  2200 weights are ~= 1. The remaining 25783 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0849  0.8640  0.9490  0.9190  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          3.57e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.007  1           1.004
## Year              1.007 16           1.000
```




```
## [1] "List of  6 outliers with residuals above  2.5"
##          ScopusId  NLCS Year OneField Fields residuals
## 1432  21344445311 3.461 1996     3308      3     2.532
## 27045 77957594024 3.536 2010     3308      1     2.505
## 27239 77953568642 3.825 2010     3308      3     2.794
## 32584 84857663622 3.693 2011     3308      1     2.628
## 34018 84866279731 3.877 2012     3308      3     2.845
## 36238 83055186560 3.802 2012     3308      3     2.770
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.1279 -0.5171  0.0273  0.4860  2.8480
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)          0.92740    0.02026   45.78  < 2e-16 ***
## FirstAuthorFemale1   0.05194    0.00889    5.84  5.3e-09 ***
## Year1997             0.00324    0.02826    0.11   0.9088
## Year1998            -0.03207    0.02837   -1.13   0.2582
## Year1999            -0.05848    0.02870   -2.04   0.0416 *
## Year2000            -0.01426    0.02749   -0.52   0.6038
## Year2001            -0.03731    0.02862   -1.30   0.1923
```

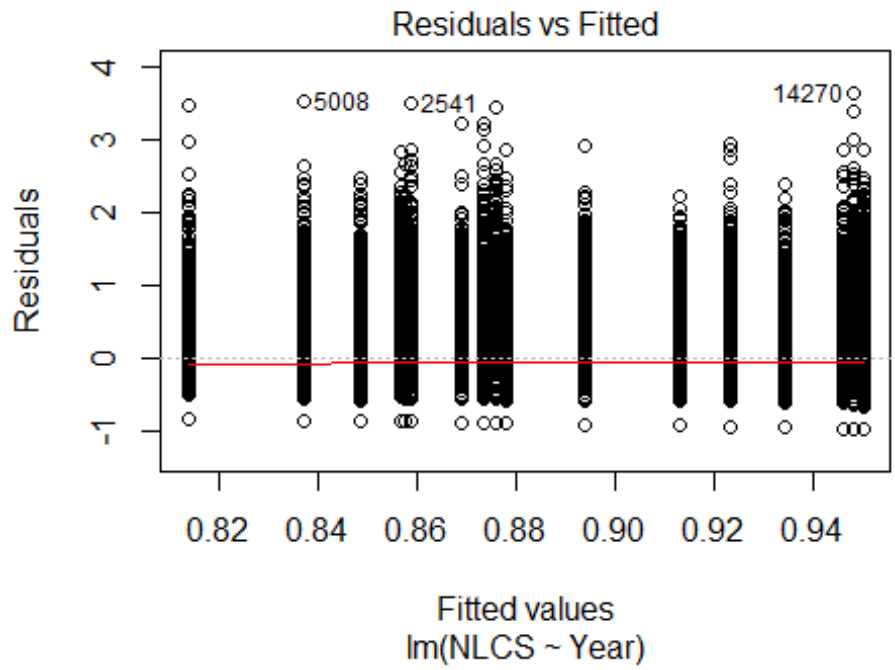
```
## Year2002            0.06483    0.02848    2.28   0.0229 *
## Year2003            0.06868    0.02944    2.33   0.0197 *
## Year2004            0.07494    0.02860    2.62   0.0088 **
## Year2005            0.07181    0.02819    2.55   0.0108 *
## Year2006            0.12307    0.02629    4.68  2.9e-06 ***
## Year2007            0.14856    0.02574    5.77  7.9e-09 ***
## Year2008            0.11932    0.02564    4.65  3.3e-06 ***
## Year2009            0.12645    0.02465    5.13  2.9e-07 ***
## Year2010            0.10079    0.02420    4.16  3.1e-05 ***
## Year2011            0.13456    0.02461    5.47  4.6e-08 ***
## Year2012            0.10156    0.02486    4.08  4.4e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.721
## Multiple R-squared:  0.00963,    Adjusted R-squared:  0.00903
## Convergence in 10 IRWLS iterations
##
## Robustness weights:
##  2227 weights are ~= 1. The remaining 25756 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.084   0.865   0.949   0.919   0.985   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          3.57e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##                   GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.006  1           1.003
## Year             1.006 16           1.000
```



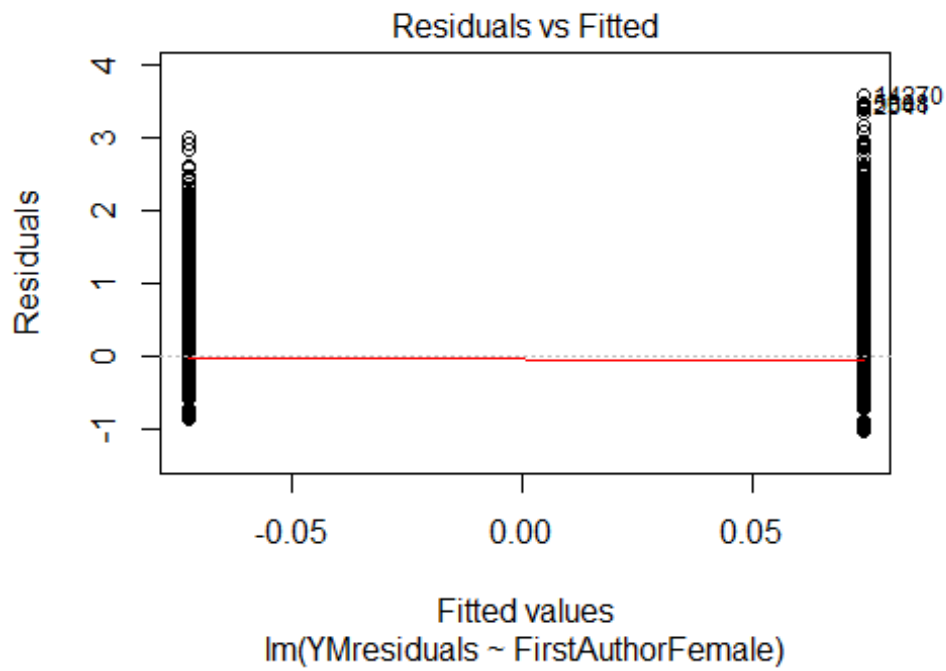

```
## [1] "List of  6 outliers with residuals above  2.5"
##          ScopusId  NLCS Year OneField Fields residuals
## 1432  21344445311 3.461 1996     3308      3     2.532
## 27045 77957594024 3.536 2010     3308      1     2.505
## 27239 77953568642 3.825 2010     3308      3     2.794
## 32584 84857663622 3.693 2011     3308      1     2.628
## 34018 84866279731 3.877 2012     3308      3     2.845
## 36238 83055186560 3.802 2012     3308      3     2.770
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.1081 -0.5230  0.0279  0.4856  2.8369
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         0.93597    0.02026   46.19  < 2e-16 ***
## LastAuthorFemale1   0.02129    0.00893    2.38   0.0171 *
## Year1997            0.00306    0.02828    0.11   0.9138
## Year1998           -0.03206    0.02837   -1.13   0.2585
## Year1999           -0.05681    0.02873   -1.98   0.0480 *
## Year2000           -0.01281    0.02752   -0.47   0.6416
## Year2001           -0.03561    0.02861   -1.24   0.2132
```

```
## Year2002           0.06609    0.02850    2.32   0.0204 *
## Year2003           0.06928    0.02946    2.35   0.0187 *
## Year2004           0.07661    0.02861    2.68   0.0074 **
## Year2005           0.07373    0.02824    2.61   0.0090 **
## Year2006           0.12491    0.02631    4.75  2.1e-06 ***
## Year2007           0.15081    0.02575    5.86  4.7e-09 ***
## Year2008           0.12220    0.02566    4.76  1.9e-06 ***
## Year2009           0.12814    0.02467    5.19  2.1e-07 ***
## Year2010           0.10340    0.02422    4.27  2.0e-05 ***
## Year2011           0.13727    0.02463    5.57  2.5e-08 ***
## Year2012           0.10411    0.02487    4.19  2.9e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.722
## Multiple R-squared:  0.00861,    Adjusted R-squared:  0.008
## Convergence in 10 IRWLS iterations
##
## Robustness weights:
##  2227 weights are ~= 1. The remaining 25756 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.088   0.863   0.949   0.919   0.985   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          3.57e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Sample size for the above analysis:  27983"
## [1] ""
## [1] ""
## [1] "####################################"
## [1] "Analysis of AJSC 3309"
## [1] "####################################"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 1119 1137 1096 1159 1329 1557 1473 1157 1337 1319 1307 1594 1832 1826 1973
## 2011 2012
```

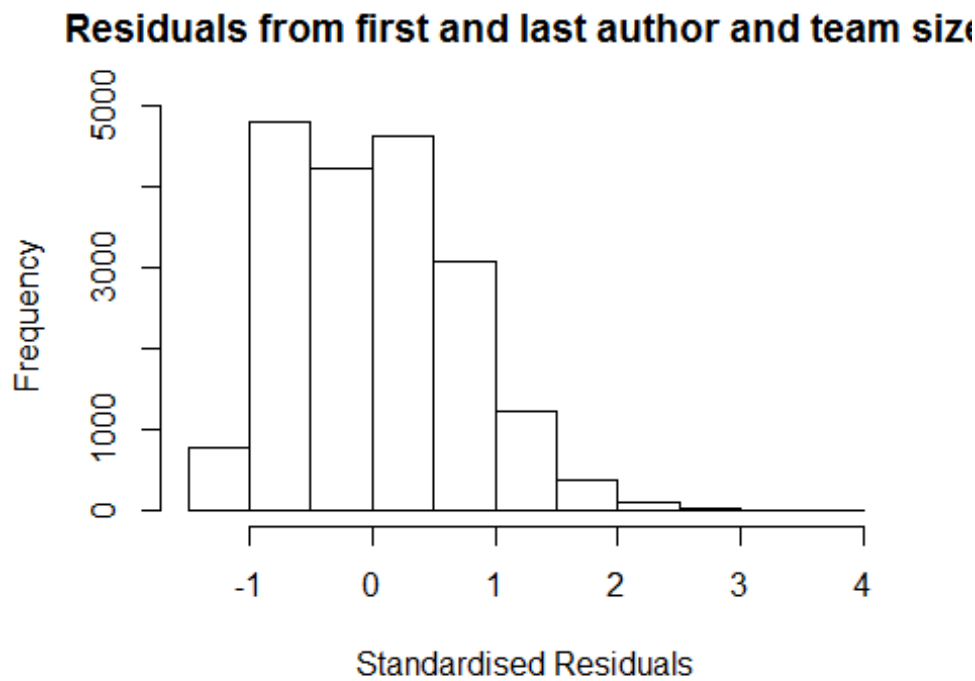
```
## 1911 1919
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  967  874  892  979 1131 1283 1274  981 1109 1098 1048 1258 1498 1476 1592
## 2011 2012
## 1520 1530
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  926  834  853  928 1072 1211 1200  935 1042 1033  970 1173 1390 1386 1485
## 2011 2012
## 1419 1400
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
##  Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 110, df = 16, p-value <2e-16
```




```
##
##  Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 190, df = 1, p-value <2e-16
```



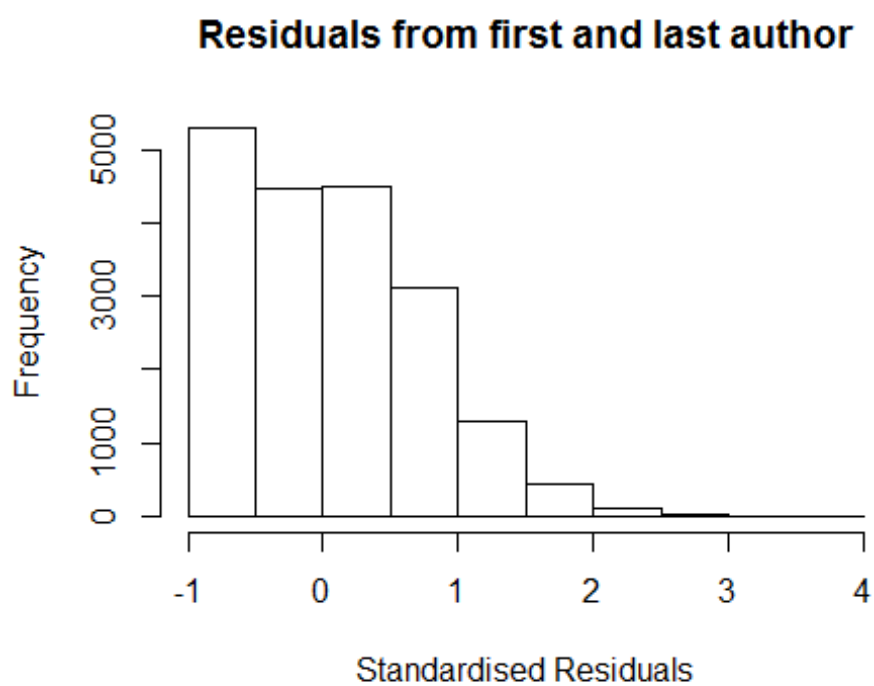

```
## [1] "Female first author team size 2018 geometric mean: 1.74227934367565"
## [1] "Male first author team size 2018 geometric mean: 1.85006176769137"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 140000, p-value = 0.1
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 1.73159617394141"
## [1] "Male last author team size 2018 geometric mean: 1.86920699366452"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 140000, p-value = 0.08
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 2.396  1           1.548
## LastAuthorFemale  2.400  1           1.549
## UniqueAuthors     1.065  4           1.008
## Year              1.064 16           1.002
```




```
## [1] "List of  25 outliers with residuals above  2.5"
##          ScopusId  NLCS Year OneField Fields residuals
## 416    0030240869 3.699 1996     1710      2     3.094
## 1254  35048881580 4.306 1997     3309      1     3.066
## 2548   0032028728 3.228 1998     1706      3     2.512
## 3362   0032339276 3.604 1998     1705      4     2.594
## 3380   0032447095 3.733 1998     1710      3     2.657
## 4062   0033449090 3.235 1999     3309      1     2.637
## 4998   0034188649 3.199 2000     1706      3     2.636
## 5041   0034315068 3.331 2000     1706      3     2.635
## 6117   0034559541 3.488 2000     1404      5     2.792
## 8416   0036611573 3.387 2002     1706      3     2.591
## 8419   0036611602 3.834 2002     1706      3     2.552
## 8733   0036005221 3.297 2002     3309      1     2.634
## 9145   0036015968 3.391 2002     1705      4     2.595
## 10794  5444237123 3.544 2004     1706      3     2.833
## 12088 29144439194 3.812 2005     1706      3     2.712
## 13526 33947416035 3.963 2006     1706      3     2.856
## 13607 33750593475 3.353 2006     1706      3     2.607
## 14270 33645712892 4.593 2006     1706      3     3.847
## 14418 31744440684 4.357 2006     1706      3     3.125
## 14980 64649083745 4.114 2007     1706      3     3.103
## 17387 48849113556 3.758 2008     1706      3     2.672
## 21253 77953568642 3.825 2010     3308      3     2.563
## 25046 84866279731 3.877 2012     3308      3     3.143
## 26460 83055176288 3.332 2012     3308      3     2.598
## 26463 83055186560 3.802 2012     3308      3     2.583
```

```
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
UniqueAuthors +
##     Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##     k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.3681 -0.5849 -0.0216  0.4989  3.8467
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         0.73735    0.02713   27.18  < 2e-16 ***
## FirstAuthorFemale1 -0.06581    0.01590   -4.14  3.5e-05 ***
## LastAuthorFemale1  -0.06699    0.01586   -4.22  2.4e-05 ***
## UniqueAuthors2      0.36027    0.01331   27.07  < 2e-16 ***
## UniqueAuthors3      0.48569    0.01954   24.86  < 2e-16 ***
## UniqueAuthors4      0.53924    0.03208   16.81  < 2e-16 ***
## UniqueAuthors5      0.59151    0.03120   18.96  < 2e-16 ***
## Year1997           -0.02178    0.03626   -0.60    0.548
## Year1998           -0.02180    0.03578   -0.61    0.542
## Year1999           -0.00655    0.03567   -0.18    0.854
## Year2000           -0.04131    0.03419   -1.21    0.227
## Year2001            0.02277    0.03459    0.66    0.510
## Year2002            0.05859    0.03448    1.70    0.089 .
## Year2003           -0.02164    0.03531   -0.61    0.540
## Year2004           -0.02675    0.03470   -0.77    0.441
## Year2005            0.00275    0.03388    0.08    0.935
## Year2006            0.00894    0.03476    0.26    0.797
## Year2007           -0.01963    0.03236   -0.61    0.544
## Year2008           -0.01193    0.03184   -0.37    0.708
## Year2009            0.02062    0.03170    0.65    0.515
## Year2010            0.03928    0.03188    1.23    0.218
## Year2011            0.04456    0.03210    1.39    0.165
## Year2012           -0.00379    0.03320   -0.11    0.909
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.699
## Multiple R-squared:  0.0942, Adjusted R-squared:  0.0931
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
##  observation 10719 is an outlier with |weight| = 0 ( < 5.2e-06);
##  1380 weights are ~= 1. The remaining 17876 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0064  0.8950  0.9370  0.9120  0.9830  0.9990
## Algorithmic parameters:
```

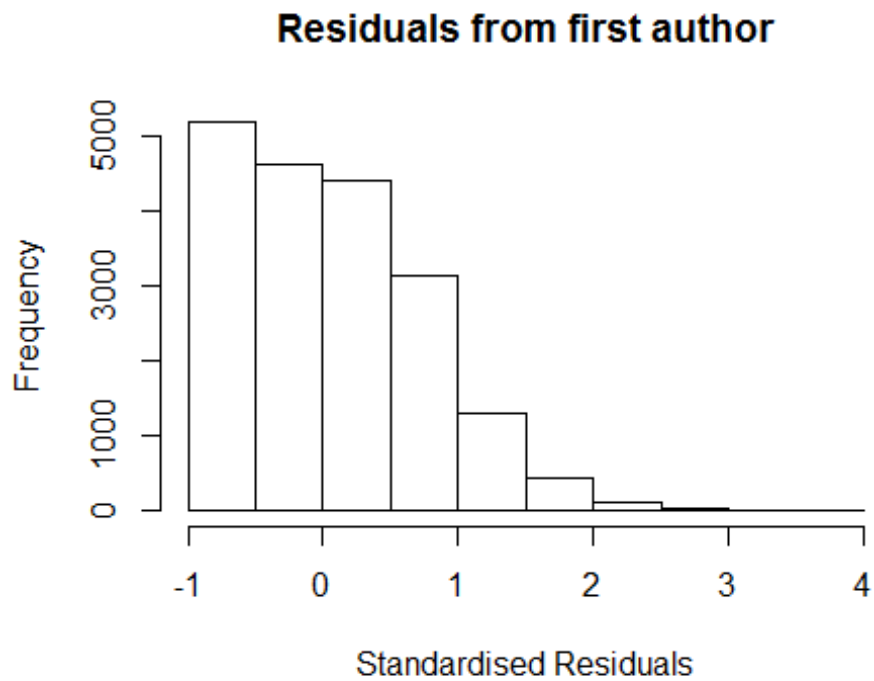
```
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          5.19e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 2.268  1           1.506
## LastAuthorFemale  2.264  1           1.505
## Year              1.010 16           1.000
```




```
## [1] "List of  31 outliers with residuals above  2.5"
##          ScopusId  NLCS Year OneField Fields residuals
## 271    0030290419 3.426 1996     1706      3     2.587
## 416    0030240869 3.699 1996     1710      2     2.995
## 1254  35048881580 4.306 1997     3309      1     3.574
```

```
## 1255  35248867366 3.351 1997     3309      1     2.565
## 2618   0032184399 3.521 1998     1706      3     2.681
## 2629   0032201980 3.365 1998     1706      3     2.525
## 3362   0032339276 3.604 1998     1705      4     2.805
## 3380   0032447095 3.733 1998     1710      3     2.893
## 3738   0032636505 3.564 1999     1706      3     2.710
## 4062   0033449090 3.235 1999     3309      1     2.516
## 4998   0034188649 3.199 2000     1706      3     2.507
## 5041   0034315068 3.331 2000     1706      3     2.504
## 6117   0034559541 3.488 2000     1404      5     2.661
## 6772   0002549585 3.549 2001     1710      2     2.670
## 8419   0036611602 3.834 2002     1706      3     2.888
## 10794  5444237123 3.544 2004     1706      3     2.680
## 12088 29144439194 3.812 2005     1706      3     2.903
## 13526 33947416035 3.963 2006     1706      3     3.034
## 13696 33947399169 3.483 2006     1706      3     2.594
## 14238 33644912045 3.569 2006     1710      2     2.640
## 14270 33645712892 4.593 2006     1706      3     3.664
## 14418 31744440684 4.357 2006     1706      3     3.428
## 14980 64649083745 4.114 2007     1706      3     3.321
## 15826 34250855167 3.383 2007     1500      4     2.631
## 17387 48849113556 3.758 2008     1706      3     2.858
## 17601 48849096944 3.369 2008     1710      2     2.510
## 21253 77953568642 3.825 2010     3308      3     2.855
## 21687 77950244328 3.504 2010     1706      3     2.534
## 25046 84866279731 3.877 2012     3308      3     2.933
## 25284 84867288643 3.693 2012     1606      4     2.749
## 26463 83055186560 3.802 2012     3308      3     2.858
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -0.9765 -0.5984 -0.0123  0.5210  3.6636
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         0.83881    0.02777   30.21  < 2e-16 ***
## FirstAuthorFemale1 -0.04064    0.01624   -2.50  0.01235 *
## LastAuthorFemale1  -0.09449    0.01614   -5.85  4.9e-09 ***
## Year1997           -0.01207    0.03659   -0.33  0.74141
## Year1998            0.00119    0.03695    0.03  0.97441
## Year1999            0.01562    0.03620    0.43  0.66617
## Year2000           -0.01213    0.03520   -0.34  0.73032
## Year2001            0.03997    0.03510    1.14  0.25486
## Year2002            0.10755    0.03556    3.02  0.00249 **
## Year2003            0.01392    0.03666    0.38  0.70421
```

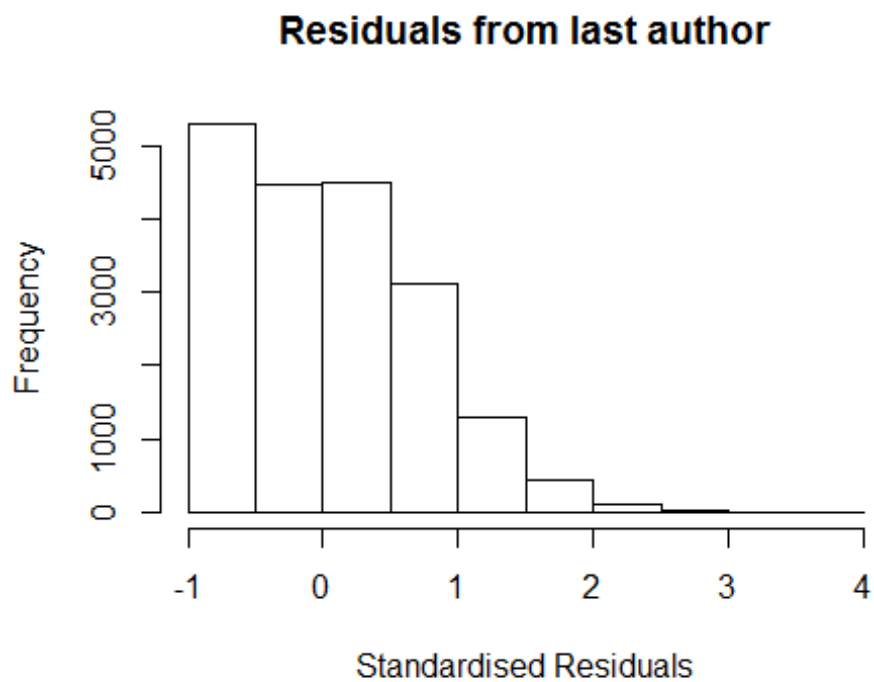
```
## Year2004            0.02510    0.03554    0.71  0.48004
## Year2005            0.07047    0.03552    1.98  0.04726 *
## Year2006            0.09064    0.03598    2.52  0.01177 *
## Year2007            0.04820    0.03361    1.43  0.15155
## Year2008            0.06126    0.03292    1.86  0.06275 .
## Year2009            0.10945    0.03286    3.33  0.00087 ***
## Year2010            0.13114    0.03276    4.00  6.3e-05 ***
## Year2011            0.13765    0.03329    4.13  3.6e-05 ***
## Year2012            0.10485    0.03430    3.06  0.00224 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.763
## Multiple R-squared:  0.0116, Adjusted R-squared:  0.0107
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
##  2 observations c(985,10719)
##   are outliers with |weight| <= 2.7e-06 ( < 5.2e-06);
##  1519 weights are ~= 1. The remaining 17736 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0067  0.8920  0.9400  0.9180  0.9840  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          5.19e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.005  1           1.003
## Year              1.005 16           1.000
```




```
## [1] "List of  31 outliers with residuals above  2.5"
##          ScopusId  NLCS Year OneField Fields residuals
## 271    0030290419 3.426 1996     1706      3     2.587
## 416    0030240869 3.699 1996     1710      2     2.995
## 1254  35048881580 4.306 1997     3309      1     3.574
## 1255  35248867366 3.351 1997     3309      1     2.565
## 2618   0032184399 3.521 1998     1706      3     2.681
## 2629   0032201980 3.365 1998     1706      3     2.525
## 3362   0032339276 3.604 1998     1705      4     2.805
## 3380   0032447095 3.733 1998     1710      3     2.893
## 3738   0032636505 3.564 1999     1706      3     2.710
## 4062   0033449090 3.235 1999     3309      1     2.516
## 4998   0034188649 3.199 2000     1706      3     2.507
## 5041   0034315068 3.331 2000     1706      3     2.504
## 6117   0034559541 3.488 2000     1404      5     2.661
## 6772   0002549585 3.549 2001     1710      2     2.670
## 8419   0036611602 3.834 2002     1706      3     2.888
## 10794  5444237123 3.544 2004     1706      3     2.680
## 12088 29144439194 3.812 2005     1706      3     2.903
## 13526 33947416035 3.963 2006     1706      3     3.034
## 13696 33947399169 3.483 2006     1706      3     2.594
## 14238 33644912045 3.569 2006     1710      2     2.640
## 14270 33645712892 4.593 2006     1706      3     3.664
## 14418 31744440684 4.357 2006     1706      3     3.428
## 14980 64649083745 4.114 2007     1706      3     3.321
## 15826 34250855167 3.383 2007     1500      4     2.631
## 17387 48849113556 3.758 2008     1706      3     2.858
```

```
## 17601 48849096944 3.369 2008     1710      2     2.510
## 21253 77953568642 3.825 2010     3308      3     2.855
## 21687 77950244328 3.504 2010     1706      3     2.534
## 25046 84866279731 3.877 2012     3308      3     2.933
## 25284 84867288643 3.693 2012     1606      4     2.749
## 26463 83055186560 3.802 2012     3308      3     2.858
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -0.9642 -0.5874 -0.0167  0.5232  3.6756
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         0.82798    0.02767   29.93  < 2e-16 ***
## FirstAuthorFemale1 -0.11029    0.01081  -10.20  < 2e-16 ***
## Year1997           -0.01449    0.03661   -0.40   0.6922
## Year1998            0.00103    0.03697    0.03   0.9778
## Year1999            0.01481    0.03618    0.41   0.6822
## Year2000           -0.01332    0.03520   -0.38   0.7051
## Year2001            0.03780    0.03510    1.08   0.2815
## Year2002            0.10754    0.03556    3.02   0.0025 **
## Year2003            0.01419    0.03671    0.39   0.6991
## Year2004            0.02706    0.03554    0.76   0.4464
## Year2005            0.07028    0.03552    1.98   0.0479 *
## Year2006            0.08944    0.03596    2.49   0.0129 *
## Year2007            0.04640    0.03362    1.38   0.1676
## Year2008            0.05965    0.03294    1.81   0.0702 .
## Year2009            0.10629    0.03286    3.23   0.0012 **
## Year2010            0.12924    0.03278    3.94  8.1e-05 ***
## Year2011            0.13623    0.03331    4.09  4.3e-05 ***
## Year2012            0.10418    0.03431    3.04   0.0024 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.764
## Multiple R-squared:  0.00972,    Adjusted R-squared:  0.00885
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
##  observation 10719 is an outlier with |weight| = 0 ( < 5.2e-06);
##  1566 weights are ~= 1. The remaining 17690 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0024  0.8910  0.9400  0.9180  0.9830  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
```

```
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          5.19e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##                   GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.004  1           1.002
## Year             1.004 16           1.000
```

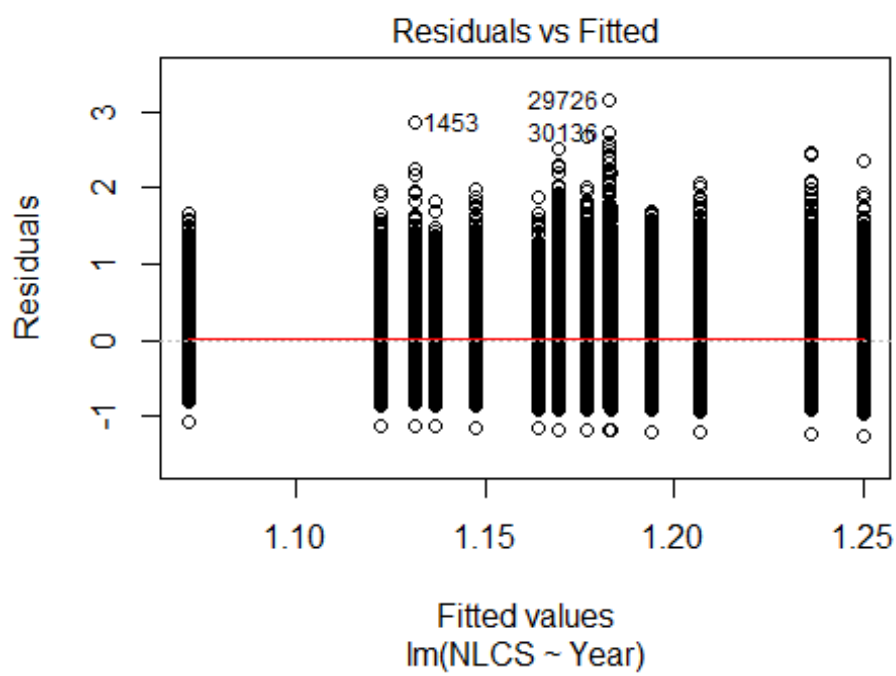



```
## [1] "List of  31 outliers with residuals above  2.5"
##          ScopusId  NLCS Year OneField Fields residuals
## 271    0030290419 3.426 1996     1706      3     2.587
## 416    0030240869 3.699 1996     1710      2     2.995
## 1254  35048881580 4.306 1997     3309      1     3.574
## 1255  35248867366 3.351 1997     3309      1     2.565
## 2618   0032184399 3.521 1998     1706      3     2.681
## 2629   0032201980 3.365 1998     1706      3     2.525
```

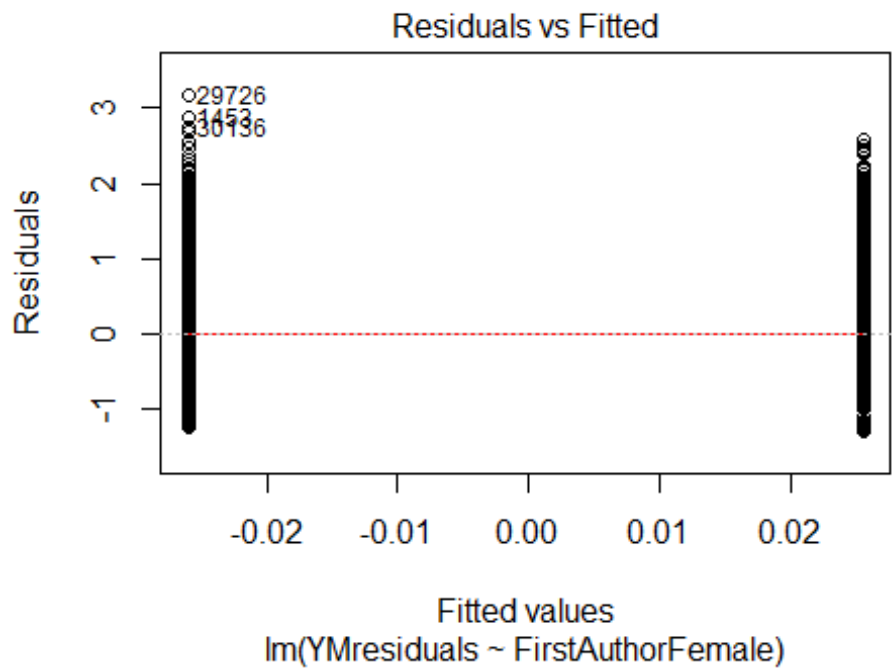
```
## 3362   0032339276 3.604 1998     1705      4     2.805
## 3380   0032447095 3.733 1998     1710      3     2.893
## 3738   0032636505 3.564 1999     1706      3     2.710
## 4062   0033449090 3.235 1999     3309      1     2.516
## 4998   0034188649 3.199 2000     1706      3     2.507
## 5041   0034315068 3.331 2000     1706      3     2.504
## 6117   0034559541 3.488 2000     1404      5     2.661
## 6772   0002549585 3.549 2001     1710      2     2.670
## 8419   0036611602 3.834 2002     1706      3     2.888
## 10794  5444237123 3.544 2004     1706      3     2.680
## 12088 29144439194 3.812 2005     1706      3     2.903
## 13526 33947416035 3.963 2006     1706      3     3.034
## 13696 33947399169 3.483 2006     1706      3     2.594
## 14238 33644912045 3.569 2006     1710      2     2.640
## 14270 33645712892 4.593 2006     1706      3     3.664
## 14418 31744440684 4.357 2006     1706      3     3.428
## 14980 64649083745 4.114 2007     1706      3     3.321
## 15826 34250855167 3.383 2007     1500      4     2.631
## 17387 48849113556 3.758 2008     1706      3     2.858
## 17601 48849096944 3.369 2008     1710      2     2.510
## 21253 77953568642 3.825 2010     3308      3     2.855
## 21687 77950244328 3.504 2010     1706      3     2.534
## 25046 84866279731 3.877 2012     3308      3     2.933
## 25284 84867288643 3.693 2012     1606      4     2.749
## 26463 83055186560 3.802 2012     3308      3     2.858
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -0.9697 -0.5934 -0.0124  0.5222  3.6696
##
## Coefficients:
##                   Estimate Std. Error t value Pr(>|t|)
## (Intercept)        0.83365    0.02769   30.11  < 2e-16 ***
## LastAuthorFemale1 -0.12436    0.01074  -11.58  < 2e-16 ***
## Year1997          -0.01167    0.03656   -0.32   0.7496
## Year1998           0.00128    0.03695    0.03   0.9724
## Year1999           0.01563    0.03621    0.43   0.6660
## Year2000          -0.01223    0.03521   -0.35   0.7283
## Year2001           0.03922    0.03512    1.12   0.2641
## Year2002           0.10624    0.03556    2.99   0.0028 **
## Year2003           0.01336    0.03665    0.36   0.7155
## Year2004           0.02307    0.03555    0.65   0.5164
## Year2005           0.06890    0.03553    1.94   0.0525 .
## Year2006           0.08974    0.03600    2.49   0.0127 *
## Year2007           0.04736    0.03362    1.41   0.1589
```

```
## Year2008           0.06028    0.03292    1.83   0.0671 .
## Year2009           0.10917    0.03287    3.32   0.0009 ***
## Year2010           0.13010    0.03275    3.97  7.1e-05 ***
## Year2011           0.13602    0.03329    4.09  4.4e-05 ***
## Year2012           0.10349    0.03430    3.02   0.0026 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.764
## Multiple R-squared:  0.0112, Adjusted R-squared:  0.0104
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
##  2 observations c(985,10719) are outliers with |weight| = 0 ( < 5.2e-06);
##  1562 weights are ~= 1. The remaining 17693 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0063  0.8920  0.9400  0.9180  0.9830  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          5.19e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Sample size for the above analysis:  19257"
## [1] ""
## [1] ""
## [1] "####################################"
## [1] "Analysis of AJSC 3310"
## [1] "####################################"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 1164 1287 1095 1236 1344 1471 1566 1298 1433 1613 1633 1843 1945 2093 2197
## 2011 2012
## 2438 2621
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  996 1063  950 1039 1127 1164 1367 1123 1225 1361 1377 1591 1653 1765 1877
## 2011 2012
```

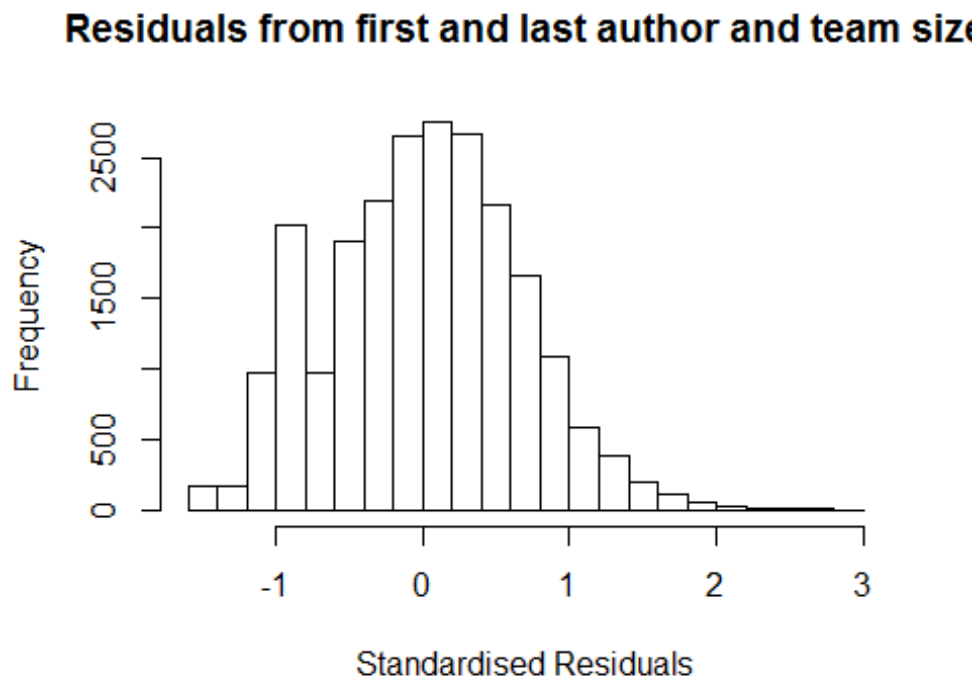
```
## 2039 2226
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  957 1032  908 1004 1073 1110 1307 1066 1162 1297 1323 1527 1575 1654 1746
## 2011 2012
## 1916 2110
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
##  Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 300, df = 16, p-value <2e-16
```




```
##
##  Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 58, df = 1, p-value = 3e-14
```



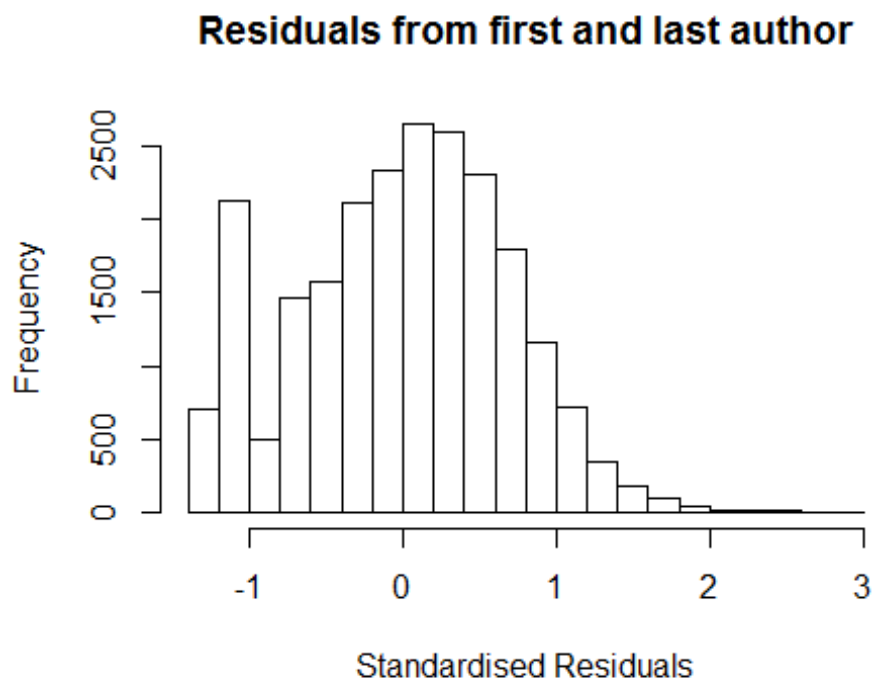

```
## [1] "Female first author team size 2018 geometric mean: 1.72291528158929"
## [1] "Male first author team size 2018 geometric mean: 1.54302395368615"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 510000, p-value = 3e-05
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 1.68143702855826"
## [1] "Male last author team size 2018 geometric mean: 1.59959231472255"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 490000, p-value = 0.04
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.460  1           1.208
## LastAuthorFemale  1.454  1           1.206
## UniqueAuthors     1.033  4           1.004
## Year              1.027 16           1.001
```




```
## [1] "List of  9 outliers with residuals above  2.5"
##          ScopusId  NLCS Year OneField Fields residuals
## 1453   0031381525 3.993 1997     1203      3     2.623
## 8538  21044438693 3.851 2002     1211      2     2.878
## 13070 32944462886 3.596 2005     1211      2     2.554
## 25941 80054769198 3.695 2011     1203      2     2.694
## 27008 79955674091 3.677 2011     1203      6     2.645
## 28379 84867298859 3.592 2012     3310      1     2.623
## 29630 84872871724 3.721 2012     1203      3     2.752
## 30136 84856523526 3.906 2012     1203      3     2.549
## 30438 84858270024 3.720 2012     3207      3     2.782
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
UniqueAuthors +
##     Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##     k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.5711 -0.4571  0.0237  0.4484  2.8781
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)          0.97505    0.02219   43.94   <2e-16 ***
## FirstAuthorFemale1  0.00655    0.01087    0.60   0.5466
```

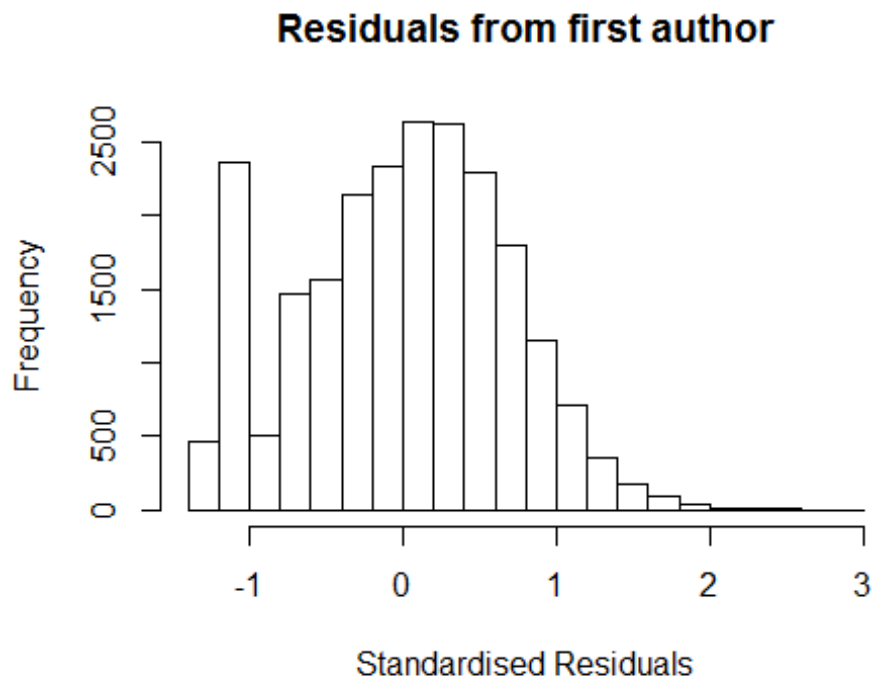
```
## LastAuthorFemale1   0.02405    0.01083    2.22   0.0264 *
## UniqueAuthors2      0.41826    0.01084   38.59   <2e-16 ***
## UniqueAuthors3      0.45123    0.01345   33.56   <2e-16 ***
## UniqueAuthors4      0.47524    0.01825   26.04   <2e-16 ***
## UniqueAuthors5      0.50482    0.02096   24.08   <2e-16 ***
## Year1997           -0.02367    0.03022   -0.78   0.4336
## Year1998            0.02265    0.03048    0.74   0.4575
## Year1999           -0.03411    0.02918   -1.17   0.2425
## Year2000           -0.03702    0.02823   -1.31   0.1897
## Year2001           -0.08879    0.02977   -2.98   0.0029 **
## Year2002           -0.00219    0.02965   -0.07   0.9410
## Year2003            0.01168    0.02952    0.40   0.6923
## Year2004            0.02794    0.03053    0.92   0.3601
## Year2005            0.06720    0.02794    2.41   0.0162 *
## Year2006            0.01109    0.02767    0.40   0.6887
## Year2007           -0.00178    0.02806   -0.06   0.9494
## Year2008           -0.02016    0.02688   -0.75   0.4532
## Year2009           -0.02300    0.02663   -0.86   0.3878
## Year2010           -0.05999    0.02632   -2.28   0.0227 *
## Year2011            0.02626    0.02691    0.98   0.3292
## Year2012           -0.03668    0.02817   -1.30   0.1929
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.667
## Multiple R-squared:  0.101,  Adjusted R-squared:   0.1
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  1904 weights are ~= 1. The remaining 20863 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0228  0.8480  0.9500  0.9100  0.9860  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          4.39e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
```

```
##                     GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.560  1           1.249
## LastAuthorFemale  1.562  1           1.250
## Year              1.005 16           1.000
```



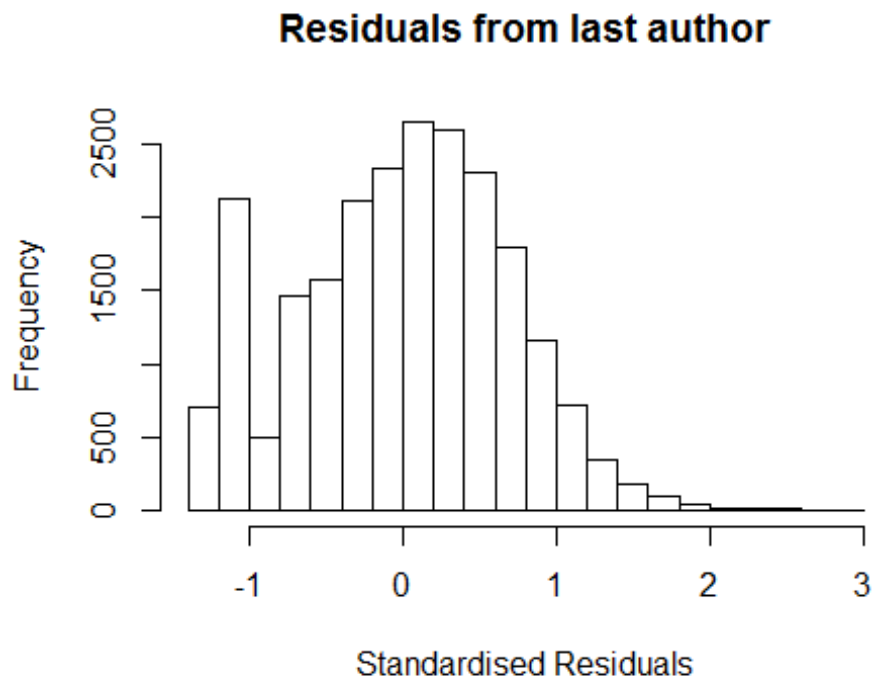

```
## [1] "List of  8 outliers with residuals above  2.5"
##          ScopusId  NLCS Year OneField Fields residuals
## 1453   0031381525 3.993 1997     1203      3     2.903
## 8538  21044438693 3.851 2002     1211      2     2.717
## 18461 60950608143 3.692 2008     1203      3     2.540
## 25941 80054769198 3.695 2011     1203      2     2.507
## 29506 84857205310 3.786 2012     1203      3     2.596
## 29630 84872871724 3.721 2012     1203      3     2.531
## 30136 84856523526 3.906 2012     1203      3     2.789
## 30438 84858270024 3.720 2012     3207      3     2.603
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.2806 -0.4824  0.0445  0.4833  2.9026
##
## Coefficients:
##                      Estimate Std. Error t value Pr(>|t|)
```

```
## (Intercept)         1.128817   0.022979   49.12   <2e-16 ***
## FirstAuthorFemale1  0.044559   0.011984    3.72   0.0002 ***
## LastAuthorFemale1   0.028998   0.011966    2.42   0.0154 *
## Year1997           -0.038380   0.031726   -1.21   0.2264
## Year1998            0.017498   0.031878    0.55   0.5831
## Year1999           -0.048524   0.030569   -1.59   0.1124
## Year2000           -0.030354   0.029511   -1.03   0.3037
## Year2001           -0.097913   0.030959   -3.16   0.0016 **
## Year2002            0.005580   0.031246    0.18   0.8583
## Year2003            0.002381   0.031282    0.08   0.9393
## Year2004            0.032503   0.032176    1.01   0.3124
## Year2005            0.078184   0.029307    2.67   0.0076 **
## Year2006            0.022009   0.029122    0.76   0.4498
## Year2007            0.001054   0.029554    0.04   0.9715
## Year2008           -0.006197   0.028349   -0.22   0.8270
## Year2009           -0.000955   0.027930   -0.03   0.9727
## Year2010           -0.035232   0.027730   -1.27   0.2039
## Year2011            0.059147   0.028391    2.08   0.0372 *
## Year2012           -0.012119   0.030193   -0.40   0.6882
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.702
## Multiple R-squared:  0.00545,    Adjusted R-squared:  0.00467
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  1886 weights are ~= 1. The remaining 20881 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0492  0.8690  0.9480  0.9110  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          4.39e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.003  1           1.001
## Year              1.003 16           1.000
```




```
## [1] "List of  8 outliers with residuals above  2.5"
##          ScopusId  NLCS Year OneField Fields residuals
## 1453   0031381525 3.993 1997     1203      3     2.903
## 8538  21044438693 3.851 2002     1211      2     2.717
## 18461 60950608143 3.692 2008     1203      3     2.540
## 25941 80054769198 3.695 2011     1203      2     2.507
## 29506 84857205310 3.786 2012     1203      3     2.596
## 29630 84872871724 3.721 2012     1203      3     2.531
## 30136 84856523526 3.906 2012     1203      3     2.789
## 30438 84858270024 3.720 2012     3207      3     2.603
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.2753 -0.4861  0.0437  0.4837  2.8982
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.13229    0.02291   49.43  < 2e-16 ***
## FirstAuthorFemale1  0.06350    0.00963    6.59  4.4e-11 ***
## Year1997           -0.03748    0.03173   -1.18   0.2375
## Year1998            0.01914    0.03185    0.60   0.5479
## Year1999           -0.04748    0.03057   -1.55   0.1204
```

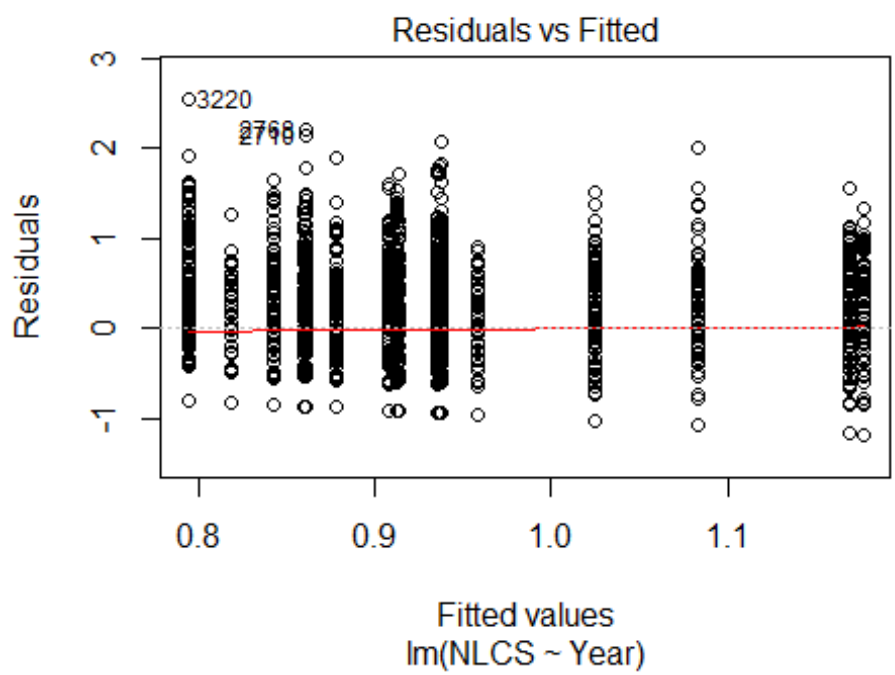
```
## Year2000           -0.02938    0.02951   -1.00   0.3194
## Year2001           -0.09638    0.03094   -3.11   0.0018 **
## Year2002            0.00698    0.03125    0.22   0.8233
## Year2003            0.00367    0.03127    0.12   0.9066
## Year2004            0.03413    0.03216    1.06   0.2886
## Year2005            0.07950    0.02929    2.71   0.0067 **
## Year2006            0.02299    0.02911    0.79   0.4298
## Year2007            0.00230    0.02955    0.08   0.9381
## Year2008           -0.00479    0.02834   -0.17   0.8658
## Year2009            0.00106    0.02791    0.04   0.9696
## Year2010           -0.03378    0.02772   -1.22   0.2230
## Year2011            0.06084    0.02838    2.14   0.0321 *
## Year2012           -0.01086    0.03018   -0.36   0.7189
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.702
## Multiple R-squared:  0.0052, Adjusted R-squared:  0.00445
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  1884 weights are ~= 1. The remaining 20883 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0503  0.8690  0.9480  0.9110  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          4.39e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##                   GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.004  1           1.002
## Year             1.004 16           1.000
```



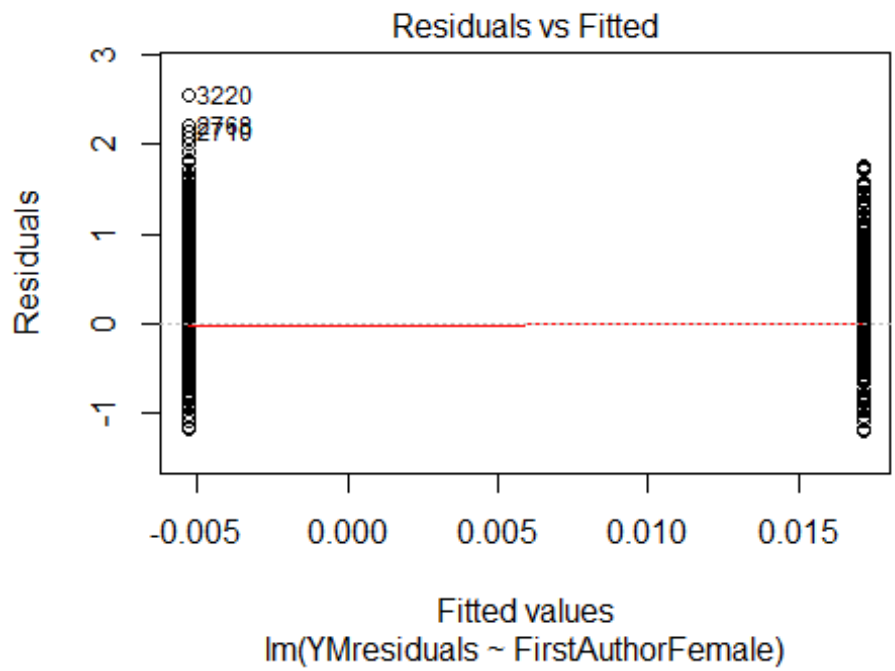

```
## [1] "List of  8 outliers with residuals above  2.5"
##          ScopusId  NLCS Year OneField Fields residuals
## 1453   0031381525 3.993 1997     1203      3     2.903
## 8538  21044438693 3.851 2002     1211      2     2.717
## 18461 60950608143 3.692 2008     1203      3     2.540
## 25941 80054769198 3.695 2011     1203      2     2.507
## 29506 84857205310 3.786 2012     1203      3     2.596
## 29630 84872871724 3.721 2012     1203      3     2.531
## 30136 84856523526 3.906 2012     1203      3     2.789
## 30438 84858270024 3.720 2012     3207      3     2.603
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.2732 -0.4832  0.0445  0.4843  2.8950
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)        1.136794   0.022876   49.69  < 2e-16 ***
## LastAuthorFemale1  0.058105   0.009622    6.04  1.6e-09 ***
## Year1997          -0.038778   0.031719   -1.22   0.2215
## Year1998           0.017633   0.031893    0.55   0.5804
## Year1999          -0.049078   0.030617   -1.60   0.1090
```

```
## Year2000          -0.029746   0.029527   -1.01   0.3138
## Year2001          -0.097902   0.030975   -3.16   0.0016 **
## Year2002           0.005380   0.031261    0.17   0.8634
## Year2003           0.002626   0.031318    0.08   0.9332
## Year2004           0.032882   0.032212    1.02   0.3074
## Year2005           0.078257   0.029338    2.67   0.0076 **
## Year2006           0.022906   0.029127    0.79   0.4316
## Year2007           0.001429   0.029589    0.05   0.9615
## Year2008          -0.006127   0.028388   -0.22   0.8291
## Year2009          -0.000949   0.027955   -0.03   0.9729
## Year2010          -0.034563   0.027754   -1.25   0.2130
## Year2011           0.059266   0.028426    2.08   0.0371 *
## Year2012          -0.011365   0.030231   -0.38   0.7070
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.703
## Multiple R-squared:  0.00487,    Adjusted R-squared:  0.00413
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  1881 weights are ~= 1. The remaining 20886 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0512  0.8690  0.9480  0.9110  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          4.39e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Sample size for the above analysis:  22767"
## [1] ""
## [1] ""
## [1] "####################################"
## [1] "Analysis of AJSC 3311"
## [1] "####################################"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
```

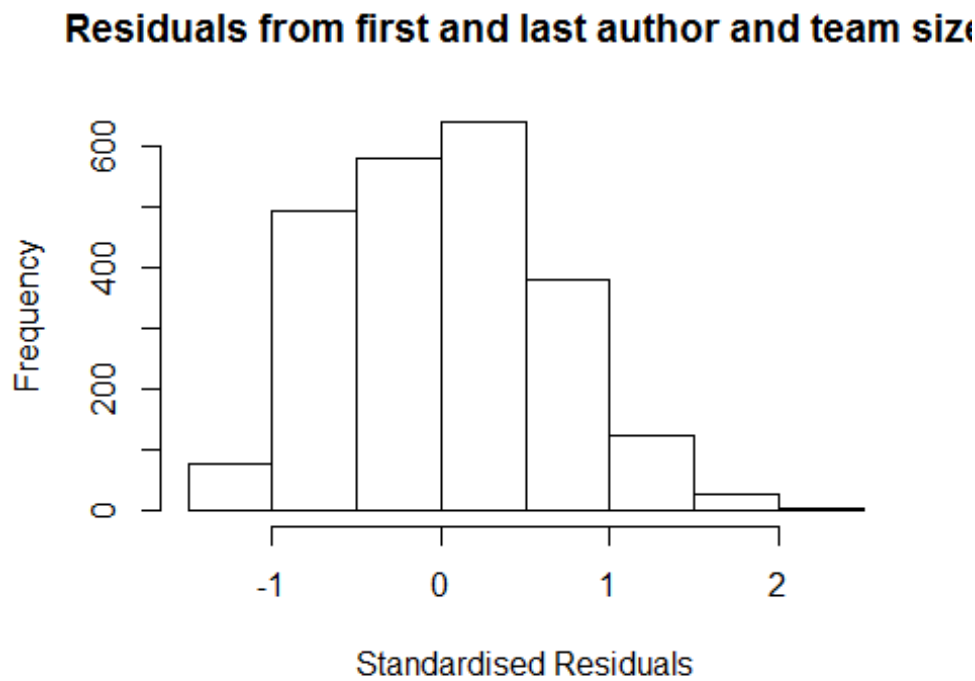
```
##   74   63   77   84  121  141  129  132  106  144  165  206  281  298  305
## 2011 2012
##  327  366
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##   58   53   65   55   91   87  109  113   92  121  139  183  220  250  248
## 2011 2012
##  270  316
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##   55   52   62   54   89   87   98  109   85  114  130  173  205  230  233
## 2011 2012
##  258  293
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
##  Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 30, df = 16, p-value = 0.02
```




```
##
##  Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 0.00093, df = 1, p-value = 1
```

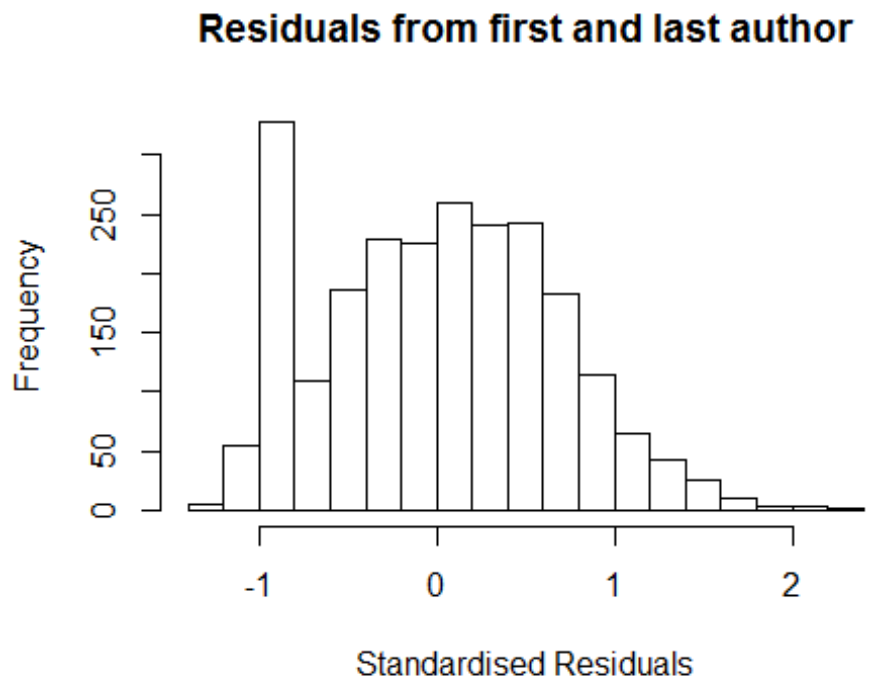



```
## [1] "Female first author team size 2018 geometric mean: 2.79575070438836"
## [1] "Male first author team size 2018 geometric mean: 2.40273053323676"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 15000, p-value = 0.04
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 2.6539543022873"
## [1] "Male last author team size 2018 geometric mean: 2.47300856085043"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 14000, p-value = 0.3
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.600  1           1.265
## LastAuthorFemale  1.573  1           1.254
## UniqueAuthors     1.207  4           1.024
## Year              1.257 16           1.007
```



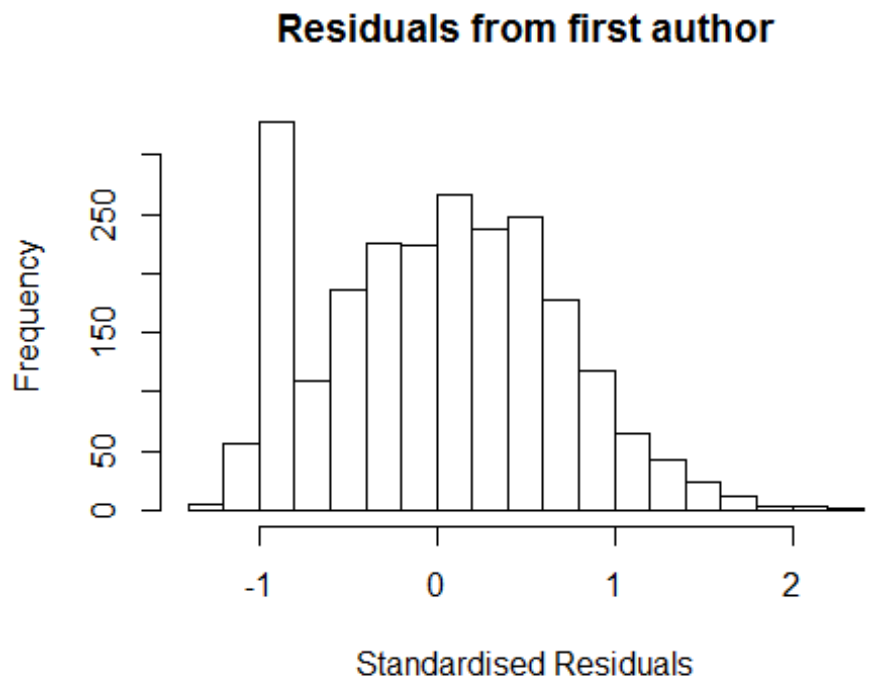

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
UniqueAuthors +
##     Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##     k.max = 1000))
##  \--> method = "MM"
## Residuals:
##    Min     1Q Median     3Q    Max
## -1.430 -0.480  0.009  0.465  2.275
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)          0.80196    0.09027    8.88  < 2e-16 ***
## FirstAuthorFemale1   0.02068    0.04117    0.50    0.615
## LastAuthorFemale1    0.01682    0.04019    0.42    0.676
## UniqueAuthors2       0.24592    0.03824    6.43  1.5e-10 ***
## UniqueAuthors3       0.24935    0.04724    5.28  1.4e-07 ***
## UniqueAuthors4       0.31187    0.05431    5.74  1.1e-08 ***
## UniqueAuthors5       0.23944    0.04700    5.09  3.8e-07 ***
## Year1997            -0.09392    0.12268   -0.77    0.444
## Year1998            -0.05063    0.12911   -0.39    0.695
## Year1999             0.05294    0.11798    0.45    0.654
```

```
## Year2000           -0.02437    0.11941   -0.20    0.838
## Year2001            0.29554    0.12331    2.40    0.017 *
## Year2002            0.19511    0.12167    1.60    0.109
## Year2003            0.11218    0.11284    0.99    0.320
## Year2004            0.12654    0.12585    1.01    0.315
## Year2005           -0.08795    0.11091   -0.79    0.428
## Year2006           -0.03288    0.10938   -0.30    0.764
## Year2007           -0.05312    0.10296   -0.52    0.606
## Year2008           -0.05937    0.09813   -0.60    0.545
## Year2009           -0.04327    0.09970   -0.43    0.664
## Year2010            0.00778    0.09948    0.08    0.938
## Year2011           -0.08485    0.09941   -0.85    0.393
## Year2012           -0.18677    0.09859   -1.89    0.058 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.686
## Multiple R-squared:  0.061,  Adjusted R-squared:  0.052
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  183 weights are ~= 1. The remaining 2144 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.248   0.888   0.948   0.918   0.984   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          4.30e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.619  1           1.272
## LastAuthorFemale  1.592  1           1.262
## Year              1.049 16           1.001
```



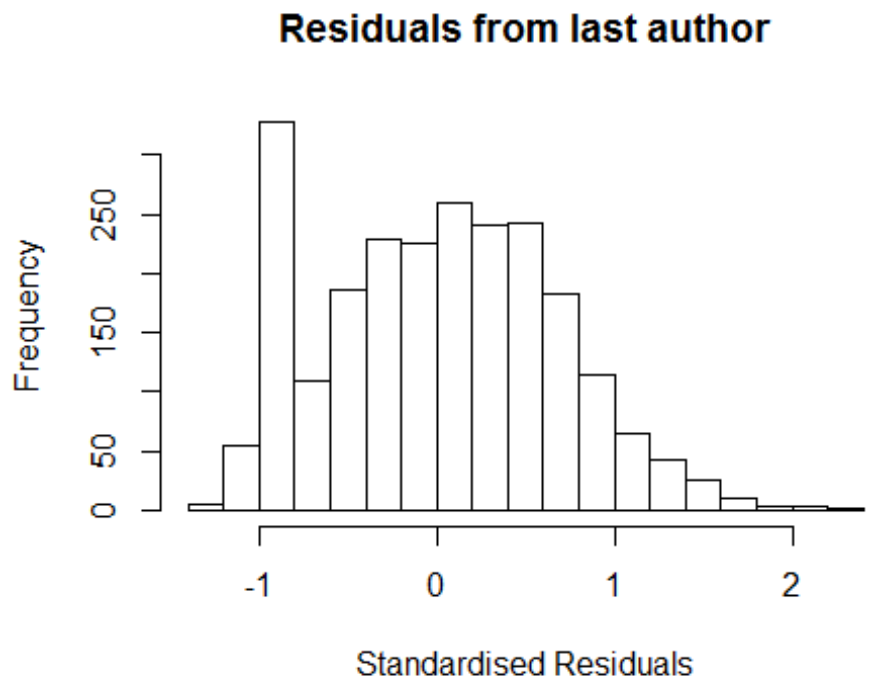

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.2286 -0.5155  0.0211  0.4899  2.2480
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)          0.87937    0.08702   10.11   <2e-16 ***
## FirstAuthorFemale1   0.02328    0.04231    0.55    0.582
## LastAuthorFemale1    0.02149    0.04125    0.52    0.602
## Year1997            -0.06437    0.12051   -0.53    0.593
## Year1998            -0.05575    0.12518   -0.45    0.656
## Year1999             0.08371    0.11495    0.73    0.467
## Year2000            -0.00589    0.11679   -0.05    0.960
## Year2001             0.30446    0.12026    2.53    0.011 *
## Year2002             0.24200    0.11892    2.04    0.042 *
## Year2003             0.15307    0.11040    1.39    0.166
## Year2004             0.19387    0.12613    1.54    0.124
## Year2005            -0.07446    0.11084   -0.67    0.502
```

```
## Year2006            0.00755    0.10849    0.07    0.945
## Year2007            0.00910    0.10047    0.09    0.928
## Year2008           -0.02132    0.09553   -0.22    0.823
## Year2009           -0.02101    0.09741   -0.22    0.829
## Year2010            0.03485    0.09685    0.36    0.719
## Year2011           -0.05933    0.09750   -0.61    0.543
## Year2012           -0.15684    0.09625   -1.63    0.103
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.711
## Multiple R-squared:  0.027,  Adjusted R-squared:  0.0194
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  198 weights are ~= 1. The remaining 2129 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.296   0.876   0.948   0.921   0.983   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          4.30e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.036  1           1.018
## Year              1.036 16           1.001
```




```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.2207 -0.5139  0.0254  0.4890  2.2454
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)          0.88003    0.08682   10.14   <2e-16 ***
## FirstAuthorFemale1   0.03655    0.03379    1.08    0.279
## Year1997            -0.06246    0.12042   -0.52    0.604
## Year1998            -0.05467    0.12510   -0.44    0.662
## Year1999             0.08437    0.11496    0.73    0.463
## Year2000            -0.00430    0.11656   -0.04    0.971
## Year2001             0.30409    0.12015    2.53    0.011 *
## Year2002             0.24233    0.11886    2.04    0.042 *
## Year2003             0.15400    0.11027    1.40    0.163
## Year2004             0.19515    0.12608    1.55    0.122
## Year2005            -0.07313    0.11076   -0.66    0.509
## Year2006             0.00805    0.10831    0.07    0.941
```

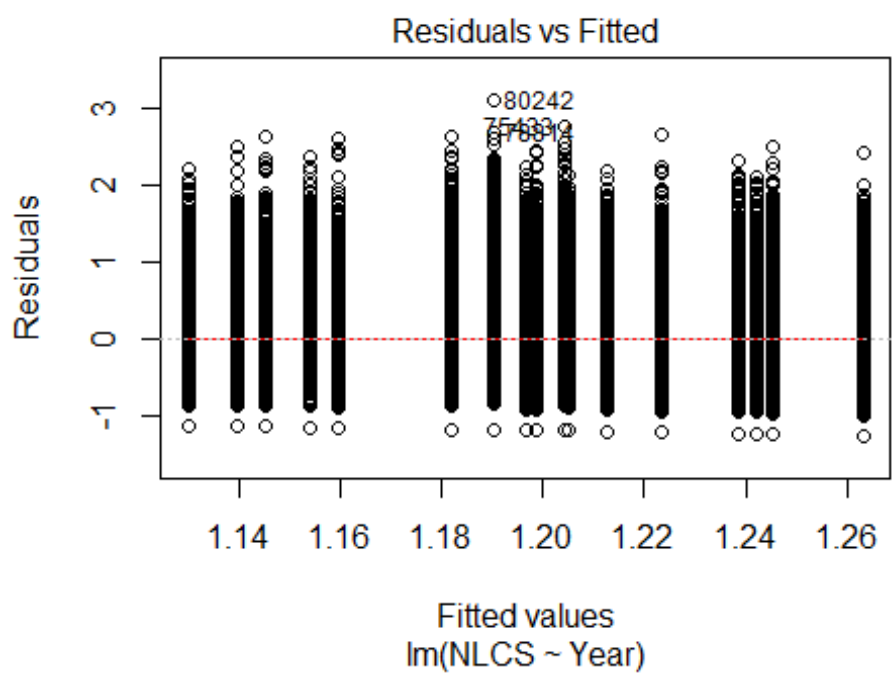
```
## Year2007            0.01131    0.10029    0.11    0.910
## Year2008           -0.01985    0.09536   -0.21    0.835
## Year2009           -0.01990    0.09727   -0.20    0.838
## Year2010            0.03669    0.09670    0.38    0.704
## Year2011           -0.05748    0.09737   -0.59    0.555
## Year2012           -0.15500    0.09610   -1.61    0.107
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.711
## Multiple R-squared:  0.0269, Adjusted R-squared:  0.0197
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  204 weights are ~= 1. The remaining 2123 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.297   0.876   0.948   0.921   0.983   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          4.30e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##                   GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.019  1           1.010
## Year             1.019 16           1.001
```



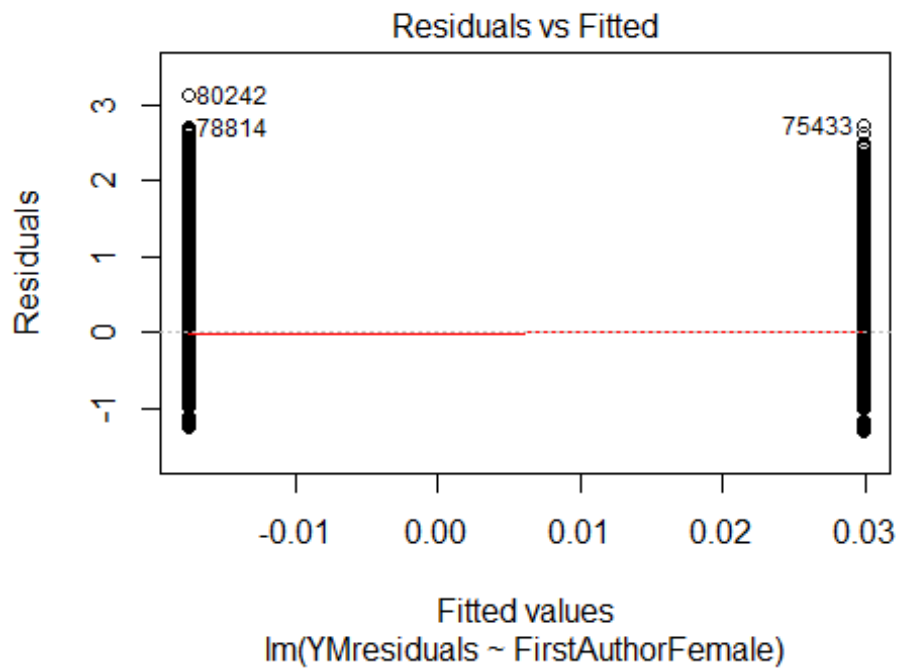

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.2230 -0.5119  0.0252  0.4895  2.2460
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         0.88185    0.08717   10.12   <2e-16 ***
## LastAuthorFemale1   0.03564    0.03294    1.08    0.279
## Year1997           -0.06699    0.12045   -0.56    0.578
## Year1998           -0.05692    0.12534   -0.45    0.650
## Year1999            0.08496    0.11496    0.74    0.460
## Year2000           -0.00692    0.11703   -0.06    0.953
## Year2001            0.30553    0.12022    2.54    0.011 *
## Year2002            0.24114    0.11886    2.03    0.043 *
## Year2003            0.15168    0.11047    1.37    0.170
## Year2004            0.19307    0.12621    1.53    0.126
## Year2005           -0.07298    0.11069   -0.66    0.510
## Year2006            0.00773    0.10868    0.07    0.943
```

```
## Year2007           0.00795    0.10063    0.08    0.937
## Year2008          -0.02162    0.09568   -0.23    0.821
## Year2009          -0.02077    0.09753   -0.21    0.831
## Year2010           0.03435    0.09698    0.35    0.723
## Year2011          -0.05987    0.09761   -0.61    0.540
## Year2012          -0.15686    0.09637   -1.63    0.104
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.713
## Multiple R-squared:  0.0268, Adjusted R-squared:  0.0197
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  196 weights are ~= 1. The remaining 2131 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.300   0.878   0.949   0.921   0.984   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          4.30e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Sample size for the above analysis:  2327"
## [1] ""
## [1] ""
## [1] "####################################"
## [1] "Analysis of AJSC 3312"
## [1] "####################################"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 3288 2929 3164 3182 3890 4054 3804 3299 3556 3702 4496 4912 5497 6556 7002
## 2011 2012
## 7092 6890
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 2843 2548 2696 2725 3204 3111 3231 2883 3109 3219 3899 4231 4745 5657 6029
## 2011 2012
```

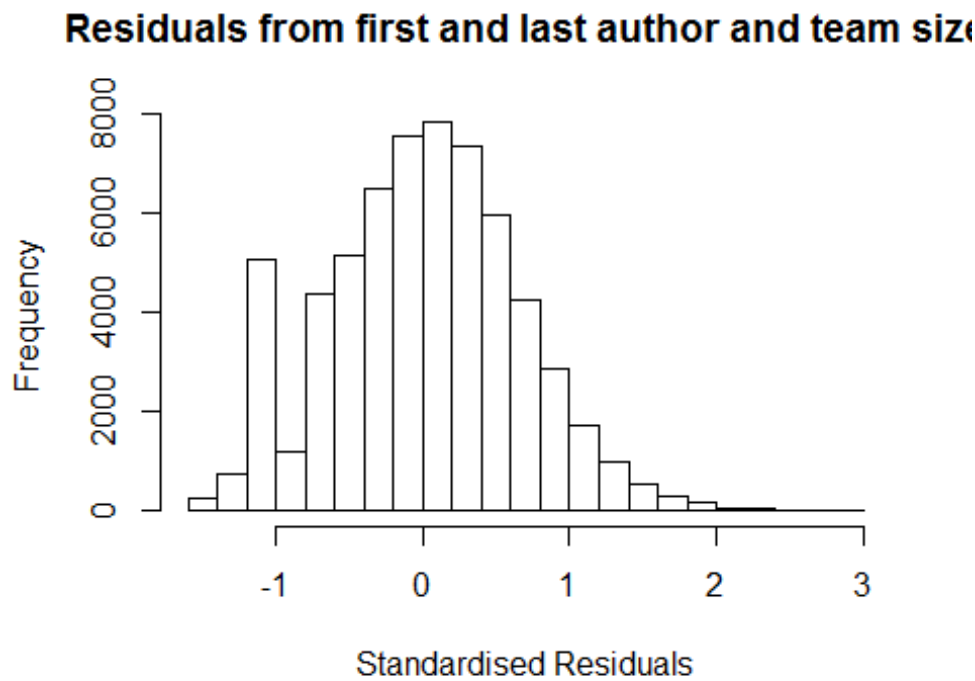
```
## 6064 5956
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 2731 2452 2583 2618 3053 2957 3087 2738 2947 3050 3685 3981 4478 5306 5686
## 2011 2012
## 5690 5593
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
##  Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 250, df = 16, p-value <2e-16
```




```
##
##  Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 200, df = 1, p-value <2e-16
```



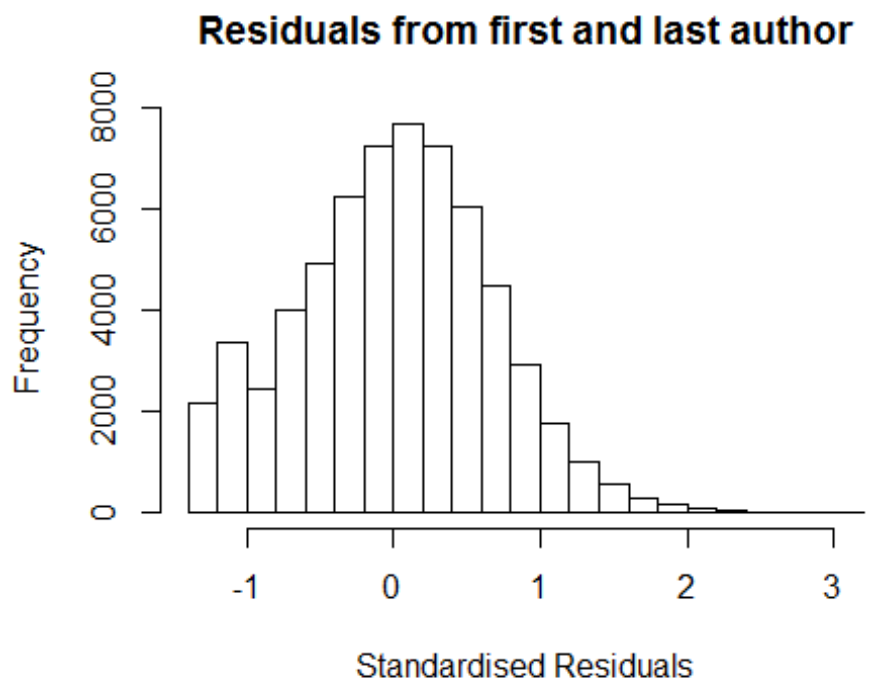

```
## [1] "Female first author team size 2018 geometric mean: 1.85668046935535"
## [1] "Male first author team size 2018 geometric mean: 1.56895438249028"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 4800000, p-value <2e-16
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 1.81749859029173"
## [1] "Male last author team size 2018 geometric mean: 1.60360161585363"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 4600000, p-value = 1e-13
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##                     GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.614  1           1.270
## LastAuthorFemale  1.603  1           1.266
## UniqueAuthors     1.046  4           1.006
## Year              1.035 16           1.001
```




```
## [1] "List of  11 outliers with residuals above  2.5"
##          ScopusId  NLCS Year OneField Fields residuals
## 1119   0030424744 3.615 1996     3312      1     2.538
## 1626   0030305209 3.653 1996     3312      2     2.576
## 8683   0032018622 3.646 1998     3312      2     2.596
## 16437  0034336793 3.778 2000     3312      2     2.750
## 16515  0034345564 3.547 2000     1202      2     2.519
## 74606 79955674091 3.677 2011     1203      6     2.610
## 75433 79951887862 3.965 2011     3312      2     2.610
## 77527 84868280635 3.541 2012     1202      3     2.531
## 78814 84866279731 3.877 2012     3308      3     2.867
## 80242 84864517335 4.300 2012     3312      1     2.953
## 83330 84856137889 3.529 2012     3312      2     2.519
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
UniqueAuthors +
##     Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##     k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.5290 -0.4344  0.0159  0.4355  2.9533
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
```

```
## (Intercept)         1.07650    0.01342   80.22  < 2e-16 ***
## FirstAuthorFemale1  0.02092    0.00686    3.05  0.00231 **
## LastAuthorFemale1   0.01527    0.00687    2.22  0.02624 *
## UniqueAuthors2      0.28751    0.00650   44.25  < 2e-16 ***
## UniqueAuthors3      0.33632    0.00852   39.45  < 2e-16 ***
## UniqueAuthors4      0.35875    0.01208   29.69  < 2e-16 ***
## UniqueAuthors5      0.37046    0.01349   27.47  < 2e-16 ***
## Year1997           -0.03903    0.01893   -2.06  0.03921 *
## Year1998           -0.06307    0.01846   -3.42  0.00063 ***
## Year1999           -0.07491    0.01788   -4.19  2.8e-05 ***
## Year2000           -0.04856    0.01710   -2.84  0.00452 **
## Year2001           -0.07042    0.01758   -4.01  6.2e-05 ***
## Year2002            0.04583    0.01761    2.60  0.00926 **
## Year2003            0.02514    0.01878    1.34  0.18063
## Year2004           -0.02951    0.01823   -1.62  0.10554
## Year2005           -0.01411    0.01758   -0.80  0.42210
## Year2006           -0.00324    0.01656   -0.20  0.84499
## Year2007            0.01238    0.01660    0.75  0.45595
## Year2008            0.00176    0.01645    0.11  0.91497
## Year2009           -0.02536    0.01606   -1.58  0.11437
## Year2010           -0.06135    0.01589   -3.86  0.00011 ***
## Year2011           -0.04524    0.01617   -2.80  0.00514 **
## Year2012           -0.06615    0.01648   -4.01  6.0e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.645
## Multiple R-squared:  0.0582, Adjusted R-squared:  0.0579
## Convergence in 10 IRWLS iterations
##
## Robustness weights:
##  5276 weights are ~= 1. The remaining 57359 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0022  0.8710  0.9500  0.9080  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.60e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                    psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
```

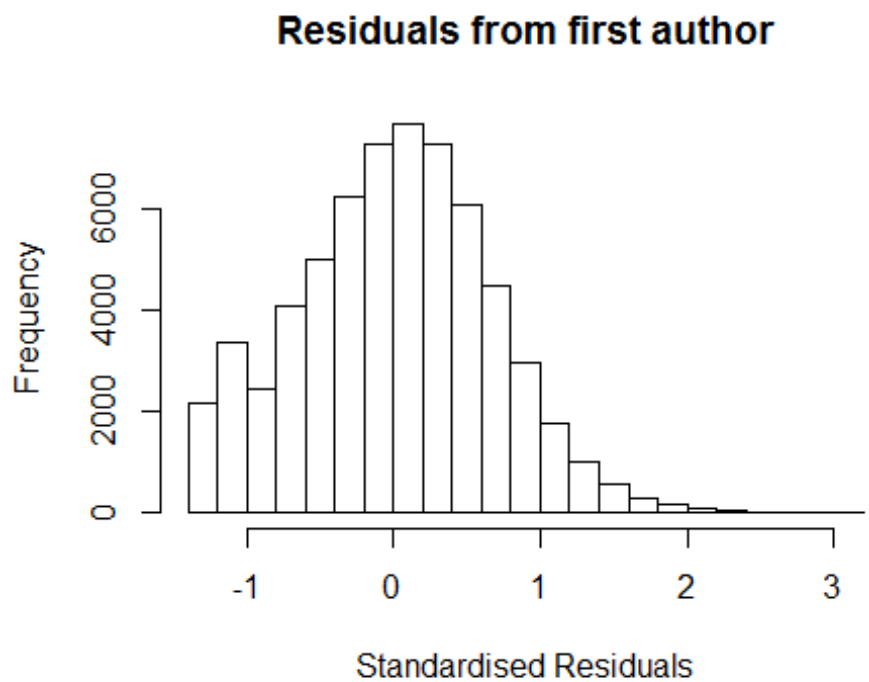
```
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##                     GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.648  1           1.284
## LastAuthorFemale  1.644  1           1.282
## Year              1.009 16           1.000
```




```
## [1] "List of  11 outliers with residuals above  2.5"
##          ScopusId  NLCS Year OneField Fields residuals
## 16437  0034336793 3.778 2000     3312      2     2.653
## 16504  0034344702 3.626 2000     1202      2     2.501
## 38752 31044445688 3.879 2006     3312      2     2.686
## 46043 34249885738 3.757 2007     3312      2     2.514
## 65124 77953568642 3.825 2010     3308      3     2.688
## 65379 77957276053 3.638 2010     3312      2     2.501
## 75433 79951887862 3.965 2011     3312      2     2.733
## 78814 84866279731 3.877 2012     3308      3     2.735
## 79687 84873204638 3.715 2012     1202      3     2.530
## 80242 84864517335 4.300 2012     3312      1     3.158
## 83957 83055186560 3.802 2012     3308      3     2.660
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
```

```
##    Min     1Q Median     3Q    Max
## -1.307 -0.448  0.024  0.449  3.158
##
## Coefficients:
##                      Estimate Std. Error t value Pr(>|t|)
## (Intercept)          1.159450   0.013787   84.10  < 2e-16 ***
## FirstAuthorFemale1   0.043275   0.007135    6.07  1.3e-09 ***
## LastAuthorFemale1    0.029408   0.007156    4.11  4.0e-05 ***
## Year1997            -0.041100   0.019439   -2.11  0.03449 *
## Year1998            -0.060543   0.018912   -3.20  0.00137 **
## Year1999            -0.069008   0.018461   -3.74  0.00019 ***
## Year2000            -0.034057   0.017660   -1.93  0.05380 .
## Year2001            -0.055488   0.018102   -3.07  0.00218 **
## Year2002             0.075319   0.018194    4.14  3.5e-05 ***
## Year2003             0.045557   0.019286    2.36  0.01817 *
## Year2004             0.001896   0.018781    0.10  0.91959
## Year2005             0.017325   0.018207    0.95  0.34131
## Year2006             0.033831   0.017122    1.98  0.04817 *
## Year2007             0.053960   0.017163    3.14  0.00167 **
## Year2008             0.045711   0.017012    2.69  0.00721 **
## Year2009             0.013838   0.016626    0.83  0.40525
## Year2010            -0.022493   0.016445   -1.37  0.17139
## Year2011            -0.000188   0.016697   -0.01  0.99100
## Year2012            -0.017574   0.017057   -1.03  0.30287
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.663
## Multiple R-squared:  0.00575,    Adjusted R-squared:  0.00547
## Convergence in 10 IRWLS iterations
##
## Robustness weights:
##  observation 59513 is an outlier with |weight| = 0 ( < 1.6e-06);
##  5321 weights are ~= 1. The remaining 57313 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0506  0.8690  0.9500  0.9080  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.60e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                    psi           subsampling                   cov
##             "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
```

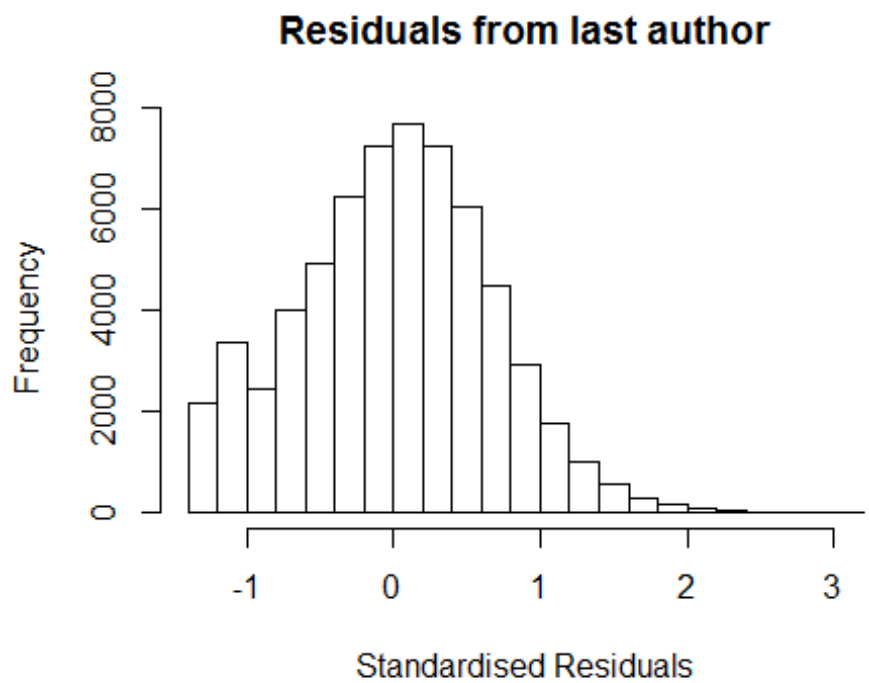
```
##                    "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##                     GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.008  1           1.004
## Year              1.008 16           1.000
```




```
## [1] "List of  11 outliers with residuals above  2.5"
##          ScopusId  NLCS Year OneField Fields residuals
## 16437  0034336793 3.778 2000     3312      2     2.653
## 16504  0034344702 3.626 2000     1202      2     2.501
## 38752 31044445688 3.879 2006     3312      2     2.686
## 46043 34249885738 3.757 2007     3312      2     2.514
## 65124 77953568642 3.825 2010     3308      3     2.688
## 65379 77957276053 3.638 2010     3312      2     2.501
## 75433 79951887862 3.965 2011     3312      2     2.733
## 78814 84866279731 3.877 2012     3308      3     2.735
## 79687 84873204638 3.715 2012     1202      3     2.530
## 80242 84864517335 4.300 2012     3312      1     3.158
## 83957 83055186560 3.802 2012     3308      3     2.660
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
```

```
##     Min      1Q  Median      3Q     Max
## -1.3011 -0.4478  0.0247  0.4496  3.1541
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.162753   0.013769   84.45  < 2e-16 ***
## FirstAuthorFemale1  0.062597   0.005580   11.22  < 2e-16 ***
## Year1997           -0.041135   0.019453   -2.11  0.03447 *
## Year1998           -0.060538   0.018918   -3.20  0.00137 **
## Year1999           -0.068415   0.018476   -3.70  0.00021 ***
## Year2000           -0.033903   0.017668   -1.92  0.05501 .
## Year2001           -0.055040   0.018114   -3.04  0.00238 **
## Year2002            0.075709   0.018199    4.16  3.2e-05 ***
## Year2003            0.045715   0.019294    2.37  0.01782 *
## Year2004            0.001846   0.018790    0.10  0.92174
## Year2005            0.017047   0.018214    0.94  0.34932
## Year2006            0.034449   0.017131    2.01  0.04433 *
## Year2007            0.054502   0.017170    3.17  0.00150 **
## Year2008            0.045995   0.017021    2.70  0.00689 **
## Year2009            0.014329   0.016633    0.86  0.38898
## Year2010           -0.021943   0.016458   -1.33  0.18243
## Year2011            0.000541   0.016707    0.03  0.97415
## Year2012           -0.016902   0.017068   -0.99  0.32203
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.663
## Multiple R-squared:  0.00549,    Adjusted R-squared:  0.00522
## Convergence in 10 IRWLS iterations
##
## Robustness weights:
##  observation 59513 is an outlier with |weight| = 0 ( < 1.6e-06);
##  5285 weights are ~= 1. The remaining 57349 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0498  0.8690  0.9500  0.9080  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.60e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
```

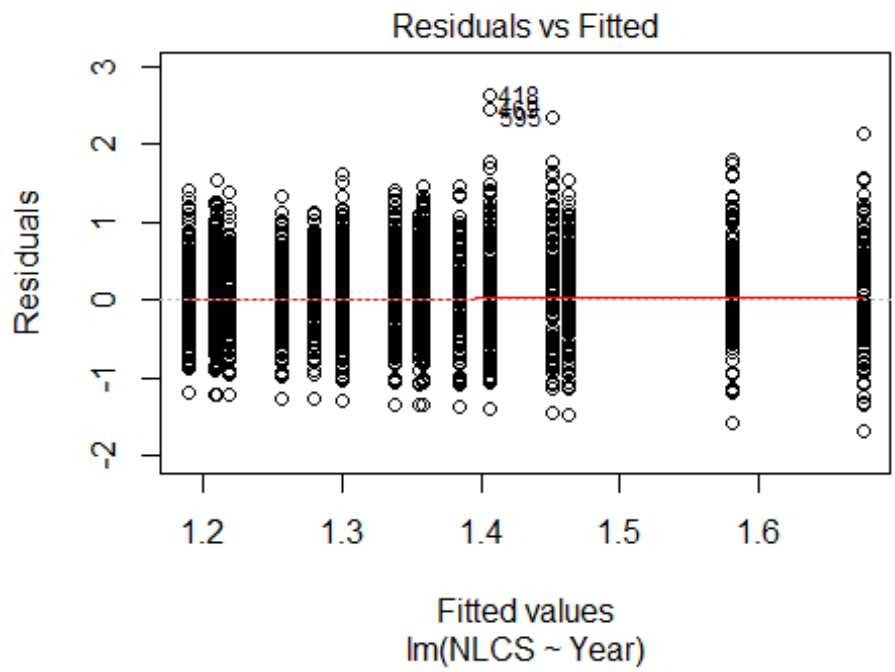
```
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.005  1           1.003
## Year             1.005 16           1.000
```



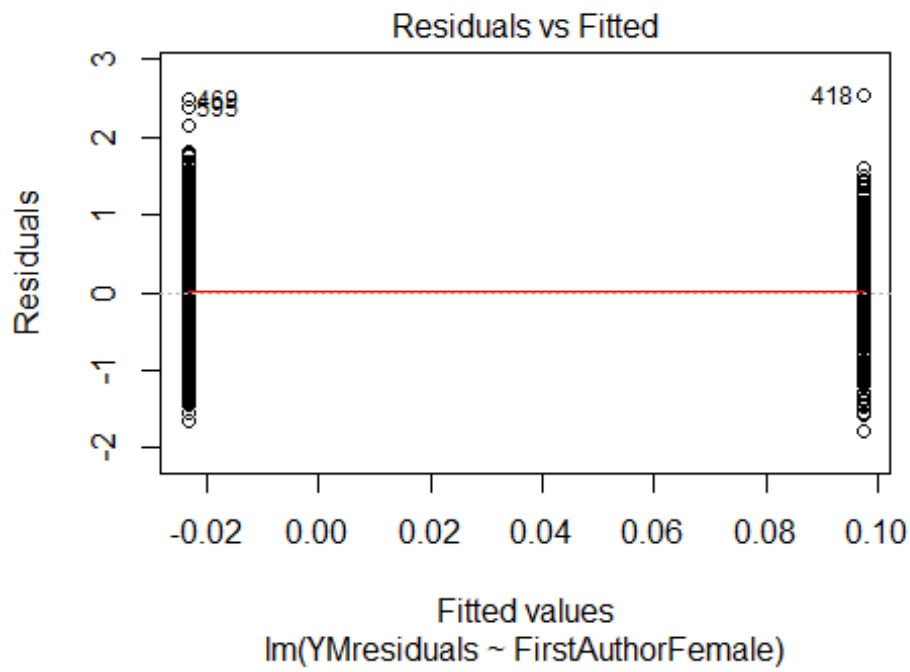

```
## [1] "List of  11 outliers with residuals above  2.5"
##          ScopusId  NLCS Year OneField Fields residuals
## 16437  0034336793 3.778 2000     3312      2     2.653
## 16504  0034344702 3.626 2000     1202      2     2.501
## 38752 31044445688 3.879 2006     3312      2     2.686
## 46043 34249885738 3.757 2007     3312      2     2.514
## 65124 77953568642 3.825 2010     3308      3     2.688
## 65379 77957276053 3.638 2010     3312      2     2.501
## 75433 79951887862 3.965 2011     3312      2     2.733
## 78814 84866279731 3.877 2012     3308      3     2.735
## 79687 84873204638 3.715 2012     1202      3     2.530
## 80242 84864517335 4.300 2012     3312      1     3.158
## 83957 83055186560 3.802 2012     3308      3     2.660
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
```

```
## -1.2976 -0.4484  0.0242  0.4498  3.1519
##
## Coefficients:
##                   Estimate Std. Error t value Pr(>|t|)
## (Intercept)        1.16432    0.01376   84.65  < 2e-16 ***
## LastAuthorFemale1  0.05794    0.00559   10.36  < 2e-16 ***
## Year1997          -0.04167    0.01943   -2.14  0.03202 *
## Year1998          -0.06054    0.01892   -3.20  0.00137 **
## Year1999          -0.06954    0.01846   -3.77  0.00017 ***
## Year2000          -0.03364    0.01766   -1.90  0.05683 .
## Year2001          -0.05519    0.01809   -3.05  0.00229 **
## Year2002           0.07535    0.01820    4.14  3.5e-05 ***
## Year2003           0.04614    0.01930    2.39  0.01680 *
## Year2004           0.00340    0.01879    0.18  0.85636
## Year2005           0.01810    0.01822    0.99  0.32048
## Year2006           0.03460    0.01712    2.02  0.04331 *
## Year2007           0.05484    0.01717    3.19  0.00140 **
## Year2008           0.04708    0.01702    2.77  0.00566 **
## Year2009           0.01490    0.01663    0.90  0.37050
## Year2010          -0.02125    0.01644   -1.29  0.19626
## Year2011           0.00151    0.01670    0.09  0.92815
## Year2012          -0.01620    0.01706   -0.95  0.34216
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.663
## Multiple R-squared:  0.0052, Adjusted R-squared:  0.00493
## Convergence in 10 IRWLS iterations
##
## Robustness weights:
##  observation 59513 is an outlier with |weight| = 0 ( < 1.6e-06);
##  5299 weights are ~= 1. The remaining 57335 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0494  0.8690  0.9500  0.9080  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.60e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                    psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
```

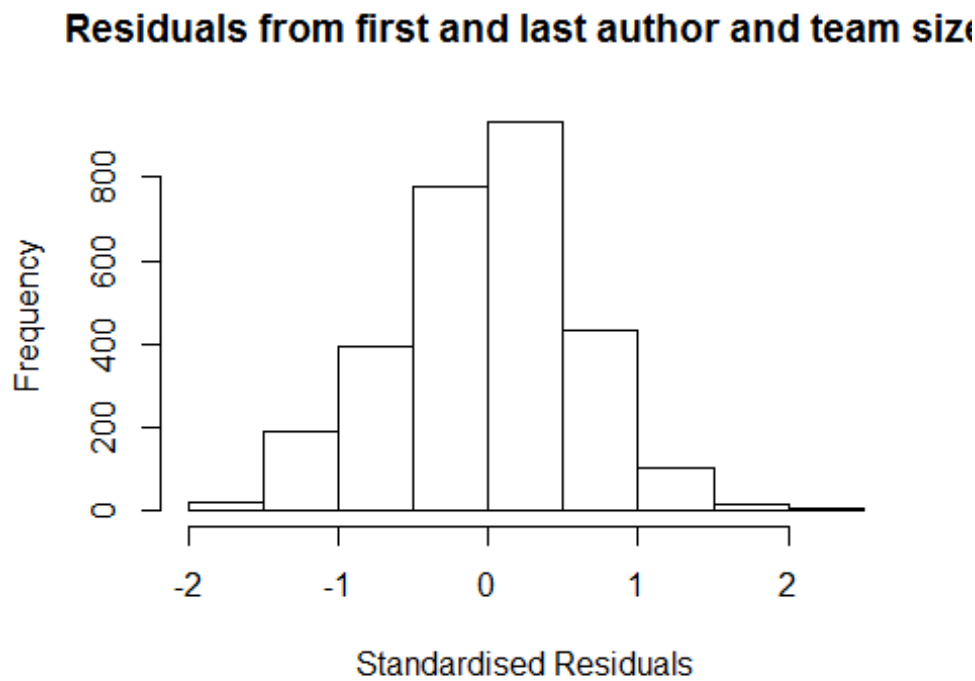
```
## [1] "Sample size for the above analysis:  62635"
## [1] ""
## [1] ""
## [1] "####################################"
## [1] "Analysis of AJSC 3313"
## [1] "####################################"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  220  250  226  231  216  247  263  271  271  290  321  337  361  318  385
## 2011 2012
##  413  456
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  158  182  154  165  143  140  198  198  200  193  202  205  236  194  229
## 2011 2012
##  255  289
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  140  159  135  150  126  124  173  177  166  165  168  180  205  158  191
## 2011 2012
##  210  238
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
##  Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 160, df = 16, p-value <2e-16
```




```
##
##  Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 0.21, df = 1, p-value = 0.6
```

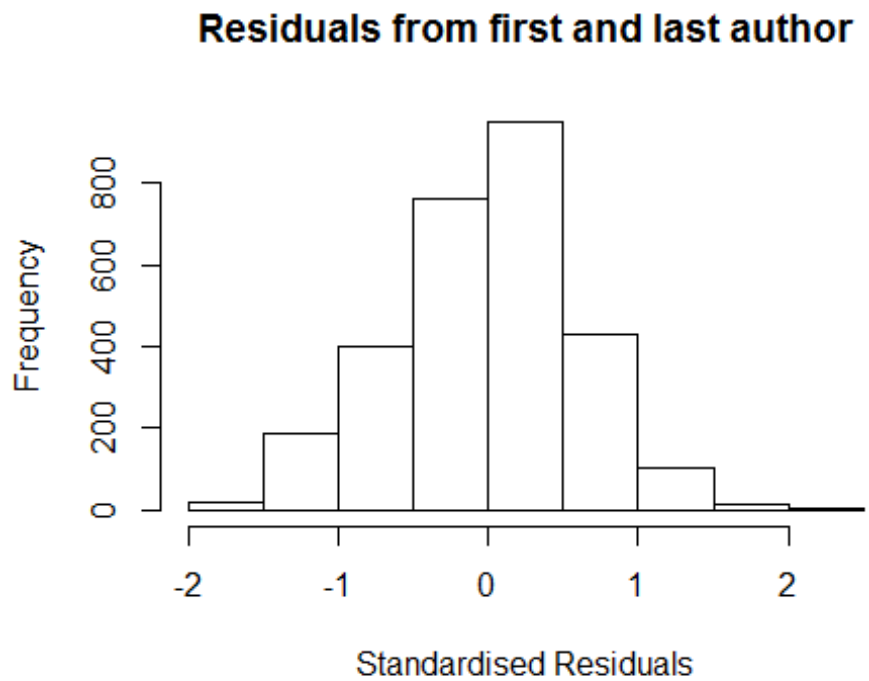



```
## [1] "Female first author team size 2018 geometric mean: 2.509322531329"
## [1] "Male first author team size 2018 geometric mean: 2.41597345535368"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 11000, p-value = 0.5
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 2.42370355152338"
## [1] "Male last author team size 2018 geometric mean: 2.4444262907142"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 11000, p-value = 0.9
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.169  1           1.081
## LastAuthorFemale  1.155  1           1.075
## UniqueAuthors     1.100  4           1.012
## Year              1.146 16           1.004
```



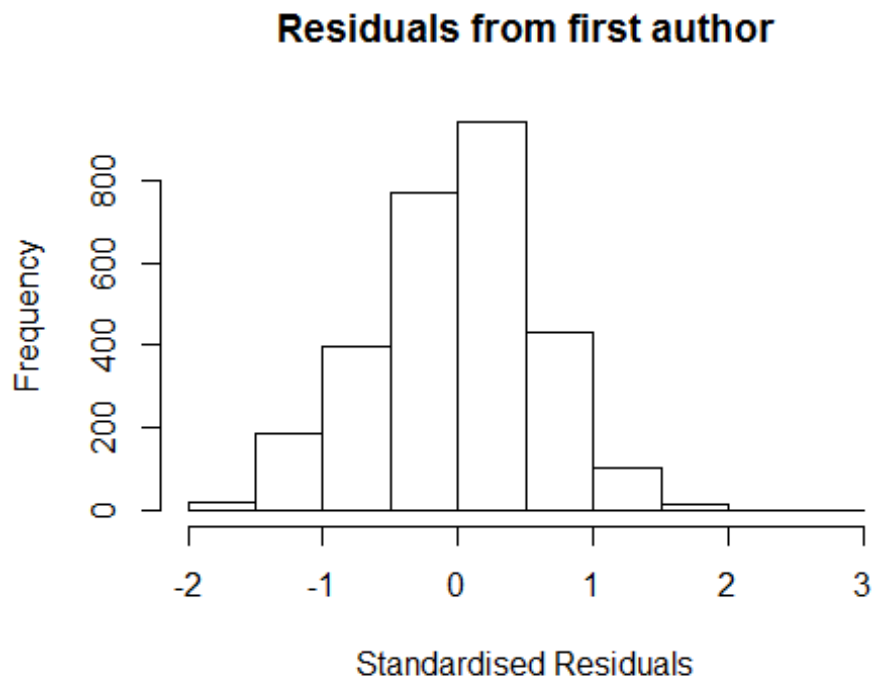

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
UniqueAuthors +
##     Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##     k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.7720 -0.4057  0.0268  0.3982  2.3703
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)          1.4453     0.0739   19.57  < 2e-16 ***
## FirstAuthorFemale1   0.1091     0.0306    3.57  0.00036 ***
## LastAuthorFemale1    0.0240     0.0314    0.77  0.44420
## UniqueAuthors2       0.2318     0.0308    7.53  6.8e-14 ***
## UniqueAuthors3       0.2459     0.0335    7.35  2.6e-13 ***
## UniqueAuthors4       0.1881     0.0472    3.99  6.8e-05 ***
## UniqueAuthors5       0.1401     0.0635    2.20  0.02755 *
## Year1997            -0.2134     0.0999   -2.14  0.03284 *
## Year1998            -0.2170     0.1025   -2.12  0.03435 *
## Year1999            -0.2096     0.0934   -2.24  0.02497 *
```

```
## Year2000            -0.3028     0.1002   -3.02  0.00252 **
## Year2001             0.0774     0.0966    0.80  0.42266
## Year2002            -0.1102     0.0891   -1.24  0.21631
## Year2003            -0.2824     0.0869   -3.25  0.00118 **
## Year2004            -0.2247     0.0822   -2.73  0.00630 **
## Year2005            -0.3163     0.0834   -3.79  0.00015 ***
## Year2006            -0.4154     0.0820   -5.07  4.3e-07 ***
## Year2007            -0.3751     0.0809   -4.64  3.7e-06 ***
## Year2008            -0.3569     0.0801   -4.46  8.6e-06 ***
## Year2009            -0.3715     0.0831   -4.47  8.0e-06 ***
## Year2010            -0.4372     0.0813   -5.38  8.2e-08 ***
## Year2011            -0.4491     0.0823   -5.46  5.3e-08 ***
## Year2012            -0.4391     0.0821   -5.35  9.5e-08 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.594
## Multiple R-squared:  0.0759, Adjusted R-squared:  0.0687
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
##  245 weights are ~= 1. The remaining 2620 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0747  0.8670  0.9490  0.9020  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          3.49e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.158  1           1.076
## LastAuthorFemale  1.151  1           1.073
## Year              1.053 16           1.002
```



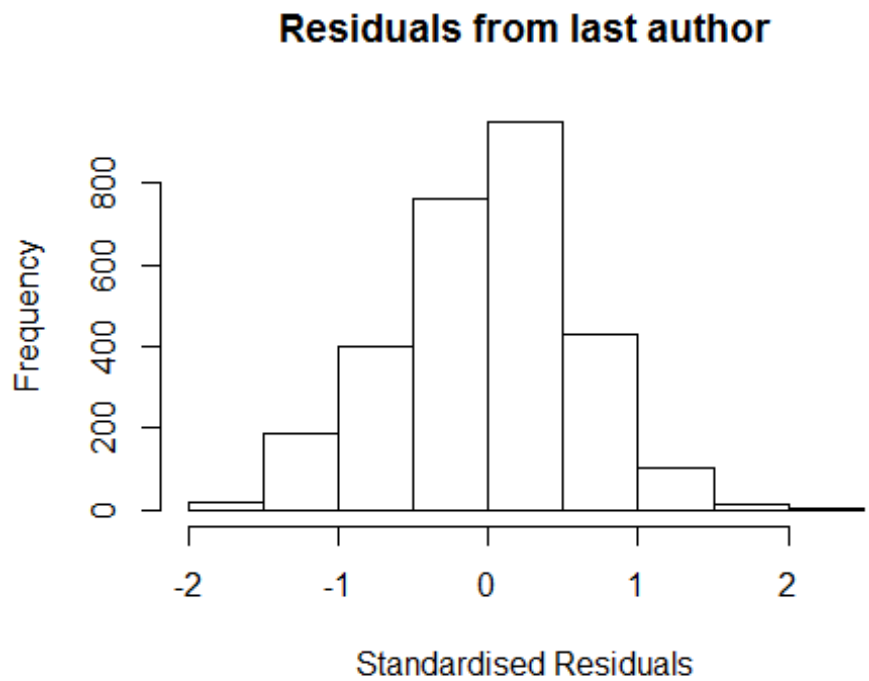

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.8028 -0.4030  0.0389  0.4108  2.4762
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)           1.5830     0.0707   22.38  < 2e-16 ***
## FirstAuthorFemale1    0.1206     0.0309    3.91  9.5e-05 ***
## LastAuthorFemale1     0.0287     0.0318    0.90  0.36573
## Year1997             -0.2299     0.1005   -2.29  0.02217 *
## Year1998             -0.2021     0.1029   -1.96  0.04968 *
## Year1999             -0.2044     0.0942   -2.17  0.03011 *
## Year2000             -0.2782     0.0990   -2.81  0.00500 **
## Year2001              0.0991     0.0976    1.02  0.30999
## Year2002             -0.1066     0.0895   -1.19  0.23390
## Year2003             -0.2657     0.0860   -3.09  0.00202 **
## Year2004             -0.2049     0.0811   -2.53  0.01157 *
## Year2005             -0.2983     0.0830   -3.60  0.00033 ***
```

```
## Year2006            -0.3979     0.0817   -4.87  1.2e-06 ***
## Year2007            -0.3398     0.0795   -4.28  2.0e-05 ***
## Year2008            -0.3396     0.0786   -4.32  1.6e-05 ***
## Year2009            -0.3502     0.0824   -4.25  2.2e-05 ***
## Year2010            -0.4086     0.0798   -5.12  3.2e-07 ***
## Year2011            -0.4145     0.0812   -5.11  3.5e-07 ***
## Year2012            -0.4024     0.0805   -5.00  6.2e-07 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.599
## Multiple R-squared:  0.0487, Adjusted R-squared:  0.0427
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
##  220 weights are ~= 1. The remaining 2645 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0487  0.8670  0.9510  0.9020  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          3.49e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.032  1           1.016
## Year              1.032 16           1.001
```




```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##    Min     1Q Median     3Q    Max
## -1.818 -0.407  0.036  0.408  2.502
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)           1.5875     0.0704   22.56  < 2e-16 ***
## FirstAuthorFemale1    0.1305     0.0291    4.48  7.6e-06 ***
## Year1997             -0.2311     0.1005   -2.30  0.02150 *
## Year1998             -0.2036     0.1030   -1.98  0.04818 *
## Year1999             -0.2066     0.0941   -2.20  0.02818 *
## Year2000             -0.2796     0.0991   -2.82  0.00480 **
## Year2001              0.0999     0.0975    1.02  0.30576
## Year2002             -0.1070     0.0895   -1.20  0.23190
## Year2003             -0.2675     0.0859   -3.12  0.00185 **
## Year2004             -0.2048     0.0811   -2.53  0.01161 *
## Year2005             -0.2992     0.0829   -3.61  0.00031 ***
## Year2006             -0.4000     0.0815   -4.91  9.8e-07 ***
```

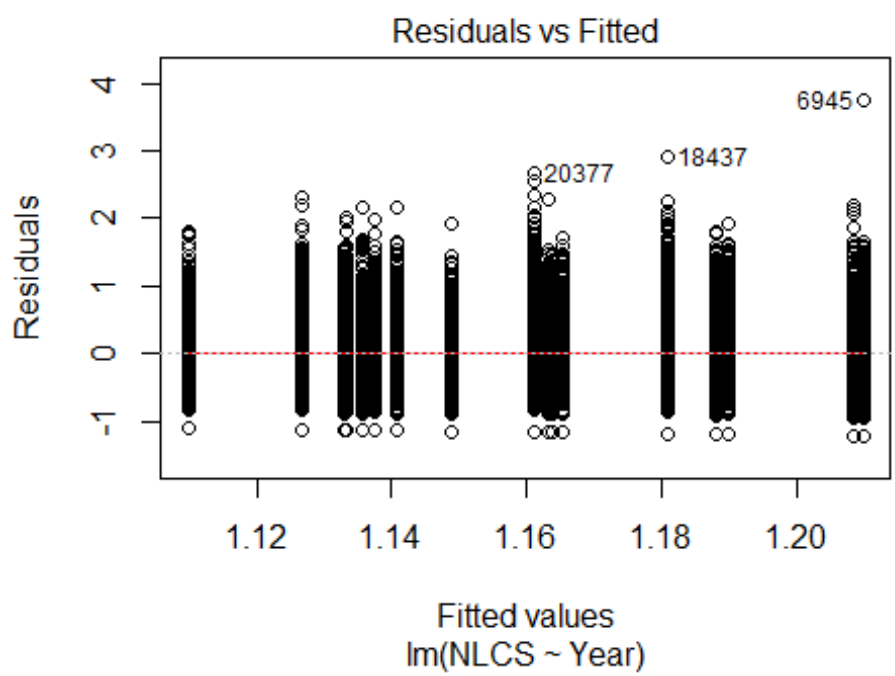
```
## Year2007            -0.3399     0.0795   -4.28  2.0e-05 ***
## Year2008            -0.3404     0.0785   -4.34  1.5e-05 ***
## Year2009            -0.3511     0.0822   -4.27  2.0e-05 ***
## Year2010            -0.4108     0.0797   -5.16  2.7e-07 ***
## Year2011            -0.4161     0.0811   -5.13  3.0e-07 ***
## Year2012            -0.4030     0.0805   -5.01  5.9e-07 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.599
## Multiple R-squared:  0.0486, Adjusted R-squared:  0.0429
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
##  219 weights are ~= 1. The remaining 2646 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0417  0.8670  0.9510  0.9010  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          3.49e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##                   GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.023  1           1.012
## Year             1.023 16           1.001
```



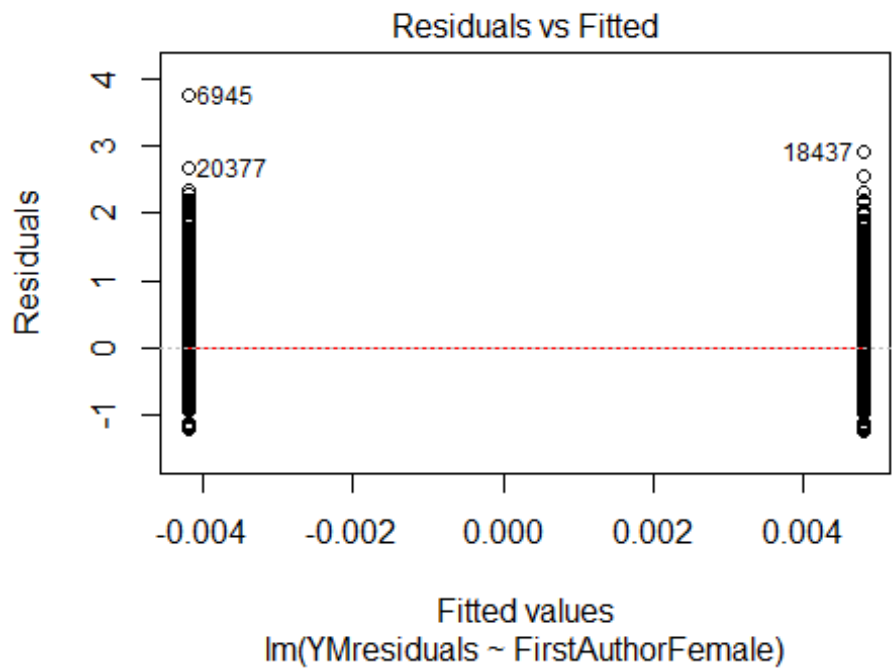

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.6963 -0.4113  0.0355  0.4176  2.4244
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)          1.5934     0.0712   22.37  < 2e-16 ***
## LastAuthorFemale1    0.0713     0.0299    2.38  0.01729 *
## Year1997            -0.2311     0.1010   -2.29  0.02223 *
## Year1998            -0.2013     0.1038   -1.94  0.05267 .
## Year1999            -0.2075     0.0945   -2.20  0.02821 *
## Year2000            -0.2739     0.0996   -2.75  0.00600 **
## Year2001             0.1029     0.0991    1.04  0.29921
## Year2002            -0.1100     0.0898   -1.22  0.22087
## Year2003            -0.2571     0.0865   -2.97  0.00297 **
## Year2004            -0.2047     0.0818   -2.50  0.01235 *
## Year2005            -0.2971     0.0836   -3.55  0.00039 ***
## Year2006            -0.3839     0.0824   -4.66  3.3e-06 ***
```

```
## Year2007           -0.3282     0.0800   -4.10  4.2e-05 ***
## Year2008           -0.3356     0.0795   -4.22  2.5e-05 ***
## Year2009           -0.3440     0.0829   -4.15  3.4e-05 ***
## Year2010           -0.3921     0.0807   -4.86  1.2e-06 ***
## Year2011           -0.4021     0.0814   -4.94  8.4e-07 ***
## Year2012           -0.3969     0.0811   -4.89  1.1e-06 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.602
## Multiple R-squared:  0.0437, Adjusted R-squared:  0.038
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
##  219 weights are ~= 1. The remaining 2646 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0679  0.8660  0.9510  0.9020  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          3.49e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Sample size for the above analysis:  2865"
## [1] ""
## [1] ""
## [1] "####################################"
## [1] "Analysis of AJSC 3314"
## [1] "####################################"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  799  788  932  936  958  958  977  860  886  891 1161 1170 1293 1472 1666
## 2011 2012
## 1703 1750
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  677  654  786  785  785  735  841  736  768  774 1000 1005 1100 1242 1411
## 2011 2012
```

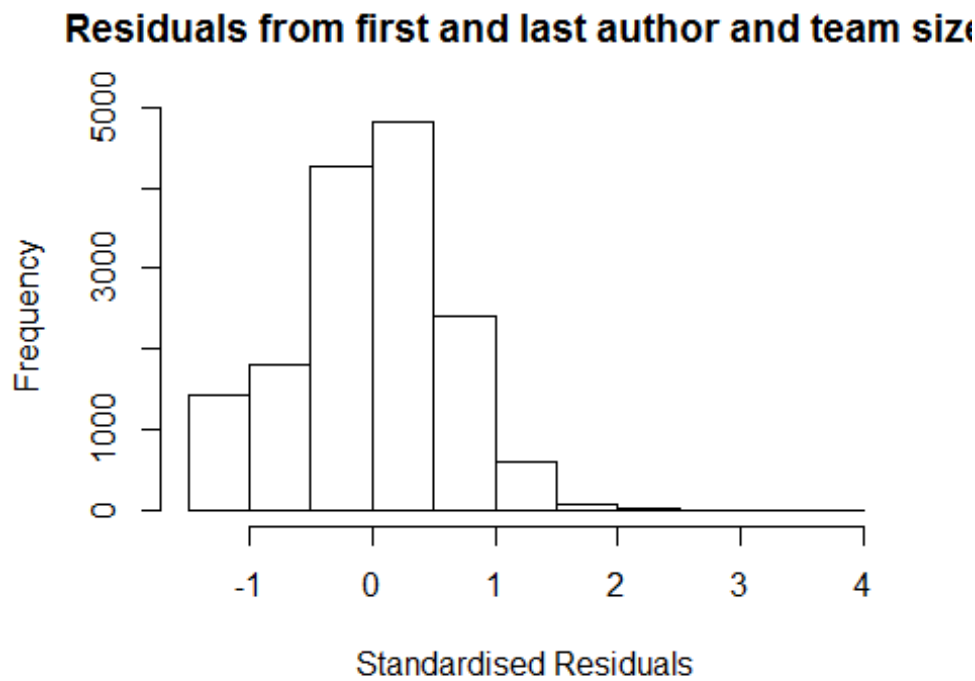
```
## 1453 1430
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  653  624  750  756  758  701  793  708  727  739  957  951 1027 1171 1346
## 2011 2012
## 1385 1369
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
##  Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 140, df = 16, p-value <2e-16
```




```
##
##  Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 36, df = 1, p-value = 2e-09
```

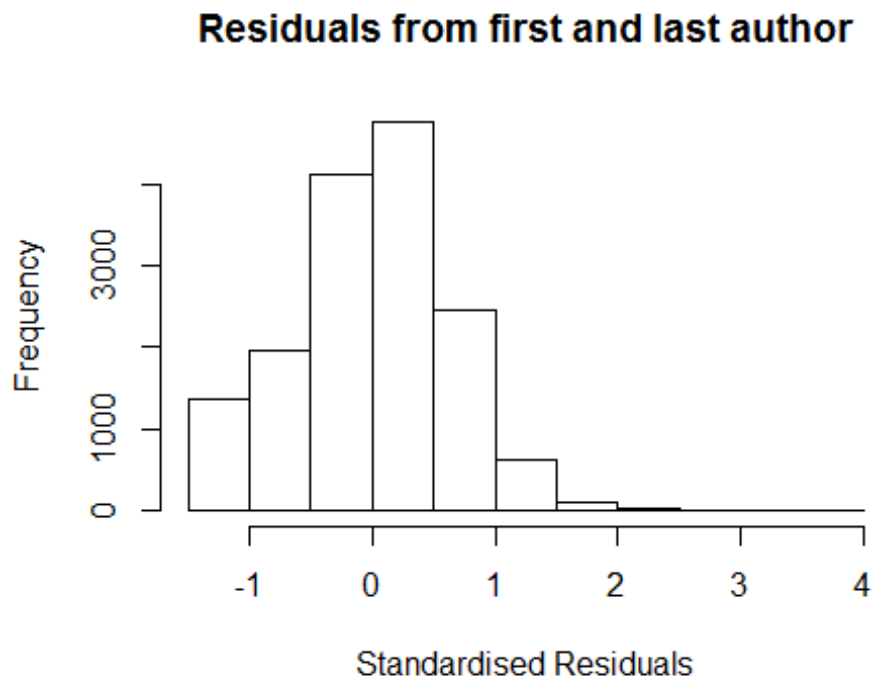



```
## [1] "Female first author team size 2018 geometric mean: 1.59882736575899"
## [1] "Male first author team size 2018 geometric mean: 1.52256898418929"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 270000, p-value = 0.1
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 1.56683094654279"
## [1] "Male last author team size 2018 geometric mean: 1.5604753597192"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 260000, p-value = 0.9
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.795  1           1.340
## LastAuthorFemale  1.789  1           1.338
## UniqueAuthors     1.026  4           1.003
## Year              1.026 16           1.001
```



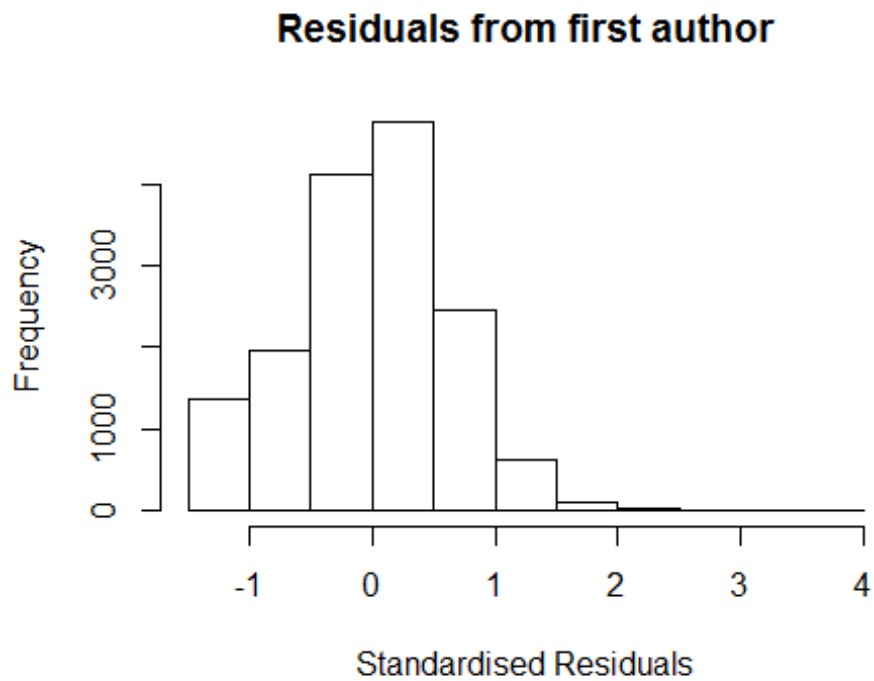

```
## [1] "List of  3 outliers with residuals above  2.5"
##          ScopusId  NLCS Year OneField Fields residuals
## 6945   3142779116 4.956 2002     1213      3     3.867
## 18437 79958705208 4.106 2011     1213      3     3.052
## 20377 84860699617 3.836 2012     3314      1     2.793
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
UniqueAuthors +
##     Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##     k.max = 1000))
##  \--> method = "MM"
## Residuals:
##    Min     1Q Median     3Q    Max
## -1.425 -0.408  0.023  0.411  3.867
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)          1.06807    0.02549   41.90   <2e-16 ***
## FirstAuthorFemale1   0.00881    0.01355    0.65    0.516
## LastAuthorFemale1    0.00359    0.01354    0.27    0.791
## UniqueAuthors2       0.27319    0.01278   21.37   <2e-16 ***
## UniqueAuthors3       0.32072    0.01689   18.99   <2e-16 ***
## UniqueAuthors4       0.32872    0.02419   13.59   <2e-16 ***
## UniqueAuthors5       0.31032    0.02769   11.21   <2e-16 ***
## Year1997            -0.00264    0.03412   -0.08    0.938
```

```
## Year1998           -0.03891    0.03296   -1.18    0.238
## Year1999           -0.07016    0.03301   -2.13    0.034 *
## Year2000           -0.03925    0.03259   -1.20    0.229
## Year2001           -0.02808    0.03343   -0.84    0.401
## Year2002            0.02117    0.03237    0.65    0.513
## Year2003           -0.02636    0.03449   -0.76    0.445
## Year2004            0.00811    0.03269    0.25    0.804
## Year2005            0.02511    0.03362    0.75    0.455
## Year2006           -0.00127    0.03106   -0.04    0.967
## Year2007           -0.00164    0.03135   -0.05    0.958
## Year2008           -0.04517    0.03167   -1.43    0.154
## Year2009           -0.04967    0.03060   -1.62    0.105
## Year2010           -0.06297    0.02997   -2.10    0.036 *
## Year2011           -0.02636    0.03067   -0.86    0.390
## Year2012           -0.02527    0.03245   -0.78    0.436
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.603
## Multiple R-squared:  0.0511, Adjusted R-squared:  0.0497
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  2 observations c(4931,13431) are outliers with |weight| = 0 ( < 6.5e-06);
##  1335 weights are ~= 1. The remaining 14078 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0005  0.8670  0.9490  0.9060  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          6.49e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.869  1           1.367
## LastAuthorFemale  1.867  1           1.366
## Year              1.008 16           1.000
```



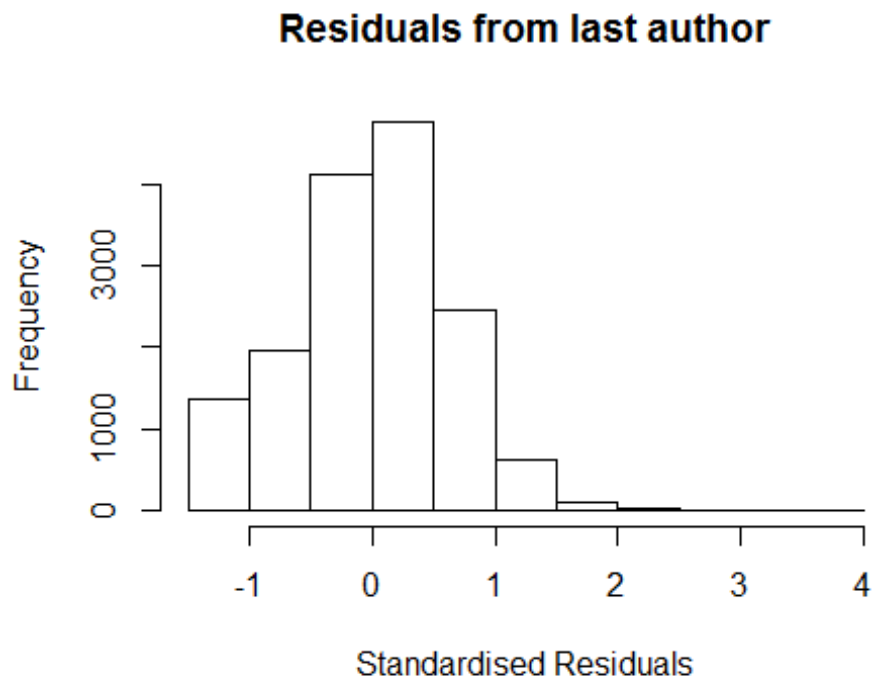

```
## [1] "List of  4 outliers with residuals above  2.5"
##          ScopusId  NLCS Year OneField Fields residuals
## 6945   3142779116 4.956 2002     1213      3     3.775
## 18437 79958705208 4.106 2011     1213      3     2.941
## 19962 84873204638 3.715 2012     1202      3     2.550
## 20377 84860699617 3.836 2012     3314      1     2.697
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.2152 -0.4149  0.0245  0.4195  3.7755
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.16313    0.02639   44.07   <2e-16 ***
## FirstAuthorFemale1  0.02634    0.01422    1.85    0.064 .
## LastAuthorFemale1  -0.00872    0.01423   -0.61    0.540
## Year1997           -0.01351    0.03571   -0.38    0.705
## Year1998           -0.03920    0.03390   -1.16    0.247
## Year1999           -0.07428    0.03470   -2.14    0.032 *
## Year2000           -0.04132    0.03368   -1.23    0.220
## Year2001           -0.02570    0.03465   -0.74    0.458
## Year2002            0.01740    0.03380    0.51    0.607
```

```
## Year2003           -0.03933    0.03587   -1.10    0.273
## Year2004            0.01315    0.03421    0.38    0.701
## Year2005            0.02573    0.03469    0.74    0.458
## Year2006           -0.00243    0.03235   -0.08    0.940
## Year2007            0.01015    0.03237    0.31    0.754
## Year2008           -0.03668    0.03297   -1.11    0.266
## Year2009           -0.04050    0.03197   -1.27    0.205
## Year2010           -0.06041    0.03145   -1.92    0.055 .
## Year2011           -0.01586    0.03190   -0.50    0.619
## Year2012           -0.02426    0.03362   -0.72    0.471
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.616
## Multiple R-squared:  0.0021, Adjusted R-squared:  0.000937
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  2 observations c(4931,13431) are outliers with |weight| = 0 ( < 6.5e-06);
##  1317 weights are ~= 1. The remaining 14096 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0163  0.8660  0.9490  0.9060  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          6.49e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##                     GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.006  1           1.003
## Year              1.006 16           1.000
```




```
## [1] "List of  4 outliers with residuals above  2.5"
##          ScopusId  NLCS Year OneField Fields residuals
## 6945   3142779116 4.956 2002     1213      3     3.775
## 18437 79958705208 4.106 2011     1213      3     2.941
## 19962 84873204638 3.715 2012     1202      3     2.550
## 20377 84860699617 3.836 2012     3314      1     2.697
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.2078 -0.4148  0.0239  0.4193  3.7764
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.16211    0.02631   44.17   <2e-16 ***
## FirstAuthorFemale1  0.02018    0.01044    1.93    0.053 .
## Year1997           -0.01364    0.03571   -0.38    0.702
## Year1998           -0.03931    0.03390   -1.16    0.246
## Year1999           -0.07430    0.03469   -2.14    0.032 *
## Year2000           -0.04146    0.03368   -1.23    0.218
## Year2001           -0.02568    0.03464   -0.74    0.458
## Year2002            0.01752    0.03379    0.52    0.604
## Year2003           -0.03935    0.03587   -1.10    0.273
```

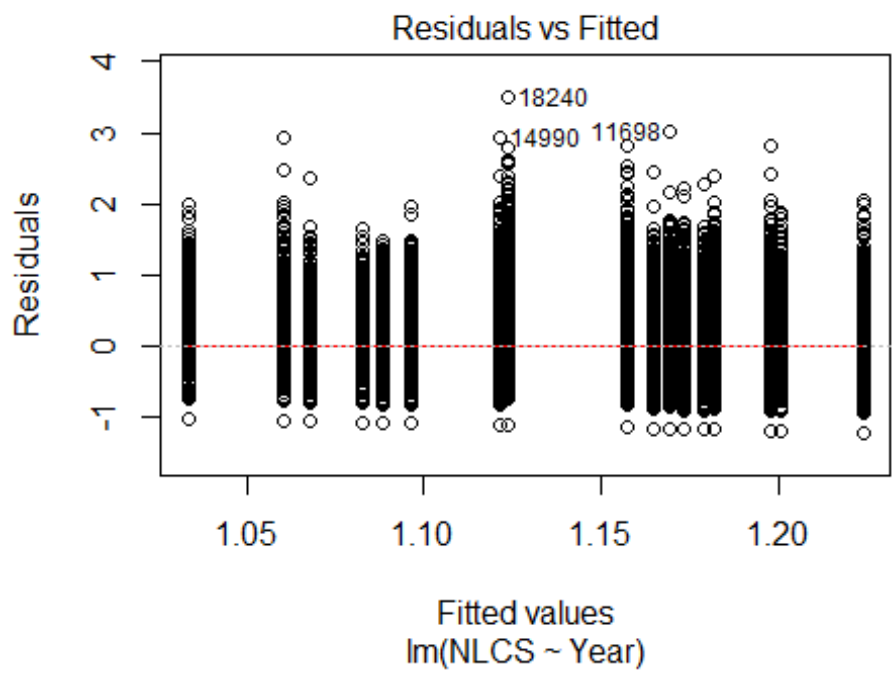
```
## Year2004            0.01305    0.03421    0.38    0.703
## Year2005            0.02551    0.03469    0.74    0.462
## Year2006           -0.00268    0.03236   -0.08    0.934
## Year2007            0.01000    0.03238    0.31    0.758
## Year2008           -0.03687    0.03297   -1.12    0.263
## Year2009           -0.04056    0.03197   -1.27    0.205
## Year2010           -0.06058    0.03145   -1.93    0.054 .
## Year2011           -0.01596    0.03191   -0.50    0.617
## Year2012           -0.02428    0.03363   -0.72    0.470
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.616
## Multiple R-squared:  0.00208,    Adjusted R-squared:  0.000978
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  2 observations c(4931,13431) are outliers with |weight| = 0 ( < 6.5e-06);
##  1313 weights are ~= 1. The remaining 14100 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0161  0.8650  0.9490  0.9060  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          6.49e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                    psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##                   GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.005  1           1.002
## Year             1.005 16           1.000
```



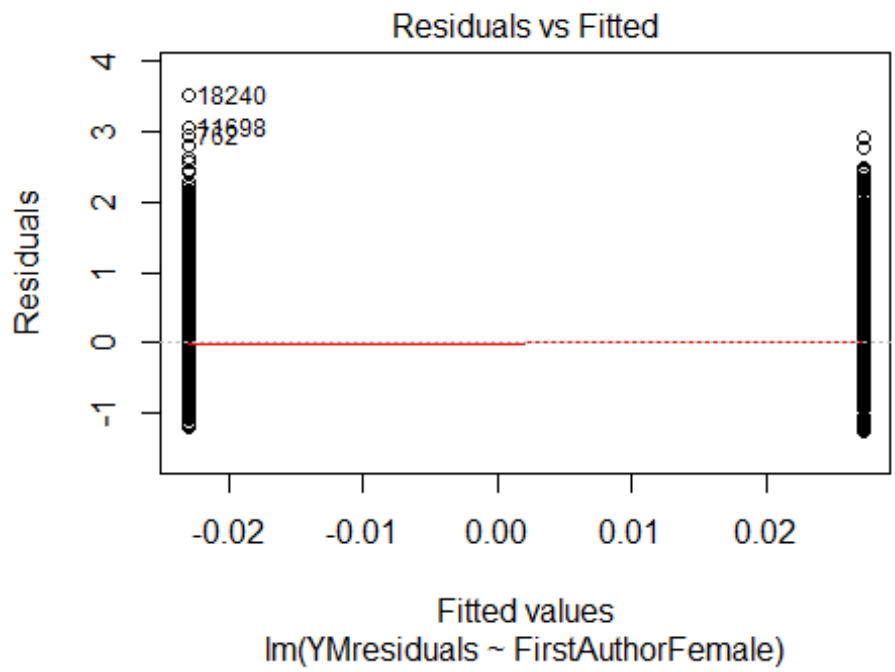

```
## [1] "List of  4 outliers with residuals above  2.5"
##          ScopusId  NLCS Year OneField Fields residuals
## 6945   3142779116 4.956 2002     1213      3     3.775
## 18437 79958705208 4.106 2011     1213      3     2.941
## 19962 84873204638 3.715 2012     1202      3     2.550
## 20377 84860699617 3.836 2012     3314      1     2.697
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.2024 -0.4154  0.0242  0.4187  3.7710
##
## Coefficients:
##                   Estimate Std. Error t value Pr(>|t|)
## (Intercept)        1.16659    0.02631   44.35   <2e-16 ***
## LastAuthorFemale1  0.00993    0.01045    0.95    0.342
## Year1997          -0.01333    0.03570   -0.37    0.709
## Year1998          -0.03922    0.03390   -1.16    0.247
## Year1999          -0.07416    0.03469   -2.14    0.033 *
## Year2000          -0.04129    0.03369   -1.23    0.220
## Year2001          -0.02548    0.03464   -0.74    0.462
## Year2002           0.01839    0.03379    0.54    0.586
## Year2003          -0.03896    0.03588   -1.09    0.278
```

```
## Year2004           0.01304    0.03420    0.38    0.703
## Year2005           0.02589    0.03470    0.75    0.456
## Year2006          -0.00253    0.03236   -0.08    0.938
## Year2007           0.01093    0.03239    0.34    0.736
## Year2008          -0.03638    0.03298   -1.10    0.270
## Year2009          -0.03996    0.03197   -1.25    0.211
## Year2010          -0.05995    0.03145   -1.91    0.057 .
## Year2011          -0.01480    0.03191   -0.46    0.643
## Year2012          -0.02322    0.03363   -0.69    0.490
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.616
## Multiple R-squared:  0.00189,    Adjusted R-squared:  0.000785
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  2 observations c(4931,13431) are outliers with |weight| = 0 ( < 6.5e-06);
##  1305 weights are ~= 1. The remaining 14108 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0171  0.8660  0.9490  0.9060  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          6.49e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Sample size for the above analysis:  15415"
## [1] ""
## [1] ""
## [1] "##################################"
## [1] "Analysis of AJSC 3315"
## [1] "##################################"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  647  588  584  648  828  800  860  723  751  862 1050 1149 1189 1454 1844
## 2011 2012
## 2001 2000
```

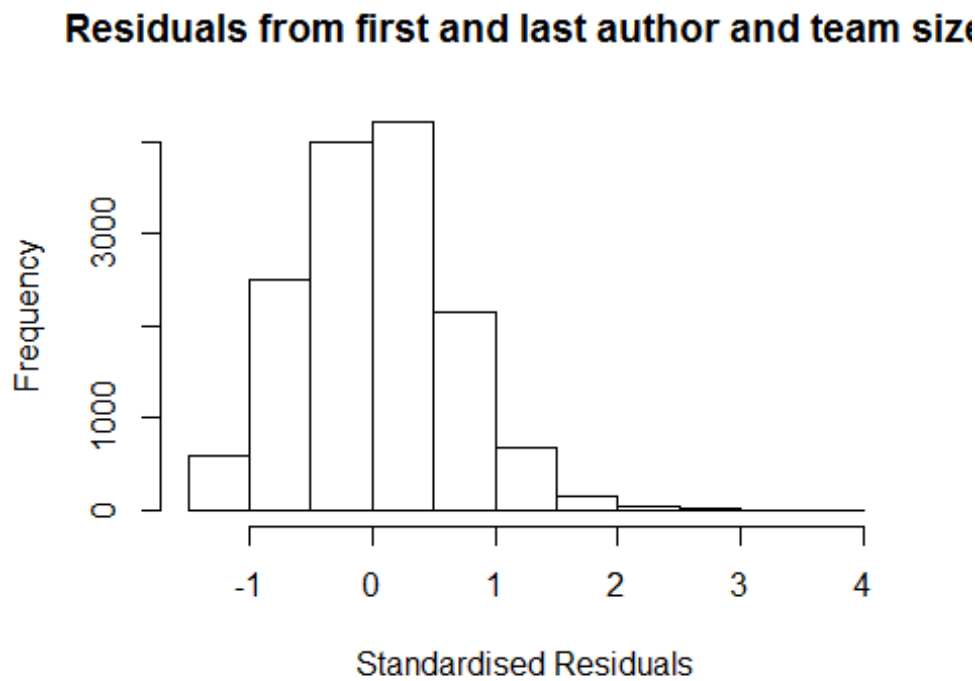
```
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  569  517  520  570  712  679  743  638  642  735  888  973  994 1179 1545
## 2011 2012
## 1645 1645
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  553  500  502  551  684  654  707  601  601  705  842  912  936 1087 1430
## 2011 2012
## 1524 1547
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
##  Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 200, df = 16, p-value <2e-16
```




```
##
##  Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 31, df = 1, p-value = 3e-08
```



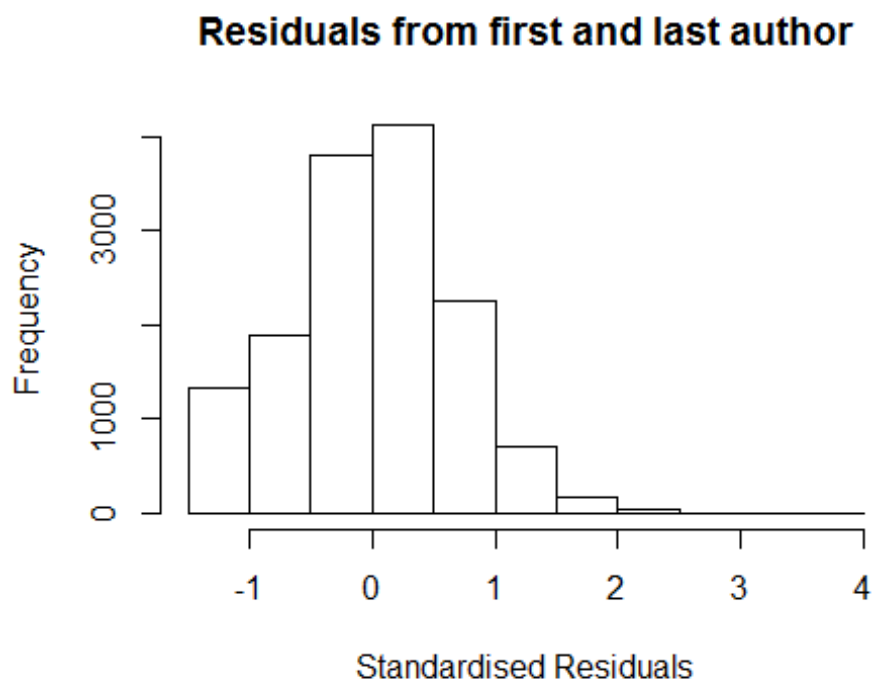

```
## [1] "Female first author team size 2018 geometric mean: 1.69265850420147"
## [1] "Male first author team size 2018 geometric mean: 1.59893210427427"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 410000, p-value = 0.06
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 1.67753714891689"
## [1] "Male last author team size 2018 geometric mean: 1.61654464029563"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 410000, p-value = 0.3
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.549  1           1.245
## LastAuthorFemale  1.540  1           1.241
## UniqueAuthors     1.055  4           1.007
## Year              1.059 16           1.002
```




```
## [1] "List of  13 outliers with residuals above  2.5"
##          ScopusId  NLCS Year OneField Fields residuals
## 726    0009431359 3.532 1997     1208      2     2.578
## 762   79959433454 3.987 1997     1208      2     3.089
## 5197  84998146028 3.599 2002     1213      2     2.600
## 7536  21544433512 4.003 2005     1208      2     2.667
## 7910  33846333428 3.605 2005     1208      3     2.594
## 10922 45349087542 3.582 2008     1208      2     2.617
## 11698 70949095633 4.199 2009     1203      2     3.247
## 14990 73249128238 4.062 2010     1208      2     2.753
## 16653 79955674091 3.677 2011     1203      6     2.682
## 16852 79951887862 3.965 2011     3312      2     2.663
## 18240 84864035512 4.620 2012     1208      2     3.723
## 19070 84856523526 3.906 2012     1203      3     2.703
## 19399 84858270024 3.720 2012     3207      3     2.823
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
UniqueAuthors +
##     Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##     k.max = 1000))
##  \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.48478 -0.41644  0.00898  0.41940  3.72281
##
```

```
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)           0.8729     0.0308   28.34  < 2e-16 ***
## FirstAuthorFemale1    0.0183     0.0134    1.37  0.16971
## LastAuthorFemale1     0.0373     0.0133    2.79  0.00520 **
## UniqueAuthors2        0.3062     0.0132   23.27  < 2e-16 ***
## UniqueAuthors3        0.3740     0.0179   20.93  < 2e-16 ***
## UniqueAuthors4        0.4277     0.0253   16.93  < 2e-16 ***
## UniqueAuthors5        0.4894     0.0235   20.80  < 2e-16 ***
## Year1997              0.0251     0.0461    0.54  0.58617
## Year1998              0.0599     0.0422    1.42  0.15534
## Year1999              0.0479     0.0383    1.25  0.21116
## Year2000              0.0534     0.0374    1.43  0.15307
## Year2001              0.0746     0.0376    1.99  0.04701 *
## Year2002              0.1266     0.0393    3.22  0.00128 **
## Year2003              0.1292     0.0402    3.22  0.00130 **
## Year2004              0.1716     0.0403    4.26  2.1e-05 ***
## Year2005              0.1381     0.0392    3.53  0.00042 ***
## Year2006              0.1146     0.0370    3.10  0.00193 **
## Year2007              0.1387     0.0363    3.82  0.00013 ***
## Year2008              0.0924     0.0358    2.58  0.00989 **
## Year2009              0.0791     0.0357    2.22  0.02659 *
## Year2010              0.0442     0.0345    1.28  0.19980
## Year2011              0.0668     0.0348    1.92  0.05481 .
## Year2012              0.0243     0.0357    0.68  0.49623
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.628
## Multiple R-squared:  0.0795, Adjusted R-squared:  0.0781
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  3 observations c(595,8772,13437) are outliers with |weight| = 0 ( < 7e-
06);
##  1252 weights are ~= 1. The remaining 13081 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0062  0.8620  0.9510  0.9070  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          6.98e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                    psi           subsampling                   cov
```

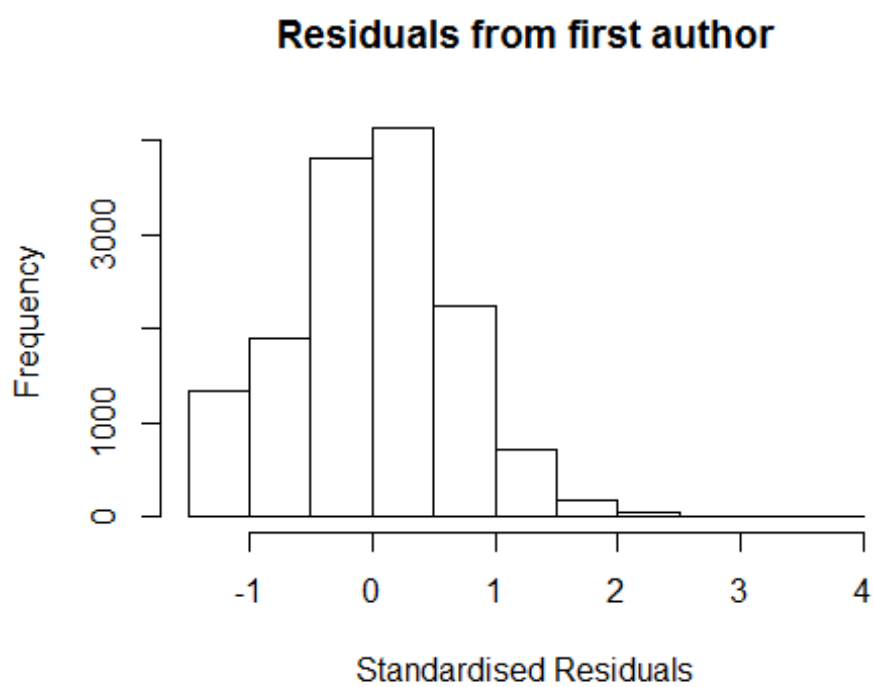
```
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.629  1           1.276
## LastAuthorFemale  1.630  1           1.277
## Year              1.012 16           1.000
```




```
## [1] "List of  10 outliers with residuals above  2.5"
##          ScopusId  NLCS Year OneField Fields residuals
## 762   79959433454 3.987 1997     1208      2     2.998
## 7536  21544433512 4.003 2005     1208      2     2.824
## 11698 70949095633 4.199 2009     1203      2     3.085
## 14990 73249128238 4.062 2010     1208      2     2.958
## 16653 79955674091 3.677 2011     1203      6     2.513
## 16852 79951887862 3.965 2011     3312      2     2.801
## 18240 84864035512 4.620 2012     1208      2     3.587
## 18689 84862335646 3.673 2012     3315      1     2.640
## 19070 84856523526 3.906 2012     1203      3     2.873
## 19399 84858270024 3.720 2012     3207      3     2.687
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
```

```
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.2497 -0.4433  0.0152  0.4429  3.5870
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         0.98110    0.03141   31.24  < 2e-16 ***
## FirstAuthorFemale1  0.04168    0.01432    2.91  0.00362 **
## LastAuthorFemale1   0.03443    0.01435    2.40  0.01648 *
## Year1997            0.00779    0.04734    0.16  0.86938
## Year1998            0.06551    0.04337    1.51  0.13099
## Year1999            0.04382    0.03941    1.11  0.26624
## Year2000            0.05871    0.03853    1.52  0.12759
## Year2001            0.06274    0.03869    1.62  0.10489
## Year2002            0.13664    0.04047    3.38  0.00074 ***
## Year2003            0.13988    0.04104    3.41  0.00066 ***
## Year2004            0.19245    0.04160    4.63  3.8e-06 ***
## Year2005            0.15608    0.04041    3.86  0.00011 ***
## Year2006            0.13864    0.03811    3.64  0.00028 ***
## Year2007            0.18132    0.03741    4.85  1.3e-06 ***
## Year2008            0.14194    0.03695    3.84  0.00012 ***
## Year2009            0.13242    0.03647    3.63  0.00028 ***
## Year2010            0.08079    0.03546    2.28  0.02272 *
## Year2011            0.10646    0.03581    2.97  0.00296 **
## Year2012            0.05187    0.03713    1.40  0.16243
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.654
## Multiple R-squared:  0.00888,    Adjusted R-squared:  0.00763
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  2 observations c(8772,13437) are outliers with |weight| = 0 ( < 7e-06);
##  1201 weights are ~= 1. The remaining 13133 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0017  0.8700  0.9500  0.9080  0.9860  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          6.98e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
```

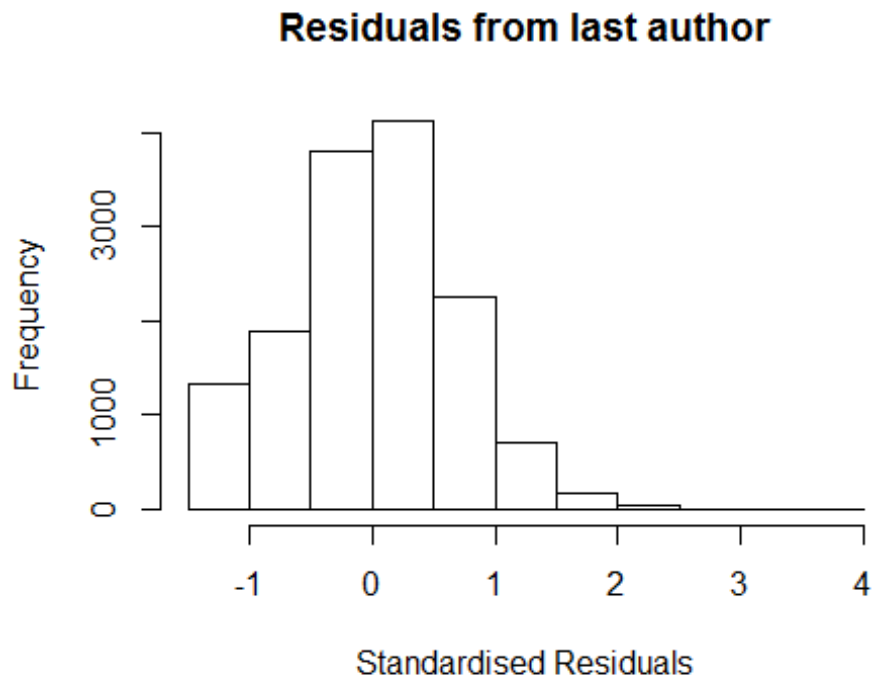
```
##                    psi           subsampling                    cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                   "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##                     GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.007  1           1.004
## Year              1.007 16           1.000
```




```
## [1] "List of  10 outliers with residuals above  2.5"
##          ScopusId  NLCS Year OneField Fields residuals
## 762   79959433454 3.987 1997     1208      2     2.998
## 7536  21544433512 4.003 2005     1208      2     2.824
## 11698 70949095633 4.199 2009     1203      2     3.085
## 14990 73249128238 4.062 2010     1208      2     2.958
## 16653 79955674091 3.677 2011     1203      6     2.513
## 16852 79951887862 3.965 2011     3312      2     2.801
## 18240 84864035512 4.620 2012     1208      2     3.587
## 18689 84862335646 3.673 2012     3315      1     2.640
## 19070 84856523526 3.906 2012     1203      3     2.873
## 19399 84858270024 3.720 2012     3207      3     2.687
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
```

```
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.2437 -0.4438  0.0126  0.4445  3.5812
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         0.98556    0.03138   31.41  < 2e-16 ***
## FirstAuthorFemale1  0.06405    0.01127    5.68  1.4e-08 ***
## Year1997            0.00834    0.04738    0.18  0.86028
## Year1998            0.06555    0.04335    1.51  0.13053
## Year1999            0.04320    0.03942    1.10  0.27318
## Year2000            0.05930    0.03858    1.54  0.12426
## Year2001            0.06370    0.03870    1.65  0.09983 .
## Year2002            0.13821    0.04050    3.41  0.00065 ***
## Year2003            0.14015    0.04105    3.41  0.00064 ***
## Year2004            0.19407    0.04160    4.66  3.1e-06 ***
## Year2005            0.15695    0.04045    3.88  0.00010 ***
## Year2006            0.13916    0.03813    3.65  0.00026 ***
## Year2007            0.18184    0.03743    4.86  1.2e-06 ***
## Year2008            0.14173    0.03696    3.84  0.00013 ***
## Year2009            0.13332    0.03650    3.65  0.00026 ***
## Year2010            0.08186    0.03548    2.31  0.02103 *
## Year2011            0.10716    0.03583    2.99  0.00278 **
## Year2012            0.05327    0.03714    1.43  0.15153
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.654
## Multiple R-squared:  0.00848,    Adjusted R-squared:  0.00731
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  2 observations c(8772,13437) are outliers with |weight| = 0 ( < 7e-06);
##  1197 weights are ~= 1. The remaining 13137 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.002   0.870   0.950   0.908   0.986   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          6.98e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                  psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
```

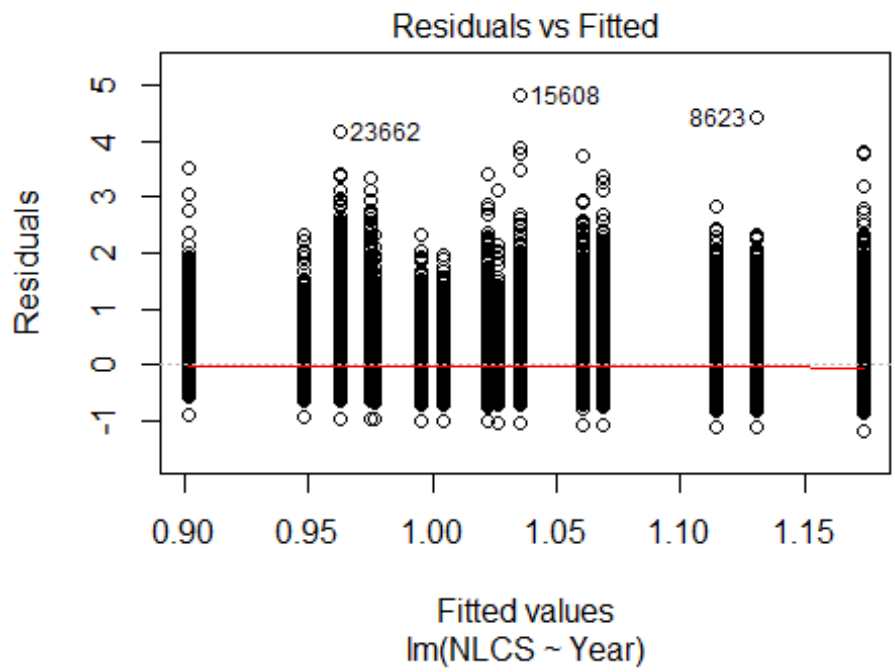
```
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##                   GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.008  1           1.004
## Year             1.008 16           1.000
```



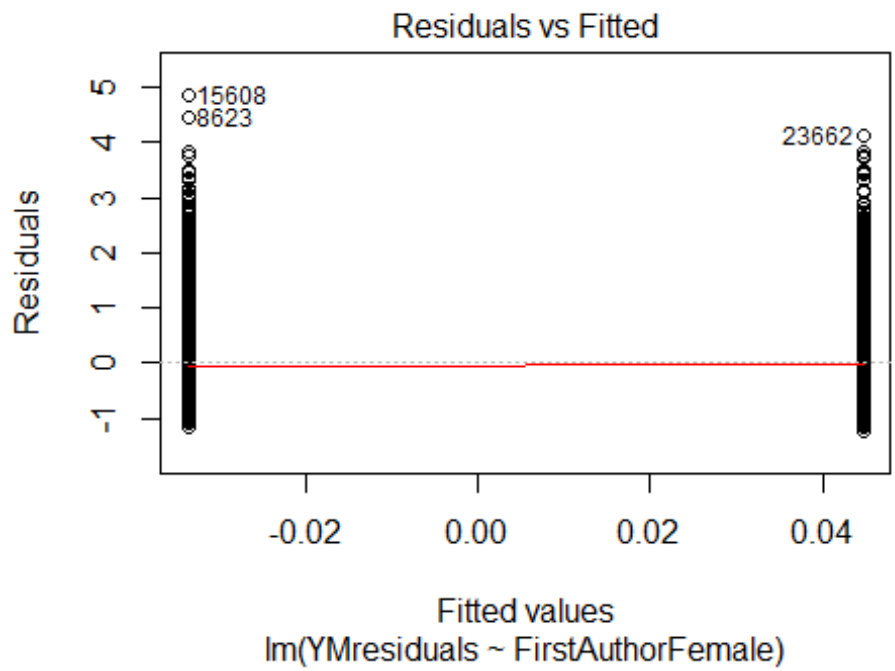

```
## [1] "List of  10 outliers with residuals above  2.5"
##          ScopusId  NLCS Year OneField Fields residuals
## 762   79959433454 3.987 1997     1208      2     2.998
## 7536  21544433512 4.003 2005     1208      2     2.824
## 11698 70949095633 4.199 2009     1203      2     3.085
## 14990 73249128238 4.062 2010     1208      2     2.958
## 16653 79955674091 3.677 2011     1203      6     2.513
## 16852 79951887862 3.965 2011     3312      2     2.801
## 18240 84864035512 4.620 2012     1208      2     3.587
## 18689 84862335646 3.673 2012     3315      1     2.640
## 19070 84856523526 3.906 2012     1203      3     2.873
## 19399 84858270024 3.720 2012     3207      3     2.687
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
```

```
##     Min      1Q  Median      3Q     Max
## -1.2416 -0.4422  0.0128  0.4442  3.5796
##
## Coefficients:
##                   Estimate Std. Error t value Pr(>|t|)
## (Intercept)        0.98775    0.03127   31.59  < 2e-16 ***
## LastAuthorFemale1  0.06162    0.01129    5.46  5.0e-08 ***
## Year1997           0.00727    0.04732    0.15  0.87792
## Year1998           0.06541    0.04341    1.51  0.13190
## Year1999           0.04436    0.03940    1.13  0.26028
## Year2000           0.05785    0.03847    1.50  0.13262
## Year2001           0.06371    0.03867    1.65  0.09949 .
## Year2002           0.13593    0.04044    3.36  0.00078 ***
## Year2003           0.14109    0.04107    3.44  0.00059 ***
## Year2004           0.19224    0.04163    4.62  3.9e-06 ***
## Year2005           0.15581    0.04039    3.86  0.00012 ***
## Year2006           0.13866    0.03810    3.64  0.00027 ***
## Year2007           0.18249    0.03741    4.88  1.1e-06 ***
## Year2008           0.14327    0.03693    3.88  0.00011 ***
## Year2009           0.13301    0.03646    3.65  0.00027 ***
## Year2010           0.08167    0.03545    2.30  0.02125 *
## Year2011           0.10727    0.03580    3.00  0.00274 **
## Year2012           0.05263    0.03712    1.42  0.15629
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.654
## Multiple R-squared:  0.00831,    Adjusted R-squared:  0.00713
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  2 observations c(8772,13437) are outliers with |weight| = 0 ( < 7e-06);
##  1205 weights are ~= 1. The remaining 13129 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0021  0.8700  0.9500  0.9080  0.9860  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          6.98e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
```

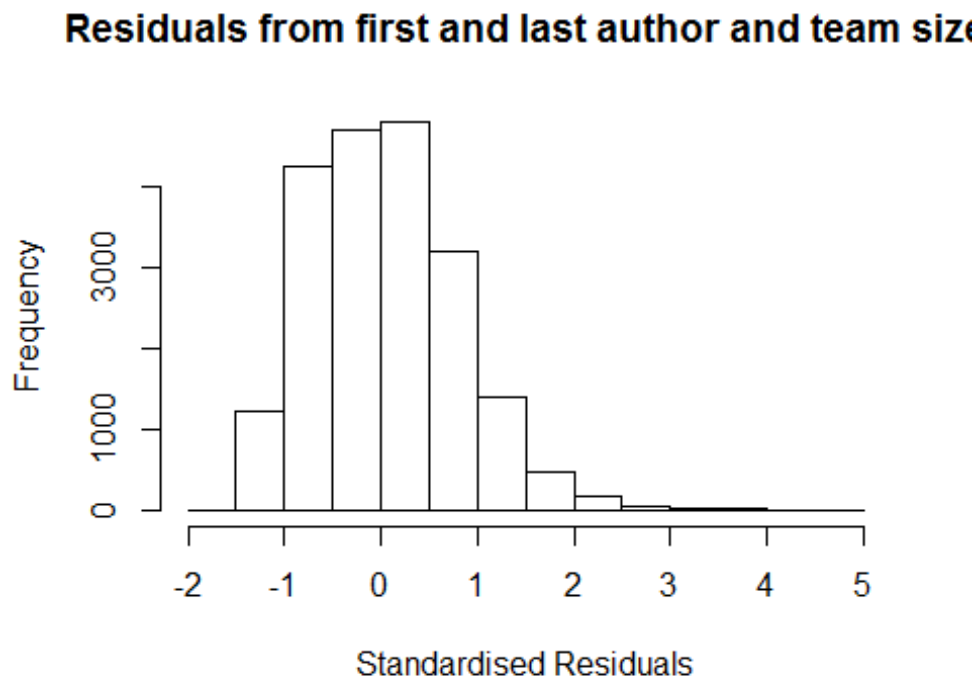
```
## seed : int(0)
## [1] "Sample size for the above analysis:  14336"
## [1] ""
## [1] ""
## [1] "####################################"
## [1] "Analysis of AJSC 3316"
## [1] "####################################"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  726  648  716  866  793  924 1201 1250 1496 1493 1657 1739 1765 1644 2179
## 2011 2012
## 2531 2869
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  631  537  614  734  657  729 1033 1089 1295 1304 1441 1507 1504 1392 1825
## 2011 2012
## 2105 2413
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  618  520  596  714  640  701 1015 1074 1275 1278 1405 1474 1471 1360 1764
## 2011 2012
## 2036 2329
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
##  Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 300, df = 16, p-value <2e-16
```




```
##
##  Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 0.86, df = 1, p-value = 0.4
```




```
## [1] "Female first author team size 2018 geometric mean: 1.30506650604224"
## [1] "Male first author team size 2018 geometric mean: 1.22364045463778"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 770000, p-value = 0.001
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 1.31085779902262"
## [1] "Male last author team size 2018 geometric mean: 1.21753259530978"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 780000, p-value = 2e-04
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 2.982  1           1.727
## LastAuthorFemale  2.978  1           1.726
## UniqueAuthors     1.029  4           1.004
## Year              1.026 16           1.001
```

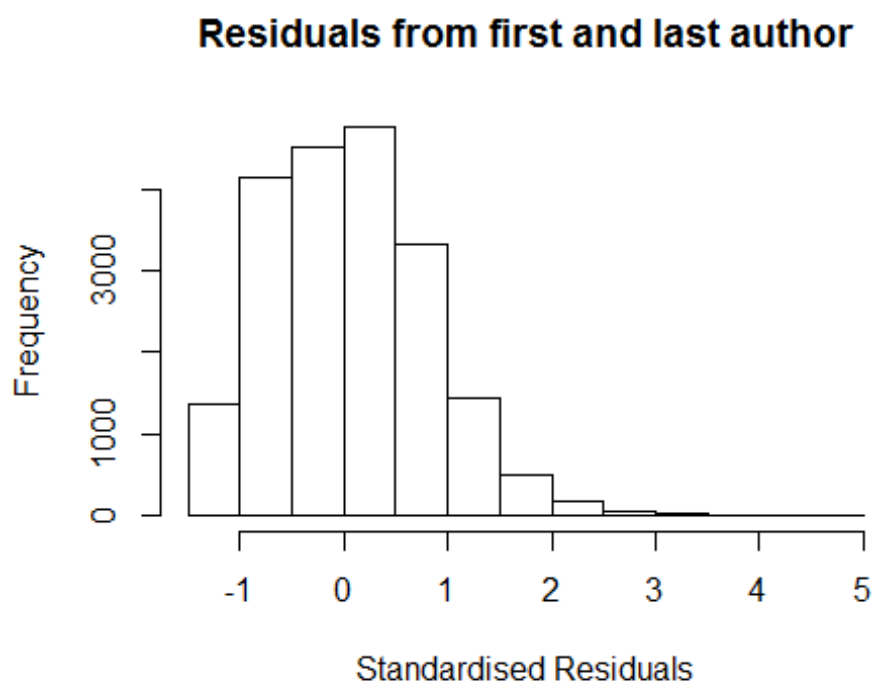



```
## [1] "List of  68 outliers with residuals above  2.5"
##          ScopusId  NLCS Year OneField Fields residuals
## 1478   0001180094 4.152 1998     1200      2     3.213
## 5679  19844382943 3.903 2002     1200      2     2.856
## 5722  61049101473 4.366 2002     1213      2     3.252
## 5807  84998183101 4.984 2002     1213      2     3.870
## 5902   3142779116 4.956 2002     1213      3     3.909
## 5931  60950279634 3.963 2002     1208      3     2.849
## 6067  85012444559 3.746 2002     1202      3     2.699
## 6116   0036526749 3.670 2002     1202      6     2.556
## 6118   1442335800 3.694 2002     1202      6     2.647
## 7137  84997941907 3.951 2003     1213      2     2.876
## 8623  84998075563 5.553 2004     1213      2     4.519
## 9033  33746417589 3.577 2005     1202      2     2.630
## 9122  60950274865 3.944 2005     1208      2     2.997
## 9271  33845250551 4.010 2005     1202      4     3.063
## 9454  63849109891 3.549 2005     1202      3     2.535
## 9719  60950504855 3.650 2005     1208      3     2.636
## 10135 43249177309 3.508 2005     1213      2     2.561
## 10212 84998090320 4.780 2005     1213      2     3.833
## 11243 33646830399 3.808 2006     3316      1     2.825
## 11834 40849109927 3.694 2006     1202      2     2.511
## 11944 84997860330 3.886 2006     1213      2     2.970
## 11950 84998153444 4.441 2006     1213      2     3.525
## 13159 34249284489 3.637 2007     3316      1     2.685
## 13194 43249145818 4.442 2007     1213      2     3.423
## 13756 84997860460 3.627 2007     1213      2     2.608
```

```
## 13760 84997943760 4.345 2007     1213      2     3.132
## 13844 51249156343 3.459 2007     1202      4     2.507
## 13869 61249239835 4.188 2007     1208      3     3.236
## 13880 61249660537 3.741 2007     1208      3     2.789
## 13886 62949099445 3.456 2007     1202      3     2.504
## 14081 60950499484 3.660 2008     1210      2     2.687
## 14222 60950666896 4.798 2008     1213      2     3.825
## 14224 60950735581 3.539 2008     1213      2     2.566
## 14914 70450082589 3.516 2008     1205      3     2.543
## 15103 60949164169 3.720 2008     1208      3     2.814
## 15580 70449355960 4.922 2008     1213      2     3.678
## 15586 77955354528 4.524 2008     1213      2     3.285
## 15607 84997884125 4.500 2008     1213      2     3.527
## 15608 84997901848 5.848 2008     1213      2     4.942
## 15614 84998153436 4.500 2008     1213      2     3.594
## 16005 61949462044 3.777 2009     1208      3     2.947
## 16472 73649148390 3.360 2009     1202      4     2.530
## 17102 67650102337 4.376 2009     1213      2     3.546
## 18554 78649433223 4.435 2010     3316      1     3.590
## 19119 79956372974 3.668 2010     1202      4     2.890
## 19188 77951809379 3.936 2010     1200      2     3.158
## 20438 80155194058 3.906 2011     1202      5     2.998
## 21405 79958705208 4.106 2011     1213      3     3.198
## 21807 79955674091 3.677 2011     1203      6     2.769
## 21958 79952719487 3.448 2011     1213      3     2.607
## 22200 79960456514 3.730 2011     1210      2     2.623
## 22232 80555156002 3.648 2011     1208      2     2.740
## 22587 84871206897 3.953 2012     1208      2     3.142
## 22589 84871229041 3.842 2012     1208      2     3.031
## 22712 84871280398 3.534 2012     3316      2     2.656
## 23171 84867392058 4.065 2012     1213      2     3.254
## 23338 84867101455 3.433 2012     3316      2     2.622
## 23541 84868540587 3.821 2012     1200      2     3.010
## 23545 84868571646 3.628 2012     1200      2     2.750
## 23632 84865713540 3.451 2012     1202      3     2.640
## 23662 84866640267 5.127 2012     1208      3     4.249
## 23833 84868305949 3.480 2012     3312      2     2.669
## 23939 84861970276 3.900 2012     1213      2     3.022
## 24011 84861899557 3.516 2012     1201      3     2.705
## 24361 84859949763 3.495 2012     1213      2     2.617
## 24840 84862115131 3.380 2012     1202      4     2.502
## 24841 84862116557 3.380 2012     1202      4     2.569
## 24924 84856468758 4.331 2012     1202      3     3.453
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
UniqueAuthors +
##     Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##     k.max = 1000))
```

```
##  \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.56710 -0.59095 -0.00241  0.53087  4.94218
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         0.84640    0.02955   28.64  < 2e-16 ***
## FirstAuthorFemale1  0.07144    0.01919    3.72   0.0002 ***
## LastAuthorFemale1  -0.00455    0.01919   -0.24   0.8126
## UniqueAuthors2      0.26636    0.01728   15.41  < 2e-16 ***
## UniqueAuthors3      0.37934    0.02511   15.11  < 2e-16 ***
## UniqueAuthors4      0.38912    0.03492   11.14  < 2e-16 ***
## UniqueAuthors5      0.46586    0.04033   11.55  < 2e-16 ***
## Year1997            0.02408    0.04230    0.57   0.5691
## Year1998            0.09281    0.03996    2.32   0.0202 *
## Year1999            0.05483    0.03931    1.40   0.1630
## Year2000            0.05425    0.03939    1.38   0.1685
## Year2001            0.02697    0.03840    0.70   0.4826
## Year2002            0.20071    0.04061    4.94  7.8e-07 ***
## Year2003            0.16180    0.03902    4.15  3.4e-05 ***
## Year2004            0.18794    0.03845    4.89  1.0e-06 ***
## Year2005            0.10079    0.03720    2.71   0.0068 **
## Year2006            0.06979    0.03558    1.96   0.0498 *
## Year2007            0.10521    0.03593    2.93   0.0034 **
## Year2008            0.05941    0.03568    1.67   0.0959 .
## Year2009           -0.01617    0.03558   -0.45   0.6494
## Year2010           -0.06851    0.03364   -2.04   0.0417 *
## Year2011           -0.00561    0.03383   -0.17   0.8684
## Year2012           -0.03515    0.03417   -1.03   0.3036
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.808
## Multiple R-squared:  0.0341, Adjusted R-squared:  0.033
## Convergence in 10 IRWLS iterations
##
## Robustness weights:
##  7 observations c(4517,4583,6841,8141,11434,12536,18803)
##   are outliers with |weight| = 0 ( < 4.9e-06);
##  1758 weights are ~= 1. The remaining 18505 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0031  0.8860  0.9450  0.9170  0.9860  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          4.93e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
```

```
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 3.363  1           1.834
## LastAuthorFemale  3.366  1           1.835
## Year              1.014 16           1.000
```

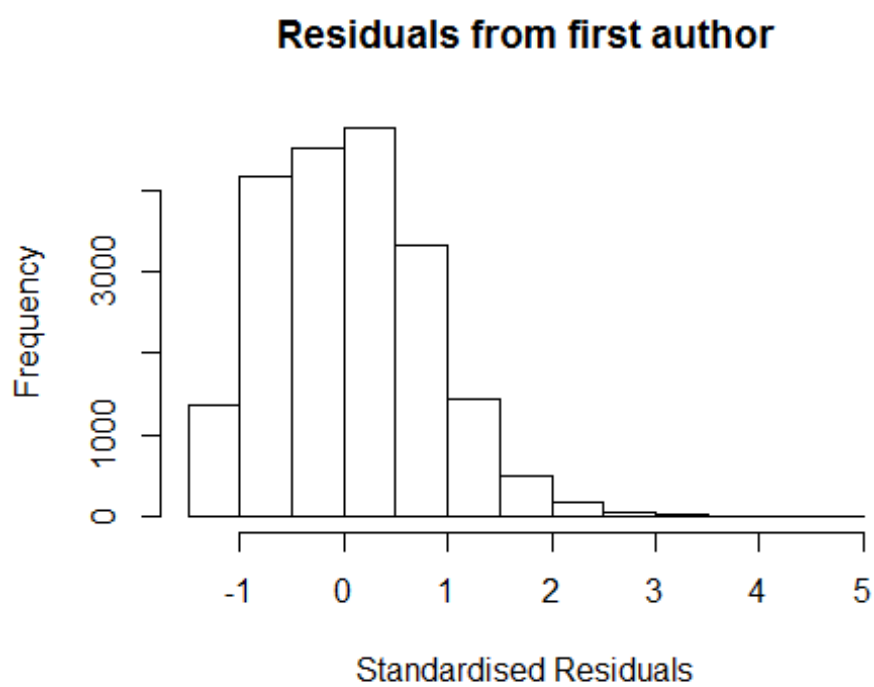



```
## [1] "List of  65 outliers with residuals above  2.5"
##          ScopusId  NLCS Year OneField Fields residuals
## 1478   0001180094 4.152 1998     1200      2     3.177
## 5679  19844382943 3.903 2002     1200      2     2.820
## 5722  61049101473 4.366 2002     1213      2     3.194
## 5807  84998183101 4.984 2002     1213      2     3.812
## 5902   3142779116 4.956 2002     1213      3     3.873
## 5931  60950279634 3.963 2002     1208      3     2.791
## 6067  85012444559 3.746 2002     1202      3     2.663
## 6118   1442335800 3.694 2002     1202      6     2.611
## 7137  84997941907 3.951 2003     1213      2     2.824
```

```
## 8623  84998075563 5.553 2004     1213      2     4.488
## 9033  33746417589 3.577 2005     1202      2     2.599
## 9122  60950274865 3.944 2005     1208      2     2.966
## 9271  33845250551 4.010 2005     1202      4     3.032
## 9719  60950504855 3.650 2005     1208      3     2.583
## 9987  34249157203 3.596 2005     1213      2     2.529
## 10135 43249177309 3.508 2005     1213      2     2.530
## 10212 84998090320 4.780 2005     1213      2     3.802
## 11243 33646830399 3.808 2006     3316      1     2.767
## 11834 40849109927 3.694 2006     1202      2     2.741
## 11944 84997860330 3.886 2006     1213      2     2.933
## 11950 84998153444 4.441 2006     1213      2     3.488
## 13159 34249284489 3.637 2007     3316      1     2.640
## 13194 43249145818 4.442 2007     1213      2     3.357
## 13756 84997860460 3.627 2007     1213      2     2.542
## 13760 84997943760 4.345 2007     1213      2     3.349
## 13869 61249239835 4.188 2007     1208      3     3.191
## 13880 61249660537 3.741 2007     1208      3     2.744
## 14081 60950499484 3.660 2008     1210      2     2.625
## 14222 60950666896 4.798 2008     1213      2     3.763
## 14224 60950735581 3.539 2008     1213      2     2.504
## 15103 60949164169 3.720 2008     1208      3     2.773
## 15580 70449355960 4.922 2008     1213      2     3.886
## 15586 77955354528 4.524 2008     1213      2     3.489
## 15607 84997884125 4.500 2008     1213      2     3.465
## 15608 84997901848 5.848 2008     1213      2     4.901
## 15614 84998153436 4.500 2008     1213      2     3.553
## 16005 61949462044 3.777 2009     1208      3     2.892
## 17102 67650102337 4.376 2009     1213      2     3.491
## 18554 78649433223 4.435 2010     3316      1     3.517
## 19119 79956372974 3.668 2010     1202      4     2.838
## 19188 77951809379 3.936 2010     1200      2     3.106
## 20289 84857935340 3.474 2011     1202      6     2.579
## 20438 80155194058 3.906 2011     1202      5     2.923
## 20708 80053564884 3.595 2011     3316      1     2.612
## 21405 79958705208 4.106 2011     1213      3     3.123
## 21807 79955674091 3.677 2011     1203      6     2.694
## 21958 79952719487 3.448 2011     1213      3     2.553
## 22200 79960456514 3.730 2011     1210      2     2.835
## 22232 80555156002 3.648 2011     1208      2     2.665
## 22587 84871206897 3.953 2012     1208      2     3.086
## 22589 84871229041 3.842 2012     1208      2     2.975
## 22712 84871280398 3.534 2012     3316      2     2.579
## 23171 84867392058 4.065 2012     1213      2     3.198
## 23338 84867101455 3.433 2012     3316      2     2.566
## 23541 84868540587 3.821 2012     1200      2     2.954
## 23545 84868571646 3.628 2012     1200      2     2.673
## 23632 84865713540 3.451 2012     1202      3     2.584
## 23662 84866640267 5.127 2012     1208      3     4.172
## 23833 84868305949 3.480 2012     3312      2     2.613
```

```
## 23939 84861970276 3.900 2012     1213      2     2.945
## 24011 84861899557 3.516 2012     1201      3     2.649
## 24361 84859949763 3.495 2012     1213      2     2.540
## 24841 84862116557 3.380 2012     1202      4     2.513
## 24924 84856468758 4.331 2012     1202      3     3.376
## 25090 84857345730 3.401 2012     3312      2     2.534
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.17267 -0.56437  0.00553  0.54851  4.90104
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)         0.893761   0.029974   29.82  < 2e-16 ***
## FirstAuthorFemale1  0.089299   0.020770    4.30  1.7e-05 ***
## LastAuthorFemale1  -0.001080   0.020789   -0.05  0.95859
## Year1997            0.028565   0.042727    0.67  0.50380
## Year1998            0.081705   0.040349    2.02  0.04289 *
## Year1999            0.057326   0.040077    1.43  0.15262
## Year2000            0.067097   0.039814    1.69  0.09195 .
## Year2001            0.039474   0.039006    1.01  0.31156
## Year2002            0.189609   0.041091    4.61  4.0e-06 ***
## Year2003            0.145118   0.039522    3.67  0.00024 ***
## Year2004            0.171475   0.038918    4.41  1.1e-05 ***
## Year2005            0.084558   0.037815    2.24  0.02536 *
## Year2006            0.058872   0.036213    1.63  0.10402
## Year2007            0.103425   0.036431    2.84  0.00453 **
## Year2008            0.053198   0.036343    1.46  0.14327
## Year2009           -0.008674   0.036312   -0.24  0.81120
## Year2010           -0.064021   0.034339   -1.86  0.06228 .
## Year2011            0.000785   0.034394    0.02  0.98178
## Year2012           -0.026758   0.034811   -0.77  0.44211
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.829
## Multiple R-squared:  0.0108, Adjusted R-squared:  0.00993
## Convergence in 10 IRWLS iterations
##
## Robustness weights:
##  4 observations c(6841,12512,12536,18803)
##   are outliers with |weight| = 0 ( < 4.9e-06);
##  1698 weights are ~= 1. The remaining 18568 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.000   0.883   0.949   0.919   0.986   0.999
```

```
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          4.93e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##                   GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.01  1           1.005
## Year              1.01 16           1.000
```

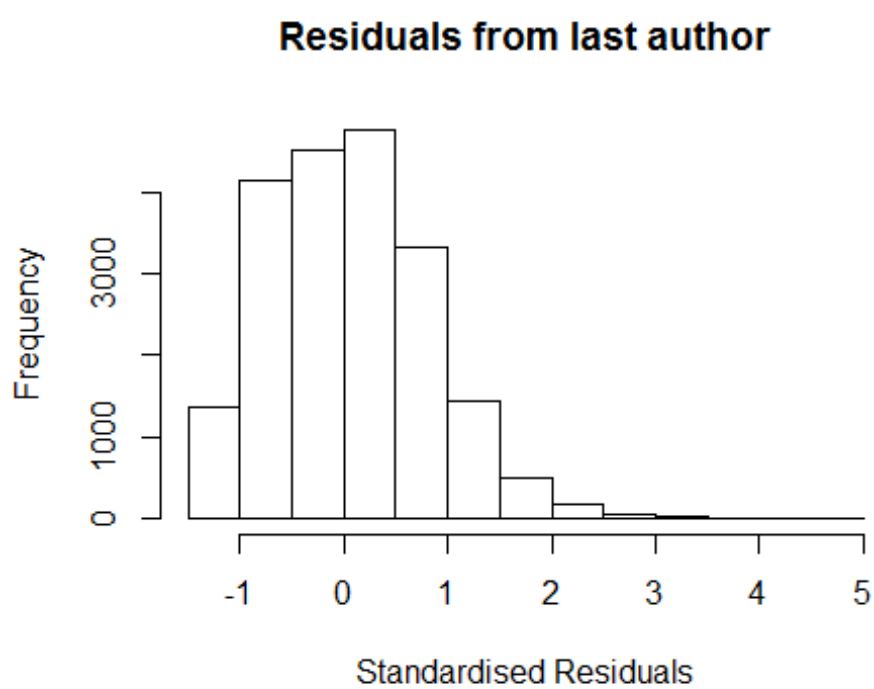



```
## [1] "List of  65 outliers with residuals above  2.5"
##          ScopusId  NLCS Year OneField Fields residuals
## 1478   0001180094 4.152 1998     1200      2     3.177
## 5679  19844382943 3.903 2002     1200      2     2.820
## 5722  61049101473 4.366 2002     1213      2     3.194
## 5807  84998183101 4.984 2002     1213      2     3.812
```

```
## 5902   3142779116 4.956 2002     1213      3     3.873
## 5931  60950279634 3.963 2002     1208      3     2.791
## 6067  85012444559 3.746 2002     1202      3     2.663
## 6118   1442335800 3.694 2002     1202      6     2.611
## 7137  84997941907 3.951 2003     1213      2     2.824
## 8623  84998075563 5.553 2004     1213      2     4.488
## 9033  33746417589 3.577 2005     1202      2     2.599
## 9122  60950274865 3.944 2005     1208      2     2.966
## 9271  33845250551 4.010 2005     1202      4     3.032
## 9719  60950504855 3.650 2005     1208      3     2.583
## 9987  34249157203 3.596 2005     1213      2     2.529
## 10135 43249177309 3.508 2005     1213      2     2.530
## 10212 84998090320 4.780 2005     1213      2     3.802
## 11243 33646830399 3.808 2006     3316      1     2.767
## 11834 40849109927 3.694 2006     1202      2     2.741
## 11944 84997860330 3.886 2006     1213      2     2.933
## 11950 84998153444 4.441 2006     1213      2     3.488
## 13159 34249284489 3.637 2007     3316      1     2.640
## 13194 43249145818 4.442 2007     1213      2     3.357
## 13756 84997860460 3.627 2007     1213      2     2.542
## 13760 84997943760 4.345 2007     1213      2     3.349
## 13869 61249239835 4.188 2007     1208      3     3.191
## 13880 61249660537 3.741 2007     1208      3     2.744
## 14081 60950499484 3.660 2008     1210      2     2.625
## 14222 60950666896 4.798 2008     1213      2     3.763
## 14224 60950735581 3.539 2008     1213      2     2.504
## 15103 60949164169 3.720 2008     1208      3     2.773
## 15580 70449355960 4.922 2008     1213      2     3.886
## 15586 77955354528 4.524 2008     1213      2     3.489
## 15607 84997884125 4.500 2008     1213      2     3.465
## 15608 84997901848 5.848 2008     1213      2     4.901
## 15614 84998153436 4.500 2008     1213      2     3.553
## 16005 61949462044 3.777 2009     1208      3     2.892
## 17102 67650102337 4.376 2009     1213      2     3.491
## 18554 78649433223 4.435 2010     3316      1     3.517
## 19119 79956372974 3.668 2010     1202      4     2.838
## 19188 77951809379 3.936 2010     1200      2     3.106
## 20289 84857935340 3.474 2011     1202      6     2.579
## 20438 80155194058 3.906 2011     1202      5     2.923
## 20708 80053564884 3.595 2011     3316      1     2.612
## 21405 79958705208 4.106 2011     1213      3     3.123
## 21807 79955674091 3.677 2011     1203      6     2.694
## 21958 79952719487 3.448 2011     1213      3     2.553
## 22200 79960456514 3.730 2011     1210      2     2.835
## 22232 80555156002 3.648 2011     1208      2     2.665
## 22587 84871206897 3.953 2012     1208      2     3.086
## 22589 84871229041 3.842 2012     1208      2     2.975
## 22712 84871280398 3.534 2012     3316      2     2.579
## 23171 84867392058 4.065 2012     1213      2     3.198
## 23338 84867101455 3.433 2012     3316      2     2.566
```

```
## 23541 84868540587 3.821 2012     1200      2     2.954
## 23545 84868571646 3.628 2012     1200      2     2.673
## 23632 84865713540 3.451 2012     1202      3     2.584
## 23662 84866640267 5.127 2012     1208      3     4.172
## 23833 84868305949 3.480 2012     3312      2     2.613
## 23939 84861970276 3.900 2012     1213      2     2.945
## 24011 84861899557 3.516 2012     1201      3     2.649
## 24361 84859949763 3.495 2012     1213      2     2.540
## 24841 84862116557 3.380 2012     1202      4     2.513
## 24924 84856468758 4.331 2012     1202      3     3.376
## 25090 84857345730 3.401 2012     3312      2     2.534
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.17169 -0.56472  0.00568  0.54833  4.90112
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)         0.893714   0.029949   29.84  < 2e-16 ***
## FirstAuthorFemale1  0.088375   0.011391    7.76  9.0e-15 ***
## Year1997            0.028560   0.042728    0.67  0.50387
## Year1998            0.081695   0.040351    2.02  0.04292 *
## Year1999            0.057328   0.040078    1.43  0.15261
## Year2000            0.067106   0.039811    1.69  0.09189 .
## Year2001            0.039461   0.039010    1.01  0.31176
## Year2002            0.189597   0.041093    4.61  4.0e-06 ***
## Year2003            0.145108   0.039524    3.67  0.00024 ***
## Year2004            0.171454   0.038917    4.41  1.1e-05 ***
## Year2005            0.084539   0.037816    2.24  0.02539 *
## Year2006            0.058845   0.036216    1.62  0.10422
## Year2007            0.103400   0.036427    2.84  0.00454 **
## Year2008            0.053168   0.036341    1.46  0.14347
## Year2009           -0.008706   0.036314   -0.24  0.81052
## Year2010           -0.064048   0.034339   -1.87  0.06217 .
## Year2011            0.000763   0.034394    0.02  0.98230
## Year2012           -0.026770   0.034815   -0.77  0.44194
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.829
## Multiple R-squared:  0.0108, Adjusted R-squared:  0.00998
## Convergence in 10 IRWLS iterations
##
## Robustness weights:
##  4 observations c(6841,12512,12536,18803)
```

```
##   are outliers with |weight| = 0 ( < 4.9e-06);
##  1697 weights are ~= 1. The remaining 18569 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.000   0.883   0.949   0.919   0.986   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          4.93e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##                   GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.011  1           1.006
## Year             1.011 16           1.000
```



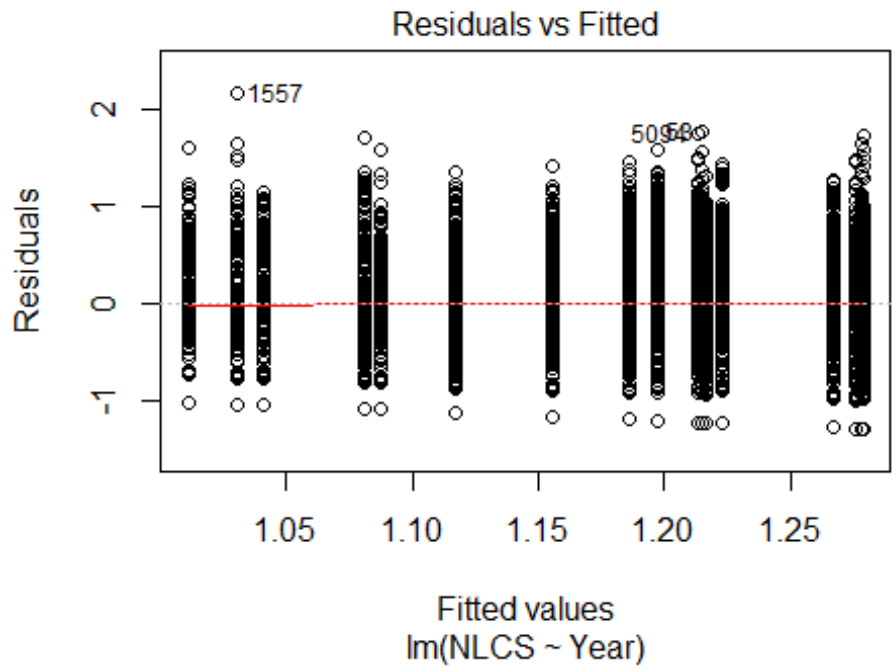

```
## [1] "List of  65 outliers with residuals above  2.5"
##          ScopusId  NLCS Year OneField Fields residuals
```

```
## 1478   0001180094 4.152 1998     1200      2     3.177
## 5679  19844382943 3.903 2002     1200      2     2.820
## 5722  61049101473 4.366 2002     1213      2     3.194
## 5807  84998183101 4.984 2002     1213      2     3.812
## 5902   3142779116 4.956 2002     1213      3     3.873
## 5931  60950279634 3.963 2002     1208      3     2.791
## 6067  85012444559 3.746 2002     1202      3     2.663
## 6118   1442335800 3.694 2002     1202      6     2.611
## 7137  84997941907 3.951 2003     1213      2     2.824
## 8623  84998075563 5.553 2004     1213      2     4.488
## 9033  33746417589 3.577 2005     1202      2     2.599
## 9122  60950274865 3.944 2005     1208      2     2.966
## 9271  33845250551 4.010 2005     1202      4     3.032
## 9719  60950504855 3.650 2005     1208      3     2.583
## 9987  34249157203 3.596 2005     1213      2     2.529
## 10135 43249177309 3.508 2005     1213      2     2.530
## 10212 84998090320 4.780 2005     1213      2     3.802
## 11243 33646830399 3.808 2006     3316      1     2.767
## 11834 40849109927 3.694 2006     1202      2     2.741
## 11944 84997860330 3.886 2006     1213      2     2.933
## 11950 84998153444 4.441 2006     1213      2     3.488
## 13159 34249284489 3.637 2007     3316      1     2.640
## 13194 43249145818 4.442 2007     1213      2     3.357
## 13756 84997860460 3.627 2007     1213      2     2.542
## 13760 84997943760 4.345 2007     1213      2     3.349
## 13869 61249239835 4.188 2007     1208      3     3.191
## 13880 61249660537 3.741 2007     1208      3     2.744
## 14081 60950499484 3.660 2008     1210      2     2.625
## 14222 60950666896 4.798 2008     1213      2     3.763
## 14224 60950735581 3.539 2008     1213      2     2.504
## 15103 60949164169 3.720 2008     1208      3     2.773
## 15580 70449355960 4.922 2008     1213      2     3.886
## 15586 77955354528 4.524 2008     1213      2     3.489
## 15607 84997884125 4.500 2008     1213      2     3.465
## 15608 84997901848 5.848 2008     1213      2     4.901
## 15614 84998153436 4.500 2008     1213      2     3.553
## 16005 61949462044 3.777 2009     1208      3     2.892
## 17102 67650102337 4.376 2009     1213      2     3.491
## 18554 78649433223 4.435 2010     3316      1     3.517
## 19119 79956372974 3.668 2010     1202      4     2.838
## 19188 77951809379 3.936 2010     1200      2     3.106
## 20289 84857935340 3.474 2011     1202      6     2.579
## 20438 80155194058 3.906 2011     1202      5     2.923
## 20708 80053564884 3.595 2011     3316      1     2.612
## 21405 79958705208 4.106 2011     1213      3     3.123
## 21807 79955674091 3.677 2011     1203      6     2.694
## 21958 79952719487 3.448 2011     1213      3     2.553
## 22200 79960456514 3.730 2011     1210      2     2.835
## 22232 80555156002 3.648 2011     1208      2     2.665
## 22587 84871206897 3.953 2012     1208      2     3.086
```

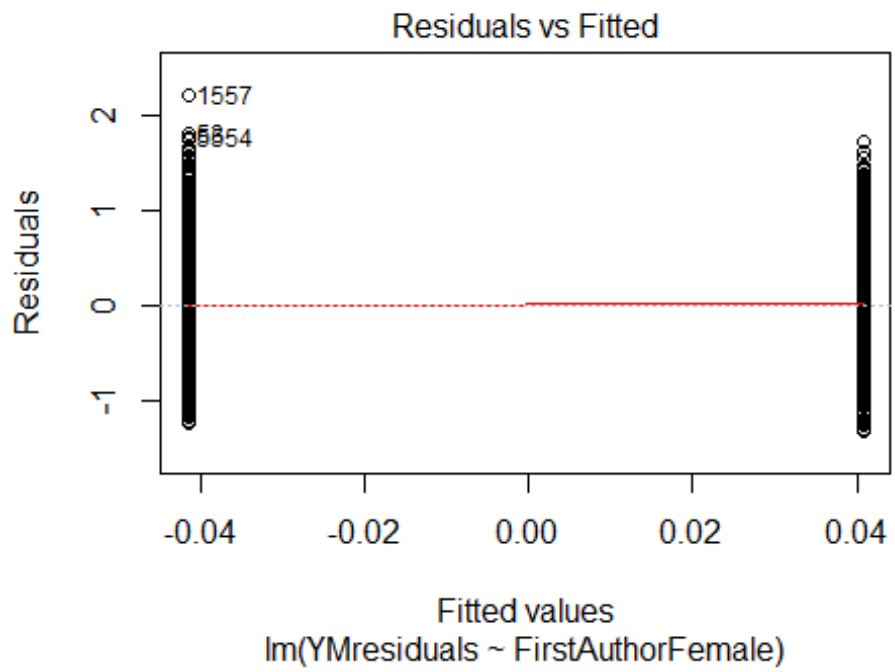
```
## 22589 84871229041 3.842 2012     1208      2     2.975
## 22712 84871280398 3.534 2012     3316      2     2.579
## 23171 84867392058 4.065 2012     1213      2     3.198
## 23338 84867101455 3.433 2012     3316      2     2.566
## 23541 84868540587 3.821 2012     1200      2     2.954
## 23545 84868571646 3.628 2012     1200      2     2.673
## 23632 84865713540 3.451 2012     1202      3     2.584
## 23662 84866640267 5.127 2012     1208      3     4.172
## 23833 84868305949 3.480 2012     3312      2     2.613
## 23939 84861970276 3.900 2012     1213      2     2.945
## 24011 84861899557 3.516 2012     1201      3     2.649
## 24361 84859949763 3.495 2012     1213      2     2.540
## 24841 84862116557 3.380 2012     1202      4     2.513
## 24924 84856468758 4.331 2012     1202      3     3.376
## 25090 84857345730 3.401 2012     3312      2     2.534
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.16426 -0.56406  0.00908  0.54751  4.89614
##
## Coefficients:
##                   Estimate Std. Error t value Pr(>|t|)
## (Intercept)        0.89930    0.03002   29.95  < 2e-16 ***
## LastAuthorFemale1  0.07542    0.01139    6.62  3.7e-11 ***
## Year1997           0.02850    0.04271    0.67  0.50467
## Year1998           0.08096    0.04043    2.00  0.04527 *
## Year1999           0.05664    0.04013    1.41  0.15813
## Year2000           0.06758    0.03988    1.69  0.09016 .
## Year2001           0.03837    0.03906    0.98  0.32589
## Year2002           0.18953    0.04113    4.61  4.1e-06 ***
## Year2003           0.14581    0.03961    3.68  0.00023 ***
## Year2004           0.17080    0.03898    4.38  1.2e-05 ***
## Year2005           0.08448    0.03789    2.23  0.02579 *
## Year2006           0.05867    0.03627    1.62  0.10575
## Year2007           0.10260    0.03650    2.81  0.00494 **
## Year2008           0.05255    0.03642    1.44  0.14909
## Year2009          -0.00881    0.03636   -0.24  0.80850
## Year2010          -0.06358    0.03440   -1.85  0.06461 .
## Year2011           0.00146    0.03448    0.04  0.96623
## Year2012          -0.02535    0.03488   -0.73  0.46727
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.83
## Multiple R-squared:  0.00997,    Adjusted R-squared:  0.00914
```

```
## Convergence in 10 IRWLS iterations
##
## Robustness weights:
##  4 observations c(6841,12512,12536,18803)
##   are outliers with |weight| = 0 ( < 4.9e-06);
##  1703 weights are ~= 1. The remaining 18563 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0001  0.8820  0.9490  0.9190  0.9860  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          4.93e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Sample size for the above analysis:  20270"
## [1] ""
## [1] ""
## [1] "##################################"
## [1] "Analysis of AJSC 3317"
## [1] "##################################"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  292  292  306  289  281  305  356  337  351  301  341  381  387  375  443
## 2011 2012
##  474  423
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  241  179  183  166  162  162  316  286  293  248  287  324  340  331  365
## 2011 2012
##  391  359
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  221  163  173  154  151  147  288  274  269  228  264  288  308  303  334
## 2011 2012
##  362  327
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
```

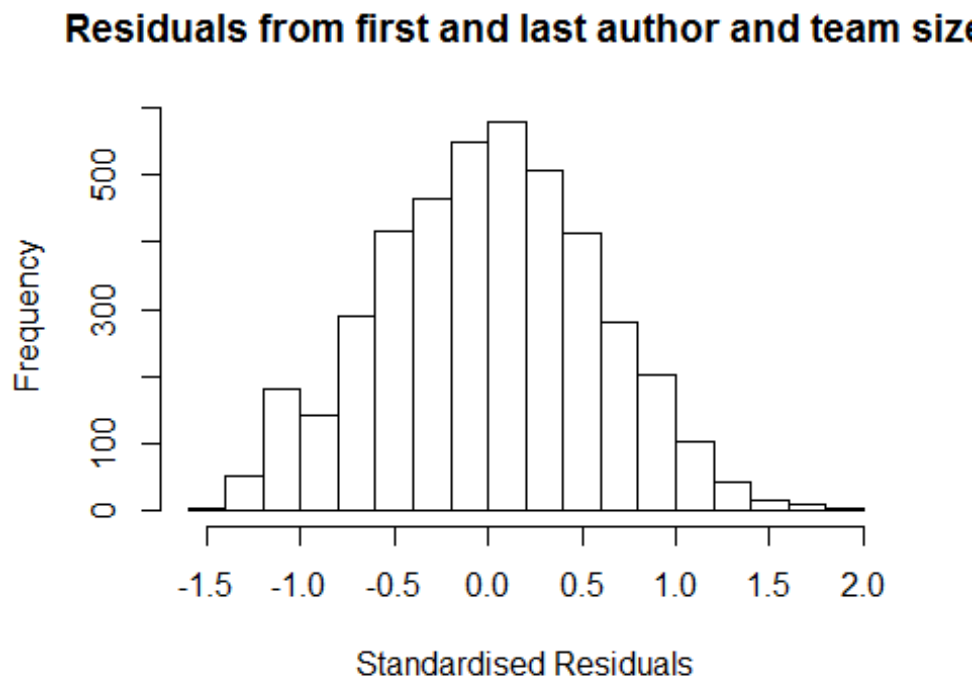
```
##  Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 24, df = 16, p-value = 0.09
```




```
##
##  Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 8, df = 1, p-value = 0.005
```

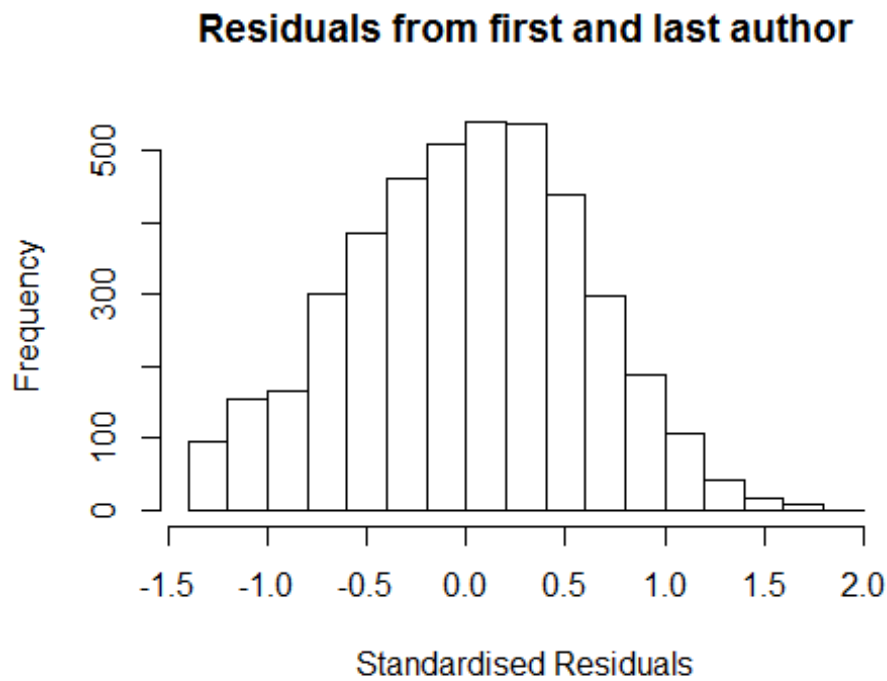



```
## [1] "Female first author team size 2018 geometric mean: 2.05327506416767"
## [1] "Male first author team size 2018 geometric mean: 1.94331668505698"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 18000, p-value = 0.4
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 2.00262612445469"
## [1] "Male last author team size 2018 geometric mean: 2.01074590814744"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 17000, p-value = 0.9
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.411  1           1.188
## LastAuthorFemale  1.407  1           1.186
## UniqueAuthors     1.125  4           1.015
## Year              1.136 16           1.004
```



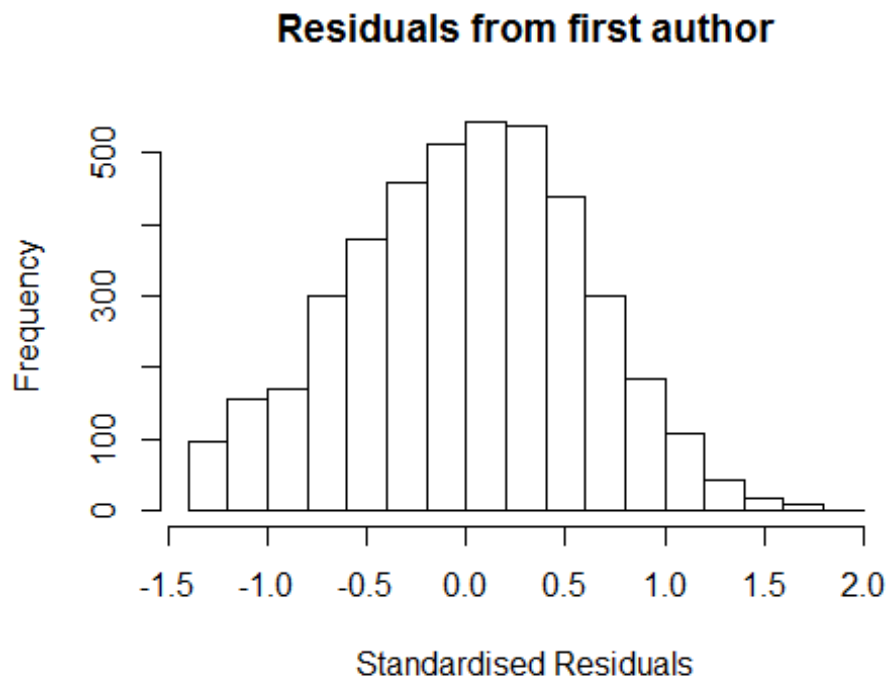

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
UniqueAuthors +
##     Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##     k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.4433 -0.4098  0.0109  0.4038  1.9432
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.04081    0.04262   24.42  < 2e-16 ***
## FirstAuthorFemale1  0.06609    0.02203    3.00   0.0027 **
## LastAuthorFemale1   0.01102    0.02189    0.50   0.6147
## UniqueAuthors2      0.17725    0.02302    7.70  1.7e-14 ***
## UniqueAuthors3      0.25428    0.02643    9.62  < 2e-16 ***
## UniqueAuthors4      0.25611    0.03648    7.02  2.6e-12 ***
## UniqueAuthors5      0.34818    0.03776    9.22  < 2e-16 ***
## Year1997           -0.09514    0.06698   -1.42   0.1555
## Year1998           -0.12376    0.06275   -1.97   0.0486 *
## Year1999           -0.11813    0.06137   -1.93   0.0543 .
```

```
## Year2000           -0.21168    0.06650   -3.18   0.0015 **
## Year2001           -0.15423    0.06677   -2.31   0.0210 *
## Year2002           -0.03795    0.05122   -0.74   0.4588
## Year2003           -0.06228    0.05622   -1.11   0.2680
## Year2004           -0.01102    0.05534   -0.20   0.8422
## Year2005            0.06930    0.05643    1.23   0.2195
## Year2006            0.00403    0.05364    0.08   0.9402
## Year2007           -0.01132    0.05634   -0.20   0.8408
## Year2008            0.05180    0.05239    0.99   0.3228
## Year2009           -0.01203    0.05414   -0.22   0.8242
## Year2010            0.04809    0.04973    0.97   0.3335
## Year2011           -0.00209    0.05167   -0.04   0.9677
## Year2012            0.05431    0.05159    1.05   0.2925
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.598
## Multiple R-squared:  0.0621, Adjusted R-squared:  0.0573
## Convergence in 10 IRWLS iterations
##
## Robustness weights:
##  364 weights are ~= 1. The remaining 3890 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.269   0.875   0.950   0.912   0.985   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          2.35e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.414  1           1.189
## LastAuthorFemale  1.412  1           1.188
## Year              1.030 16           1.001
```



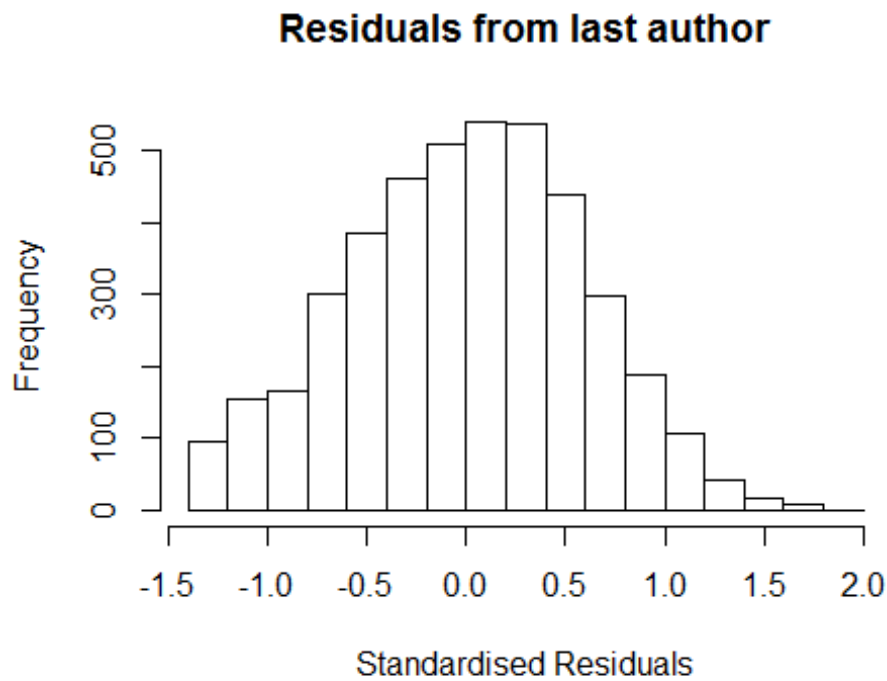

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.3348 -0.4154  0.0221  0.4159  1.8379
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)          1.14615    0.04349   26.35   <2e-16 ***
## FirstAuthorFemale1   0.08748    0.02251    3.89   0.0001 ***
## LastAuthorFemale1    0.01279    0.02243    0.57   0.5685
## Year1997            -0.13333    0.06857   -1.94   0.0519 .
## Year1998            -0.16091    0.06300   -2.55   0.0107 *
## Year1999            -0.13120    0.06223   -2.11   0.0351 *
## Year2000            -0.22197    0.06810   -3.26   0.0011 **
## Year2001            -0.15982    0.06814   -2.35   0.0190 *
## Year2002            -0.04897    0.05368   -0.91   0.3617
## Year2003            -0.07187    0.05862   -1.23   0.2203
## Year2004            -0.02678    0.05755   -0.47   0.6417
## Year2005             0.08834    0.05807    1.52   0.1283
```

```
## Year2006           -0.00149    0.05646   -0.03   0.9790
## Year2007           -0.01746    0.05867   -0.30   0.7660
## Year2008            0.07012    0.05400    1.30   0.1942
## Year2009            0.01924    0.05671    0.34   0.7345
## Year2010            0.07527    0.05185    1.45   0.1466
## Year2011            0.02359    0.05365    0.44   0.6602
## Year2012            0.07870    0.05392    1.46   0.1445
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.61
## Multiple R-squared:  0.0259, Adjusted R-squared:  0.0217
## Convergence in 10 IRWLS iterations
##
## Robustness weights:
##  348 weights are ~= 1. The remaining 3906 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.344   0.877   0.950   0.913   0.985   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          2.35e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.019  1           1.009
## Year              1.019 16           1.001
```




```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.3318 -0.4146  0.0204  0.4139  1.8351
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)          1.14892    0.04324   26.57  < 2e-16 ***
## FirstAuthorFemale1  0.09466    0.01912    4.95  7.7e-07 ***
## Year1997           -0.13348    0.06859   -1.95   0.0517 .
## Year1998           -0.16142    0.06304   -2.56   0.0105 *
## Year1999           -0.13208    0.06217   -2.12   0.0337 *
## Year2000           -0.22173    0.06804   -3.26   0.0011 **
## Year2001           -0.15933    0.06814   -2.34   0.0194 *
## Year2002           -0.04917    0.05372   -0.92   0.3601
## Year2003           -0.07220    0.05864   -1.23   0.2183
## Year2004           -0.02723    0.05755   -0.47   0.6361
## Year2005            0.08818    0.05812    1.52   0.1293
## Year2006           -0.00137    0.05649   -0.02   0.9807
```

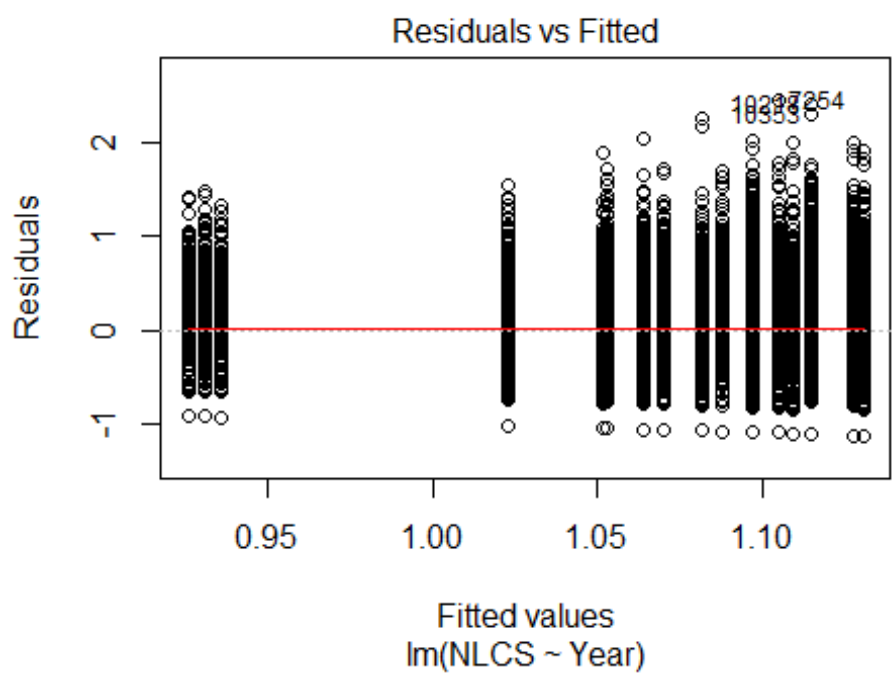
```
## Year2007           -0.01754    0.05872   -0.30   0.7651
## Year2008            0.07002    0.05404    1.30   0.1952
## Year2009            0.01913    0.05673    0.34   0.7360
## Year2010            0.07537    0.05190    1.45   0.1466
## Year2011            0.02341    0.05370    0.44   0.6629
## Year2012            0.07821    0.05398    1.45   0.1474
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.61
## Multiple R-squared:  0.0258, Adjusted R-squared:  0.0219
## Convergence in 10 IRWLS iterations
##
## Robustness weights:
##  360 weights are ~= 1. The remaining 3894 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.345   0.876   0.950   0.912   0.984   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          2.35e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##                   GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.018  1           1.009
## Year             1.018 16           1.001
```



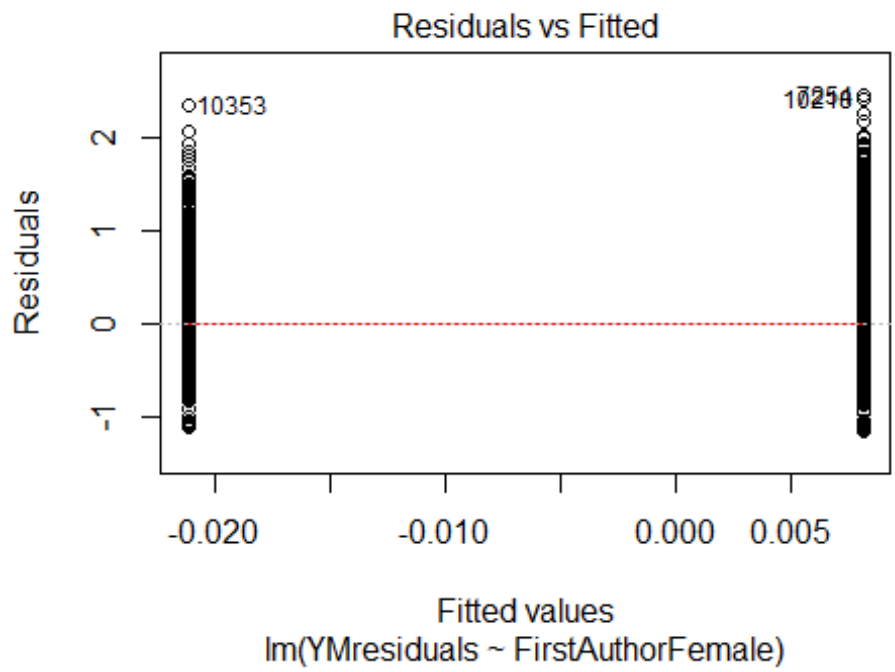

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.3144 -0.4168  0.0172  0.4176  1.8239
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.16009    0.04324   26.83   <2e-16 ***
## LastAuthorFemale1   0.06166    0.01909    3.23   0.0012 **
## Year1997           -0.13086    0.06833   -1.92   0.0556 .
## Year1998           -0.15542    0.06322   -2.46   0.0140 *
## Year1999           -0.12320    0.06236   -1.98   0.0483 *
## Year2000           -0.21700    0.06889   -3.15   0.0016 **
## Year2001           -0.15643    0.06794   -2.30   0.0214 *
## Year2002           -0.04531    0.05355   -0.85   0.3976
## Year2003           -0.06635    0.05864   -1.13   0.2579
## Year2004           -0.01700    0.05756   -0.30   0.7677
## Year2005            0.09266    0.05804    1.60   0.1104
## Year2006            0.00632    0.05642    0.11   0.9109
```

```
## Year2007          -0.00837    0.05862   -0.14   0.8865
## Year2008           0.07784    0.05403    1.44   0.1497
## Year2009           0.02873    0.05681    0.51   0.6131
## Year2010           0.08261    0.05170    1.60   0.1102
## Year2011           0.03308    0.05359    0.62   0.5371
## Year2012           0.09005    0.05379    1.67   0.0942 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.61
## Multiple R-squared:  0.0224, Adjusted R-squared:  0.0185
## Convergence in 10 IRWLS iterations
##
## Robustness weights:
##  351 weights are ~= 1. The remaining 3903 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.352   0.879   0.950   0.912   0.984   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          2.35e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Sample size for the above analysis:  4254"
## [1] ""
## [1] ""
## [1] "####################################"
## [1] "Analysis of AJSC 3318"
## [1] "####################################"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  409  313  345  362  476  479  520  461  472  599  684  759  782  859  948
## 2011 2012
##  985  952
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  359  269  310  320  393  393  451  410  404  525  603  669  686  735  827
## 2011 2012
```

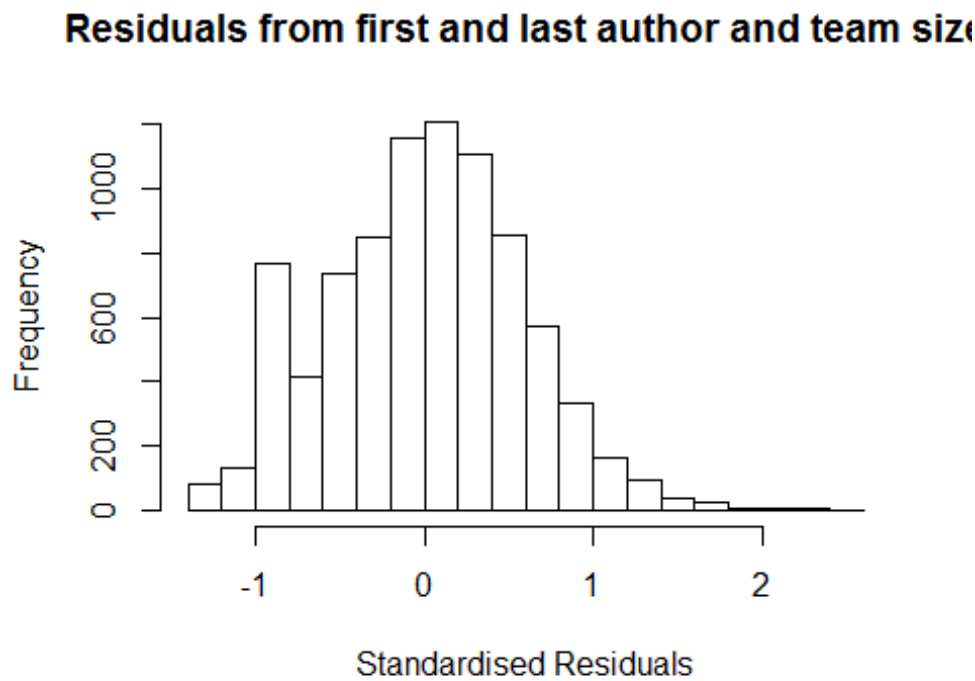
```
##  841  827
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  346  264  300  310  374  383  427  392  384  498  574  639  639  688  784
## 2011 2012
##  791  772
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
##  Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 50, df = 16, p-value = 2e-05
```




```
##
##  Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 0.018, df = 1, p-value = 0.9
```

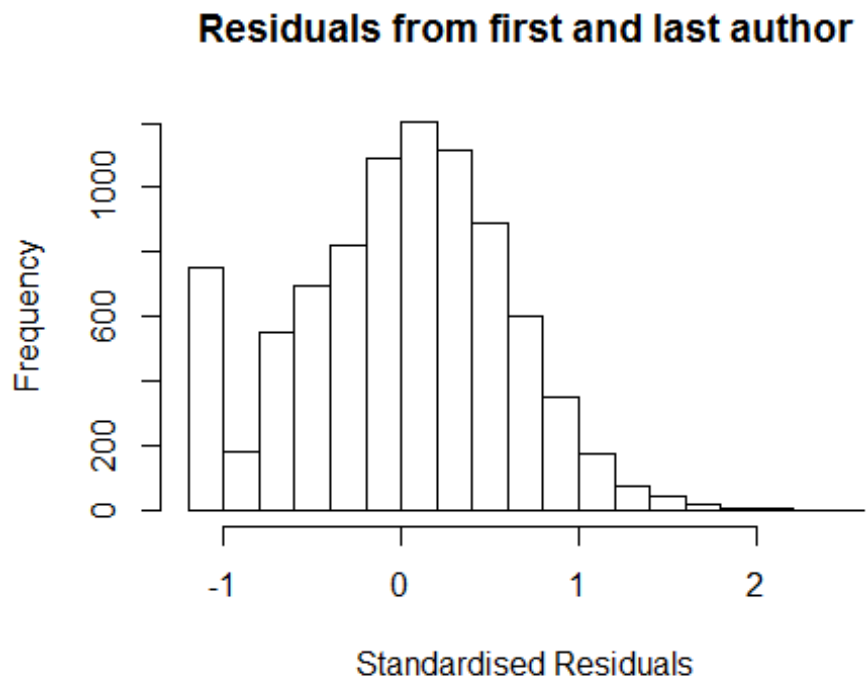



```
## [1] "Female first author team size 2018 geometric mean: 1.86779180774448"
## [1] "Male first author team size 2018 geometric mean: 1.91478159523238"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 69000, p-value = 0.6
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 1.79038369291272"
## [1] "Male last author team size 2018 geometric mean: 2.09124982652509"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 66000, p-value = 4e-04
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##                     GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.387  1           1.178
## LastAuthorFemale  1.404  1           1.185
## UniqueAuthors     1.077  4           1.009
## Year              1.067 16           1.002
```



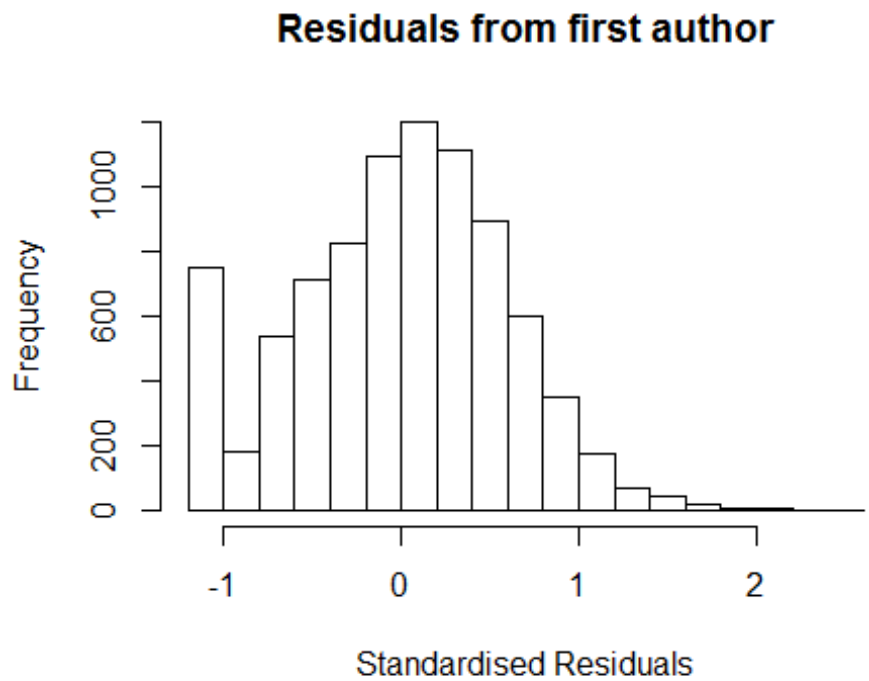

```
## [1] "List of  2 outliers with residuals above  2.5"
##          ScopusId  NLCS Year OneField Fields residuals
## 10218 84871280398 3.534 2012     3316      2     2.562
## 10353 84867101455 3.433 2012     3316      2     2.534
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
UniqueAuthors +
##     Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##     k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.3497 -0.3960  0.0259  0.3893  2.5624
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         0.93345    0.03823   24.41   <2e-16 ***
## FirstAuthorFemale1  0.01931    0.01700    1.14   0.2560
## LastAuthorFemale1   0.05279    0.01680    3.14   0.0017 **
## UniqueAuthors2      0.25429    0.01567   16.22   <2e-16 ***
## UniqueAuthors3      0.29901    0.01956   15.28   <2e-16 ***
## UniqueAuthors4      0.36800    0.02513   14.64   <2e-16 ***
## UniqueAuthors5      0.34202    0.02885   11.86   <2e-16 ***
## Year1997           -0.14778    0.05137   -2.88   0.0040 **
## Year1998           -0.15702    0.05090   -3.08   0.0020 **
```

```
## Year1999           -0.15021    0.05030   -2.99   0.0028 **
## Year2000           -0.05323    0.04685   -1.14   0.2559
## Year2001           -0.06533    0.04653   -1.40   0.1603
## Year2002           -0.03985    0.04654   -0.86   0.3918
## Year2003           -0.04542    0.04668   -0.97   0.3306
## Year2004           -0.01661    0.04696   -0.35   0.7236
## Year2005           -0.02038    0.04466   -0.46   0.6481
## Year2006           -0.02412    0.04290   -0.56   0.5740
## Year2007           -0.06543    0.04234   -1.55   0.1223
## Year2008           -0.02386    0.04199   -0.57   0.5699
## Year2009           -0.01730    0.04258   -0.41   0.6845
## Year2010           -0.03541    0.04269   -0.83   0.4069
## Year2011           -0.00524    0.04247   -0.12   0.9018
## Year2012           -0.03399    0.04356   -0.78   0.4352
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.579
## Multiple R-squared:  0.0659, Adjusted R-squared:  0.0635
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  711 weights are ~= 1. The remaining 7854 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0113  0.8650  0.9490  0.9070  0.9860  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.17e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                    psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.408  1           1.187
## LastAuthorFemale  1.411  1           1.188
## Year              1.012 16           1.000
```



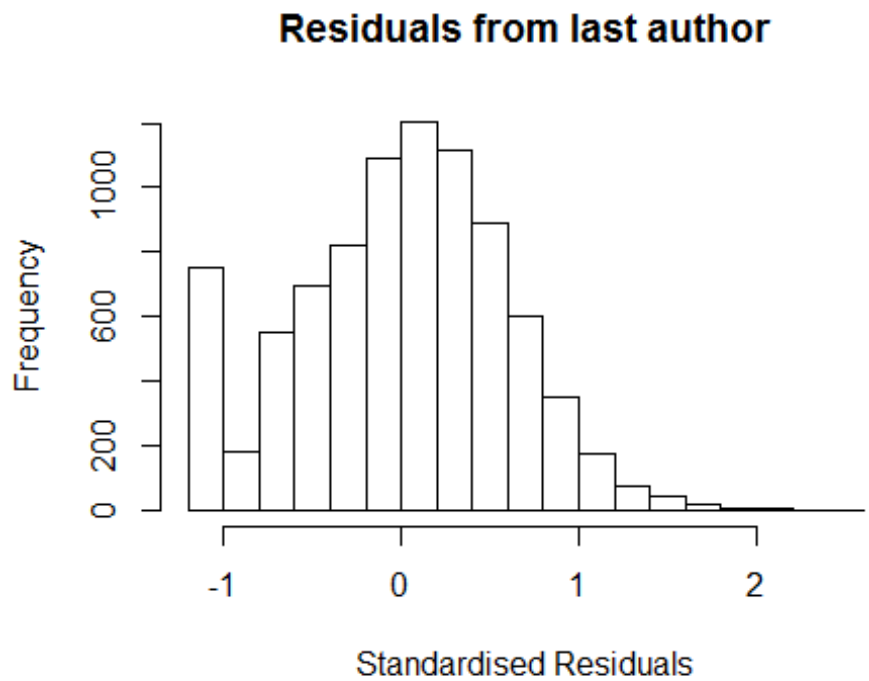

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.1354 -0.4069  0.0297  0.4040  2.4736
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)          1.04629    0.03933   26.60   <2e-16 ***
## FirstAuthorFemale1   0.02012    0.01777    1.13   0.2575
## LastAuthorFemale1    0.01747    0.01740    1.00   0.3155
## Year1997            -0.14492    0.05288   -2.74   0.0061 **
## Year1998            -0.14737    0.05173   -2.85   0.0044 **
## Year1999            -0.14590    0.05226   -2.79   0.0053 **
## Year2000            -0.02315    0.04825   -0.48   0.6313
## Year2001            -0.04561    0.04846   -0.94   0.3466
## Year2002            -0.00527    0.04817   -0.11   0.9129
## Year2003            -0.01439    0.04819   -0.30   0.7653
## Year2004             0.02287    0.04787    0.48   0.6328
## Year2005             0.02763    0.04585    0.60   0.5467
```

```
## Year2006            0.00751    0.04484    0.17   0.8671
## Year2007           -0.01357    0.04423   -0.31   0.7591
## Year2008            0.02403    0.04367    0.55   0.5822
## Year2009            0.04848    0.04383    1.11   0.2687
## Year2010            0.02284    0.04459    0.51   0.6085
## Year2011            0.05151    0.04420    1.17   0.2439
## Year2012            0.03846    0.04519    0.85   0.3948
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.593
## Multiple R-squared:  0.009,  Adjusted R-squared:  0.00691
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
##  717 weights are ~= 1. The remaining 7848 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0427  0.8640  0.9490  0.9060  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.17e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.007  1           1.003
## Year              1.007 16           1.000
```




```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##    Min     1Q Median     3Q    Max
## -1.133 -0.407  0.030  0.405  2.458
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)          1.05023    0.03908   26.88   <2e-16 ***
## FirstAuthorFemale1  0.03093    0.01510    2.05   0.0406 *
## Year1997           -0.14476    0.05292   -2.74   0.0062 **
## Year1998           -0.14661    0.05174   -2.83   0.0046 **
## Year1999           -0.14514    0.05224   -2.78   0.0055 **
## Year2000           -0.02276    0.04824   -0.47   0.6371
## Year2001           -0.04456    0.04847   -0.92   0.3579
## Year2002           -0.00449    0.04816   -0.09   0.9258
## Year2003           -0.01426    0.04821   -0.30   0.7674
## Year2004            0.02424    0.04783    0.51   0.6124
## Year2005            0.02824    0.04583    0.62   0.5377
## Year2006            0.00849    0.04483    0.19   0.8498
```

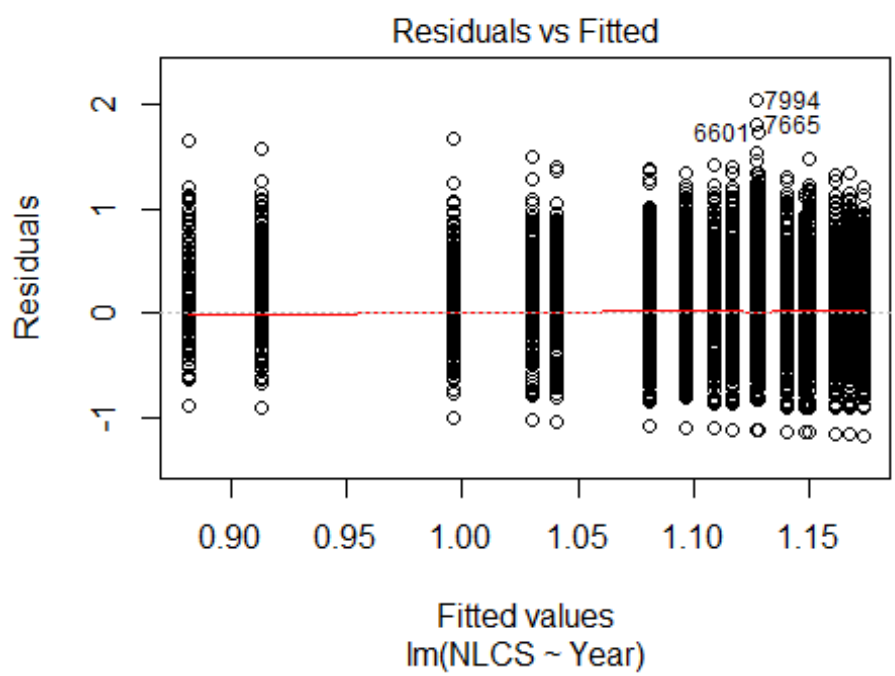
```
## Year2007           -0.01276    0.04423   -0.29   0.7730
## Year2008            0.02465    0.04367    0.56   0.5724
## Year2009            0.04886    0.04382    1.11   0.2649
## Year2010            0.02324    0.04459    0.52   0.6023
## Year2011            0.05173    0.04420    1.17   0.2419
## Year2012            0.03926    0.04518    0.87   0.3849
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.593
## Multiple R-squared:  0.00888,    Adjusted R-squared:  0.00691
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
##  717 weights are ~= 1. The remaining 7848 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0469  0.8640  0.9490  0.9060  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.17e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##                   GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.008  1           1.004
## Year             1.008 16           1.000
```



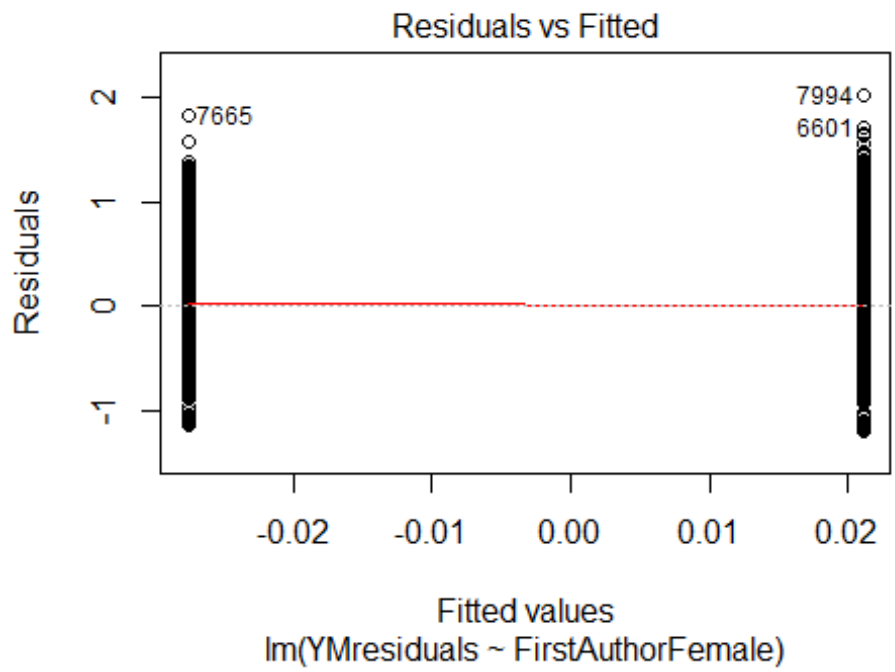

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##    Min     1Q Median     3Q    Max
## -1.133 -0.409  0.030  0.405  2.487
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.05229    0.03881   27.12   <2e-16 ***
## LastAuthorFemale1   0.02950    0.01480    1.99   0.0462 *
## Year1997           -0.14479    0.05285   -2.74   0.0062 **
## Year1998           -0.14735    0.05176   -2.85   0.0044 **
## Year1999           -0.14561    0.05226   -2.79   0.0053 **
## Year2000           -0.02297    0.04825   -0.48   0.6341
## Year2001           -0.04564    0.04845   -0.94   0.3462
## Year2002           -0.00495    0.04818   -0.10   0.9182
## Year2003           -0.01396    0.04820   -0.29   0.7721
## Year2004            0.02319    0.04789    0.48   0.6282
## Year2005            0.02785    0.04587    0.61   0.5439
## Year2006            0.00747    0.04484    0.17   0.8676
```

```
## Year2007          -0.01364    0.04423   -0.31   0.7578
## Year2008           0.02423    0.04367    0.55   0.5790
## Year2009           0.04816    0.04384    1.10   0.2719
## Year2010           0.02279    0.04460    0.51   0.6094
## Year2011           0.05170    0.04421    1.17   0.2422
## Year2012           0.03837    0.04521    0.85   0.3961
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.593
## Multiple R-squared:  0.00884,    Adjusted R-squared:  0.00687
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
##  721 weights are ~= 1. The remaining 7844 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0391  0.8640  0.9490  0.9060  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.17e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Sample size for the above analysis:  8565"
## [1] ""
## [1] ""
## [1] "####################################"
## [1] "Analysis of AJSC 3319"
## [1] "####################################"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  328  281  334  308  358  394  342  334  351  372  397  507  534  586  590
## 2011 2012
##  608  639
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  289  189  231  199  261  255  312  296  309  325  344  440  453  520  510
## 2011 2012
```

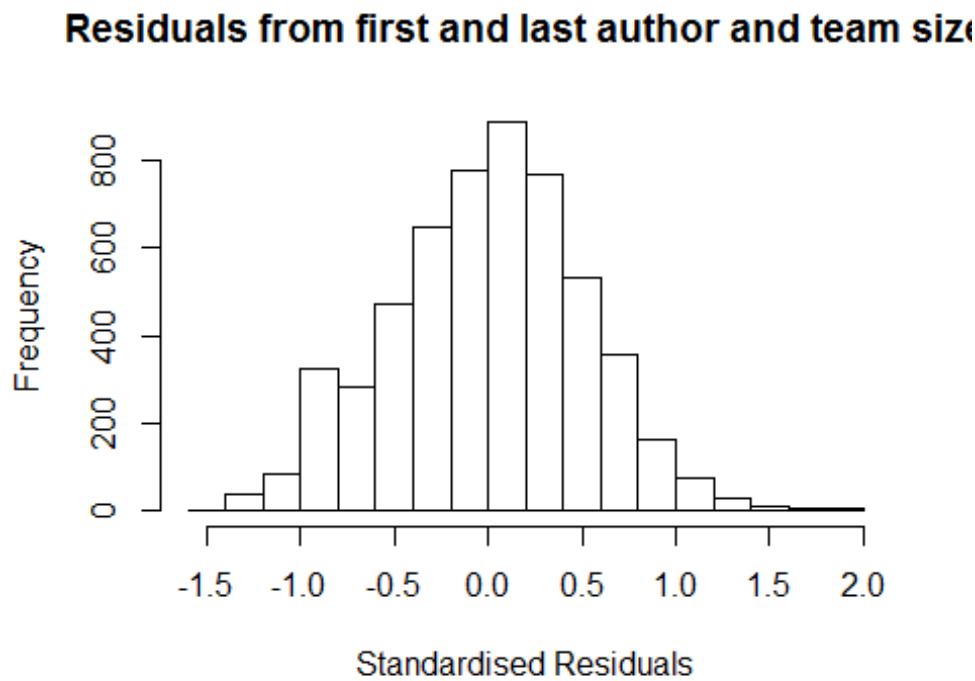
```
##  516  543
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  269  176  214  183  244  238  278  268  275  297  316  390  412  471  467
## 2011 2012
##  461  487
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
##  Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 33, df = 16, p-value = 0.007
```




```
##
##  Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 3.6, df = 1, p-value = 0.06
```

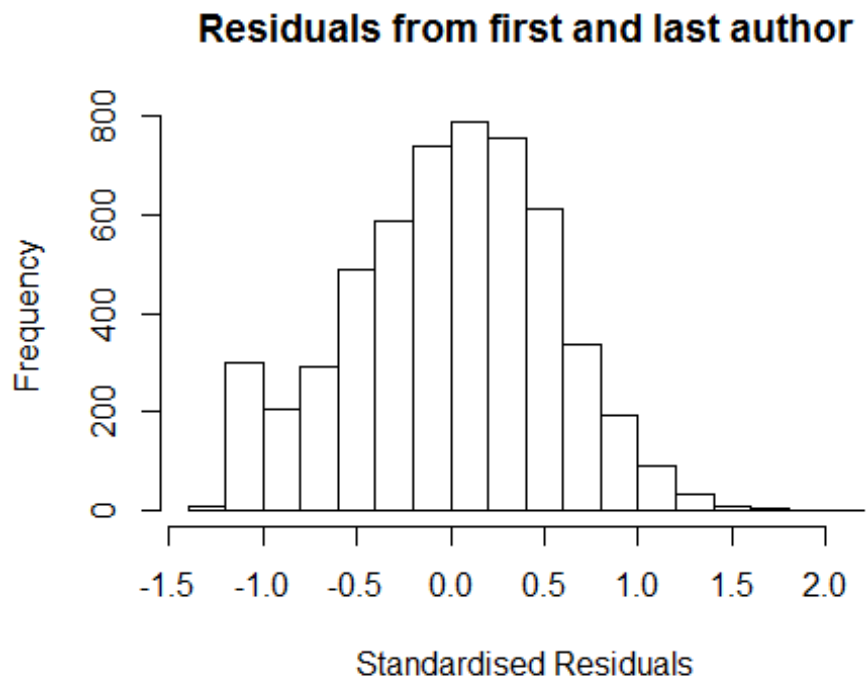



```
## [1] "Female first author team size 2018 geometric mean: 3.03607381737895"
## [1] "Male first author team size 2018 geometric mean: 2.88426988437604"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 35000, p-value = 0.5
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 2.89628692058512"
## [1] "Male last author team size 2018 geometric mean: 3.14140523278513"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 34000, p-value = 0.07
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.152  1           1.073
## LastAuthorFemale  1.154  1           1.074
## UniqueAuthors     1.120  4           1.014
## Year              1.143 16           1.004
```



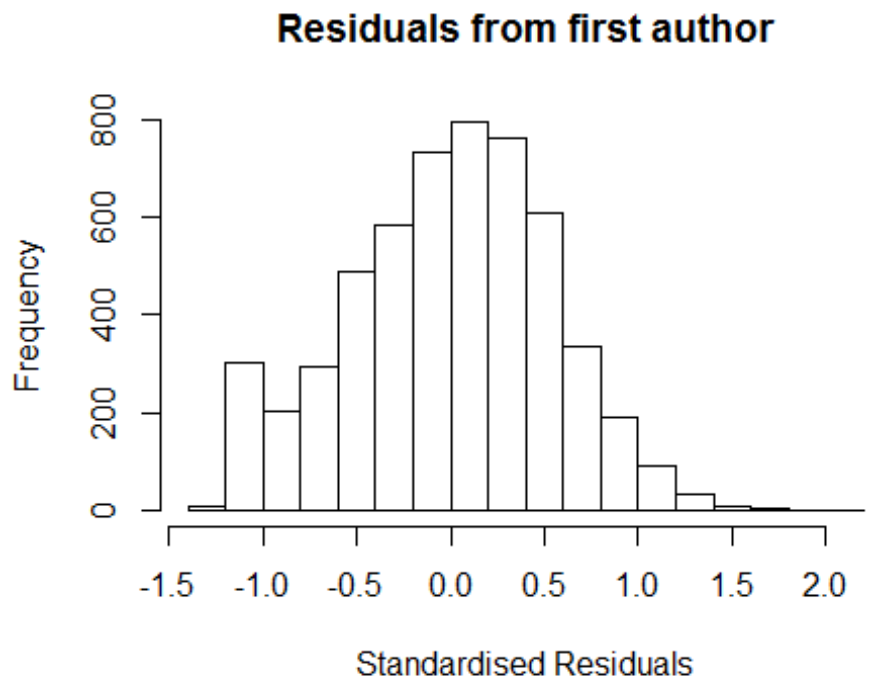

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
UniqueAuthors +
##     Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##     k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.4049 -0.3443  0.0224  0.3378  1.8665
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         0.87787    0.03905   22.48  < 2e-16 ***
## FirstAuthorFemale1  0.01446    0.01557    0.93  0.35301
## LastAuthorFemale1   0.00873    0.01539    0.57  0.57059
## UniqueAuthors2      0.31101    0.01999   15.55  < 2e-16 ***
## UniqueAuthors3      0.41804    0.02181   19.16  < 2e-16 ***
## UniqueAuthors4      0.48423    0.02444   19.81  < 2e-16 ***
## UniqueAuthors5      0.51164    0.02433   21.03  < 2e-16 ***
## Year1997           -0.24968    0.05287   -4.72  2.4e-06 ***
## Year1998           -0.22264    0.05045   -4.41  1.0e-05 ***
## Year1999           -0.15693    0.04761   -3.30  0.00099 ***
```

```
## Year2000           -0.09154    0.05034   -1.82  0.06904 .
## Year2001           -0.06556    0.04885   -1.34  0.17964
## Year2002           -0.01513    0.04792   -0.32  0.75227
## Year2003            0.00663    0.04860    0.14  0.89154
## Year2004           -0.01693    0.04943   -0.34  0.73192
## Year2005           -0.07290    0.04806   -1.52  0.12934
## Year2006           -0.04932    0.04774   -1.03  0.30164
## Year2007           -0.09160    0.04535   -2.02  0.04346 *
## Year2008           -0.04235    0.04275   -0.99  0.32190
## Year2009           -0.07166    0.04393   -1.63  0.10287
## Year2010           -0.06890    0.04436   -1.55  0.12045
## Year2011           -0.09595    0.04448   -2.16  0.03104 *
## Year2012           -0.07502    0.04442   -1.69  0.09127 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.512
## Multiple R-squared:  0.142,  Adjusted R-squared:  0.139
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  448 weights are ~= 1. The remaining 4998 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.156   0.868   0.951   0.908   0.986   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.84e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.121  1           1.059
## LastAuthorFemale  1.120  1           1.058
## Year              1.024 16           1.001
```



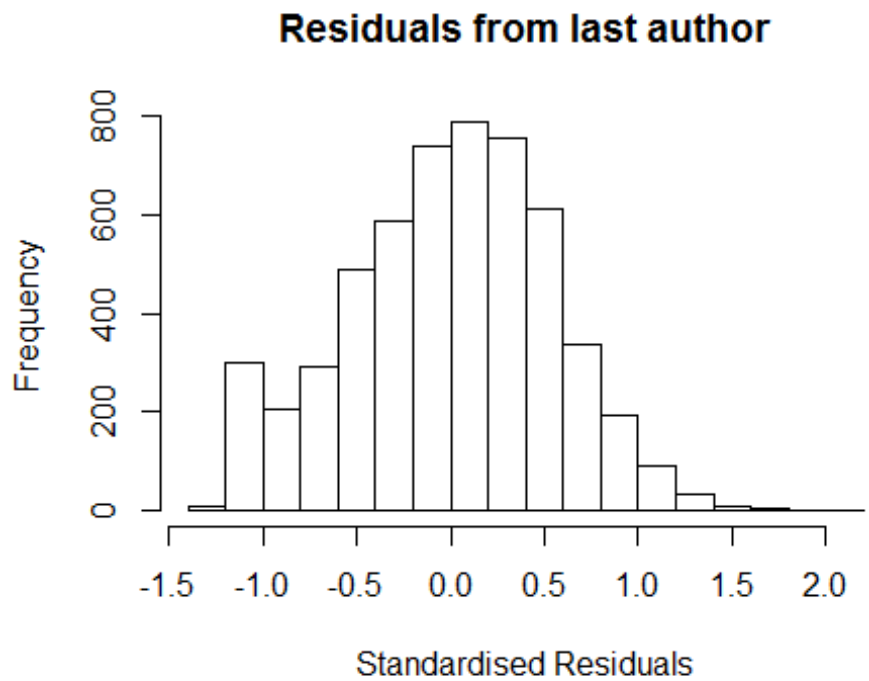

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.2113 -0.3730  0.0265  0.3760  2.0283
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)          1.10926    0.04018   27.61  < 2e-16 ***
## FirstAuthorFemale1   0.04609    0.01656    2.78   0.0054 **
## LastAuthorFemale1    0.00631    0.01630    0.39   0.6985
## Year1997            -0.28398    0.05671   -5.01  5.7e-07 ***
## Year1998            -0.24360    0.05447   -4.47  7.9e-06 ***
## Year1999            -0.15806    0.05133   -3.08   0.0021 **
## Year2000            -0.09327    0.05366   -1.74   0.0822 .
## Year2001            -0.11684    0.05167   -2.26   0.0238 *
## Year2002             0.02497    0.05125    0.49   0.6262
## Year2003             0.04967    0.05139    0.97   0.3339
## Year2004             0.02732    0.05317    0.51   0.6073
## Year2005            -0.03462    0.05181   -0.67   0.5040
```

```
## Year2006           -0.00191    0.05042   -0.04   0.9699
## Year2007           -0.05681    0.04898   -1.16   0.2461
## Year2008            0.02829    0.04581    0.62   0.5369
## Year2009           -0.00377    0.04682   -0.08   0.9359
## Year2010           -0.02161    0.04743   -0.46   0.6487
## Year2011           -0.03286    0.04766   -0.69   0.4905
## Year2012           -0.01567    0.04748   -0.33   0.7414
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.544
## Multiple R-squared:  0.0218, Adjusted R-squared:  0.0185
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  437 weights are ~= 1. The remaining 5009 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.134   0.872   0.949   0.907   0.985   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.84e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.011  1           1.006
## Year              1.011 16           1.000
```




```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.2095 -0.3719  0.0255  0.3764  2.0241
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)          1.11115    0.03987   27.87  < 2e-16 ***
## FirstAuthorFemale1  0.04846    0.01577    3.07   0.0021 **
## Year1997           -0.28407    0.05673   -5.01  5.7e-07 ***
## Year1998           -0.24370    0.05449   -4.47  7.9e-06 ***
## Year1999           -0.15779    0.05135   -3.07   0.0021 **
## Year2000           -0.09298    0.05366   -1.73   0.0832 .
## Year2001           -0.11667    0.05168   -2.26   0.0240 *
## Year2002            0.02509    0.05128    0.49   0.6246
## Year2003            0.04989    0.05141    0.97   0.3319
## Year2004            0.02704    0.05317    0.51   0.6111
## Year2005           -0.03457    0.05182   -0.67   0.5047
## Year2006           -0.00178    0.05044   -0.04   0.9719
```

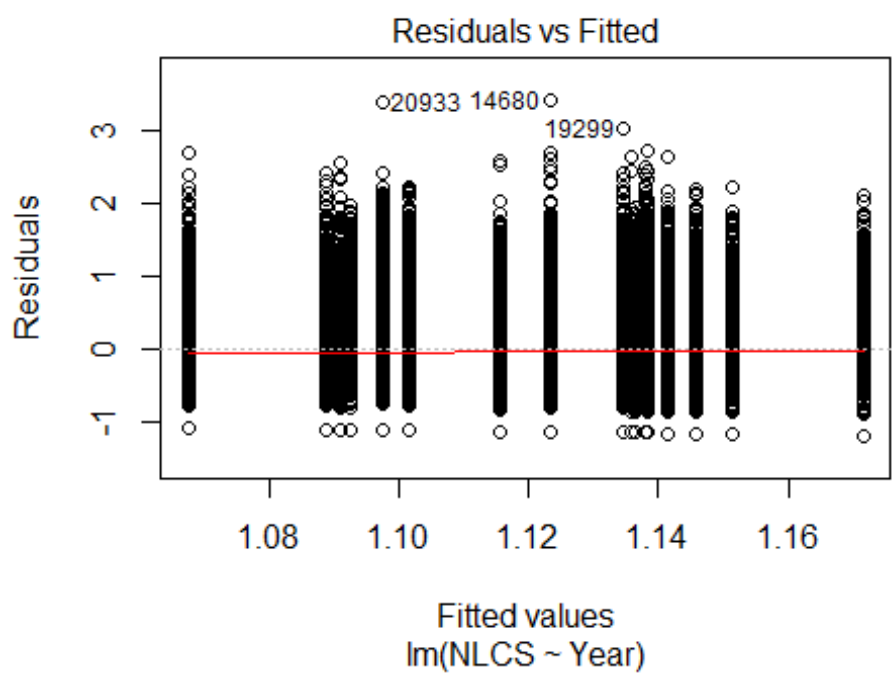
```
## Year2007           -0.05668    0.04900   -1.16   0.2474
## Year2008            0.02838    0.04583    0.62   0.5358
## Year2009           -0.00368    0.04685   -0.08   0.9374
## Year2010           -0.02131    0.04744   -0.45   0.6533
## Year2011           -0.03258    0.04767   -0.68   0.4944
## Year2012           -0.01565    0.04751   -0.33   0.7418
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.544
## Multiple R-squared:  0.0217, Adjusted R-squared:  0.0187
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  436 weights are ~= 1. The remaining 5010 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.136   0.873   0.949   0.907   0.985   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.84e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##                  GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.01  1           1.005
## Year             1.01 16           1.000
```



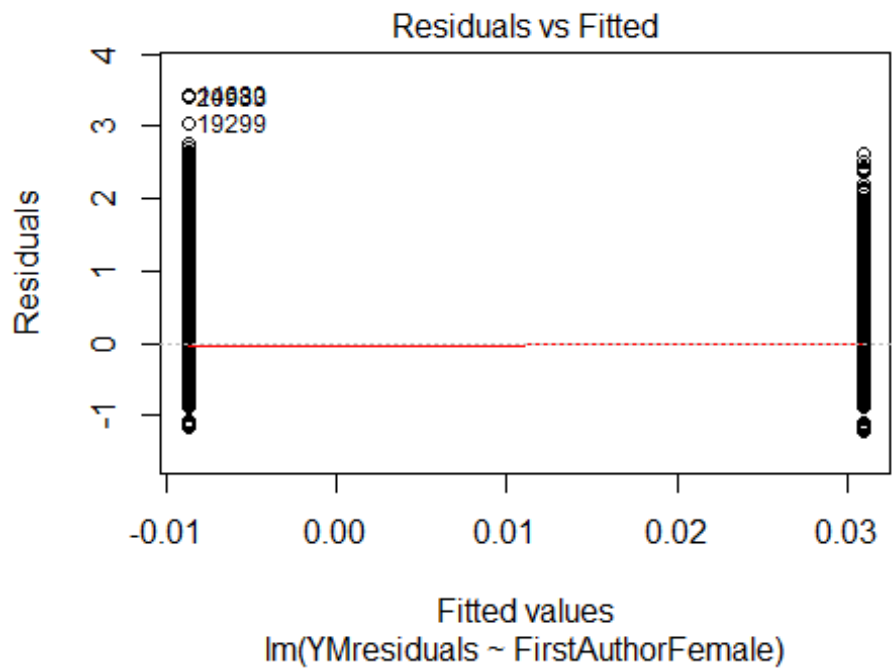

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.1996 -0.3771  0.0244  0.3731  2.0546
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.124550   0.039658   28.36  < 2e-16 ***
## LastAuthorFemale1   0.023227   0.015541    1.49    0.135
## Year1997           -0.284507   0.056862   -5.00  5.8e-07 ***
## Year1998           -0.243180   0.054589   -4.45  8.6e-06 ***
## Year1999           -0.158606   0.051291   -3.09    0.002 **
## Year2000           -0.092931   0.053728   -1.73    0.084 .
## Year2001           -0.115543   0.051687   -2.24    0.025 *
## Year2002            0.023578   0.051288    0.46    0.646
## Year2003            0.051850   0.051361    1.01    0.313
## Year2004            0.031953   0.053202    0.60    0.548
## Year2005           -0.032973   0.051828   -0.64    0.525
## Year2006            0.000574   0.050345    0.01    0.991
```

```
## Year2007          -0.053159   0.049015   -1.08    0.278
## Year2008           0.031077   0.045844    0.68    0.498
## Year2009          -0.002174   0.046841   -0.05    0.963
## Year2010          -0.018586   0.047396   -0.39    0.695
## Year2011          -0.028714   0.047622   -0.60    0.547
## Year2012          -0.011163   0.047396   -0.24    0.814
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.544
## Multiple R-squared:  0.0204, Adjusted R-squared:  0.0173
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  443 weights are ~= 1. The remaining 5003 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.122   0.874   0.949   0.907   0.985   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.84e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Sample size for the above analysis:  5446"
## [1] ""
## [1] ""
## [1] "####################################"
## [1] "Analysis of AJSC 3320"
## [1] "####################################"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  637  669  644  708  887 1001 1060 1100 1063 1080 1187 1374 1500 1818 1950
## 2011 2012
## 1984 1873
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  545  563  525  535  684  723  914  940  934  927 1039 1170 1283 1555 1649
## 2011 2012
```

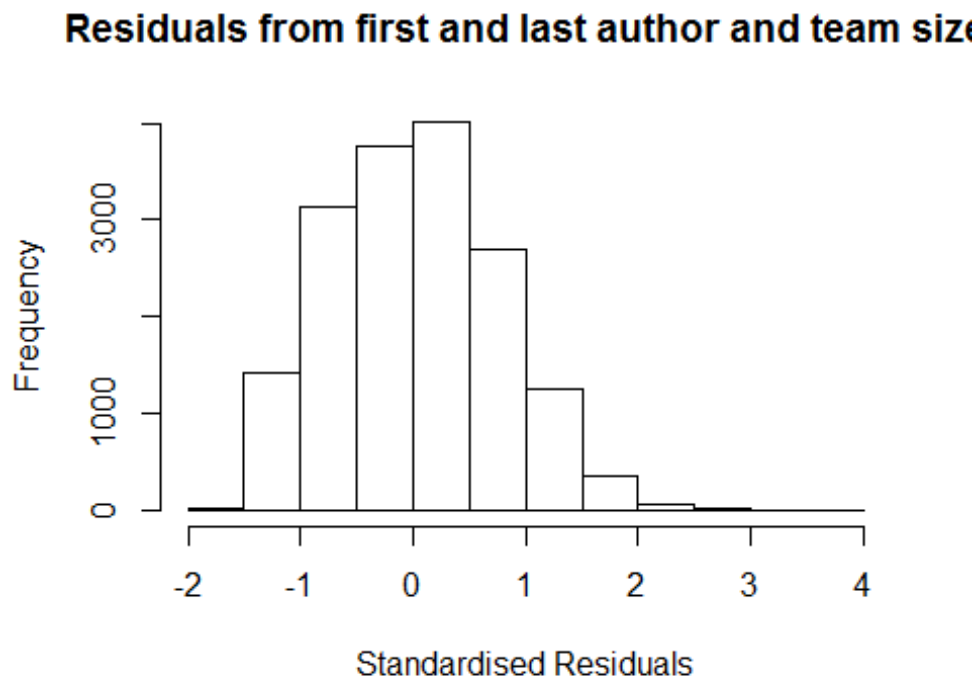
```
## 1672 1607
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  534  552  513  530  664  711  892  916  911  892 1013 1132 1255 1492 1588
## 2011 2012
## 1604 1542
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
##  Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 17, df = 16, p-value = 0.4
```




```
##
##  Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 6.8, df = 1, p-value = 0.009
```



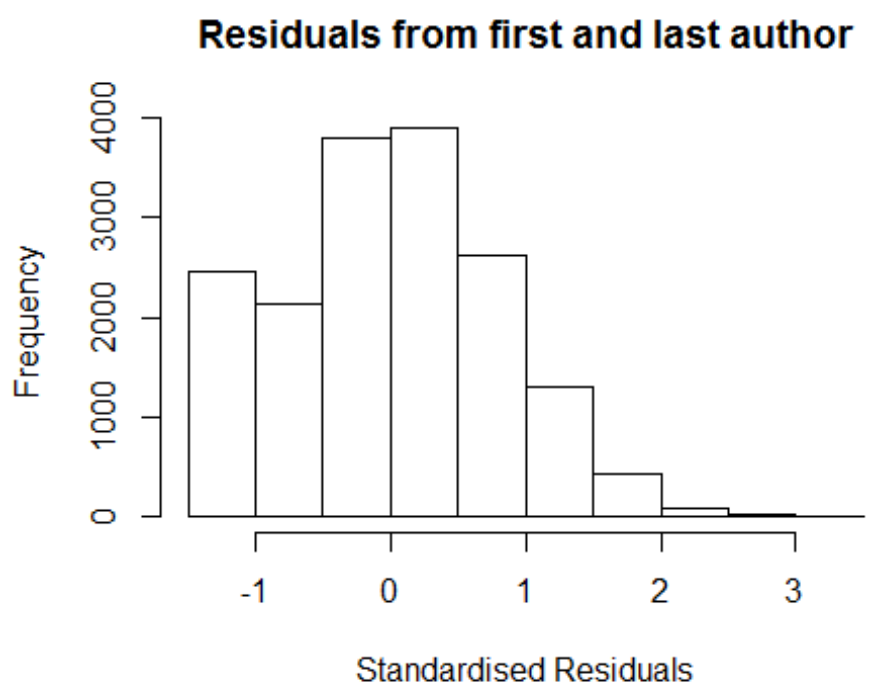

```
## [1] "Female first author team size 2018 geometric mean: 1.25800751434567"
## [1] "Male first author team size 2018 geometric mean: 1.32956594727615"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 230000, p-value = 0.01
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 1.30300439940843"
## [1] "Male last author team size 2018 geometric mean: 1.31002007042379"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 250000, p-value = 0.7
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 2.974  1           1.725
## LastAuthorFemale  2.977  1           1.725
## UniqueAuthors     1.025  4           1.003
## Year              1.027 16           1.001
```




```
## [1] "List of  14 outliers with residuals above  2.5"
##          ScopusId  NLCS Year OneField Fields residuals
## 334    0030305209 3.653 1996     3312      2     2.667
## 1323  21744459955 3.522 1997     1400      3     2.545
## 1697   0000490567 3.695 1998     1208      2     2.681
## 1739   0032018622 3.646 1998     3312      2     2.578
## 3237   0034336793 3.778 2000     3312      2     2.759
## 10378 33846152354 3.582 2006     1208      2     2.586
## 11977 51249156343 3.459 2007     1202      4     2.522
## 14678 68949083793 3.825 2009     1208      2     2.831
## 14680 68949090683 4.540 2009     1208      2     3.159
## 14682 68949092620 3.727 2009     1208      2     2.733
## 14687 68949135917 3.636 2009     1208      2     2.642
## 19299 79960371692 4.157 2011     3320      1     2.776
## 20933 84866718537 4.477 2012     1208      2     3.534
## 21583 84856137889 3.529 2012     3312      2     2.586
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
UniqueAuthors +
##     Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##     k.max = 1000))
##  \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.73637 -0.56001  0.00255  0.52993  3.53389
```

```
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         0.98648    0.03542   27.85  < 2e-16 ***
## FirstAuthorFemale1  0.02746    0.02465    1.11     0.27
## LastAuthorFemale1   0.02703    0.02451    1.10     0.27
## UniqueAuthors2      0.38704    0.01718   22.53  < 2e-16 ***
## UniqueAuthors3      0.48925    0.03196   15.31  < 2e-16 ***
## UniqueAuthors4      0.67452    0.06630   10.17  < 2e-16 ***
## UniqueAuthors5      0.60081    0.11733    5.12  3.1e-07 ***
## Year1997           -0.00982    0.04800   -0.20     0.84
## Year1998            0.02726    0.04887    0.56     0.58
## Year1999           -0.00733    0.04849   -0.15     0.88
## Year2000            0.03248    0.04523    0.72     0.47
## Year2001            0.05146    0.04667    1.10     0.27
## Year2002            0.03159    0.04427    0.71     0.48
## Year2003            0.04142    0.04295    0.96     0.33
## Year2004            0.04834    0.04370    1.11     0.27
## Year2005            0.06678    0.04388    1.52     0.13
## Year2006            0.00972    0.04261    0.23     0.82
## Year2007           -0.04926    0.04279   -1.15     0.25
## Year2008           -0.01860    0.04142   -0.45     0.65
## Year2009            0.00753    0.04039    0.19     0.85
## Year2010            0.01408    0.04009    0.35     0.73
## Year2011            0.00797    0.04016    0.20     0.84
## Year2012           -0.04338    0.04077   -1.06     0.29
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.795
## Multiple R-squared:  0.0535, Adjusted R-squared:  0.0523
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  1368 weights are ~= 1. The remaining 15373 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.010   0.867   0.950   0.919   0.985   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          5.97e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                  psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
```

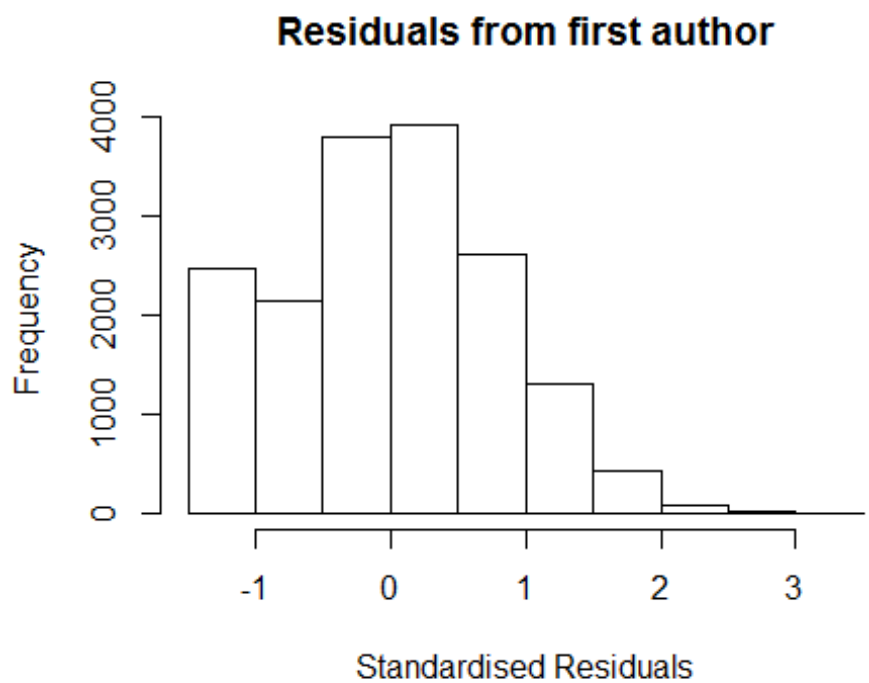
```
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##                     GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 3.255  1           1.804
## LastAuthorFemale  3.252  1           1.803
## Year              1.011 16           1.000
```




```
## [1] "List of  15 outliers with residuals above  2.5"
##          ScopusId  NLCS Year OneField Fields residuals
## 334    0030305209 3.653 1996     3312      2     2.612
## 1697   0000490567 3.695 1998     1208      2     2.620
## 1739   0032018622 3.646 1998     3312      2     2.511
## 3237   0034336793 3.778 2000     3312      2     2.688
## 3307   0034409890 3.593 2000     3312      2     2.503
## 7693  18044385112 3.777 2004     3320      1     2.647
## 9307  31044445688 3.879 2006     3312      2     2.774
## 11387 34249885738 3.757 2007     3312      2     2.687
## 14678 68949083793 3.825 2009     1208      2     2.740
## 14680 68949090683 4.540 2009     1208      2     3.455
## 14682 68949092620 3.727 2009     1208      2     2.642
## 14687 68949135917 3.636 2009     1208      2     2.551
## 16628 77957276053 3.638 2010     3312      2     2.541
## 19299 79960371692 4.157 2011     3320      1     3.067
## 20933 84866718537 4.477 2012     1208      2     3.425
```

```
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.19793 -0.57933 -0.00424  0.54065  3.45539
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.041243   0.035259   29.53   <2e-16 ***
## FirstAuthorFemale1  0.016969   0.026218    0.65    0.517
## LastAuthorFemale1   0.042534   0.026110    1.63    0.103
## Year1997           -0.000677   0.048457   -0.01    0.989
## Year1998            0.033945   0.049651    0.68    0.494
## Year1999            0.014277   0.049172    0.29    0.772
## Year2000            0.048834   0.045927    1.06    0.288
## Year2001            0.082413   0.047258    1.74    0.081 .
## Year2002            0.061877   0.044506    1.39    0.164
## Year2003            0.062854   0.043042    1.46    0.144
## Year2004            0.072228   0.043791    1.65    0.099 .
## Year2005            0.097183   0.044119    2.20    0.028 *
## Year2006            0.063875   0.042860    1.49    0.136
## Year2007           -0.014089   0.043301   -0.33    0.745
## Year2008            0.022840   0.041880    0.55    0.586
## Year2009            0.043365   0.040616    1.07    0.286
## Year2010            0.055602   0.040115    1.39    0.166
## Year2011            0.048345   0.040170    1.20    0.229
## Year2012            0.011089   0.040944    0.27    0.787
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.817
## Multiple R-squared:  0.00233,    Adjusted R-squared:  0.00126
## Convergence in 10 IRWLS iterations
##
## Robustness weights:
##  1416 weights are ~= 1. The remaining 15325 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.034   0.857   0.949   0.919   0.985   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          5.97e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
```

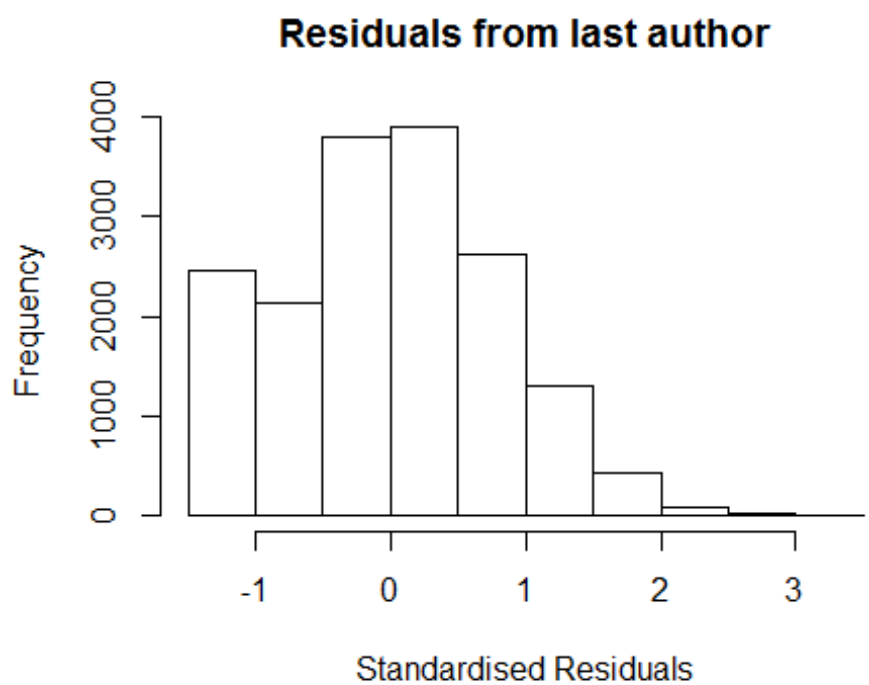
```
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                    cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.009  1           1.005
## Year              1.009 16           1.000
```




```
## [1] "List of  15 outliers with residuals above  2.5"
##          ScopusId  NLCS Year OneField Fields residuals
## 334    0030305209 3.653 1996     3312      2     2.612
## 1697   0000490567 3.695 1998     1208      2     2.620
## 1739   0032018622 3.646 1998     3312      2     2.511
## 3237   0034336793 3.778 2000     3312      2     2.688
## 3307   0034409890 3.593 2000     3312      2     2.503
## 7693  18044385112 3.777 2004     3320      1     2.647
## 9307  31044445688 3.879 2006     3312      2     2.774
## 11387 34249885738 3.757 2007     3312      2     2.687
## 14678 68949083793 3.825 2009     1208      2     2.740
## 14680 68949090683 4.540 2009     1208      2     3.455
## 14682 68949092620 3.727 2009     1208      2     2.642
## 14687 68949135917 3.636 2009     1208      2     2.551
```

```
## 16628 77957276053 3.638 2010     3312      2     2.541
## 19299 79960371692 4.157 2011     3320      1     3.067
## 20933 84866718537 4.477 2012     1208      2     3.425
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.19163 -0.58083 -0.00392  0.54204  3.45317
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.04e+00   3.53e-02   29.55  < 2e-16 ***
## FirstAuthorFemale1  5.14e-02   1.46e-02    3.52  0.00043 ***
## Year1997           -9.72e-05   4.85e-02    0.00  0.99840
## Year1998            3.45e-02   4.96e-02    0.69  0.48758
## Year1999            1.49e-02   4.92e-02    0.30  0.76209
## Year2000            4.92e-02   4.59e-02    1.07  0.28381
## Year2001            8.26e-02   4.73e-02    1.75  0.08035 .
## Year2002            6.24e-02   4.45e-02    1.40  0.16115
## Year2003            6.34e-02   4.31e-02    1.47  0.14079
## Year2004            7.29e-02   4.38e-02    1.66  0.09608 .
## Year2005            9.77e-02   4.41e-02    2.21  0.02682 *
## Year2006            6.49e-02   4.29e-02    1.51  0.12994
## Year2007           -1.38e-02   4.33e-02   -0.32  0.75003
## Year2008            2.35e-02   4.19e-02    0.56  0.57474
## Year2009            4.43e-02   4.06e-02    1.09  0.27557
## Year2010            5.64e-02   4.01e-02    1.41  0.15979
## Year2011            4.93e-02   4.02e-02    1.23  0.21975
## Year2012            1.21e-02   4.10e-02    0.29  0.76811
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.818
## Multiple R-squared:  0.00215,    Adjusted R-squared:  0.00114
## Convergence in 10 IRWLS iterations
##
## Robustness weights:
##  1423 weights are ~= 1. The remaining 15318 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0352  0.8570  0.9490  0.9190  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          5.97e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
```

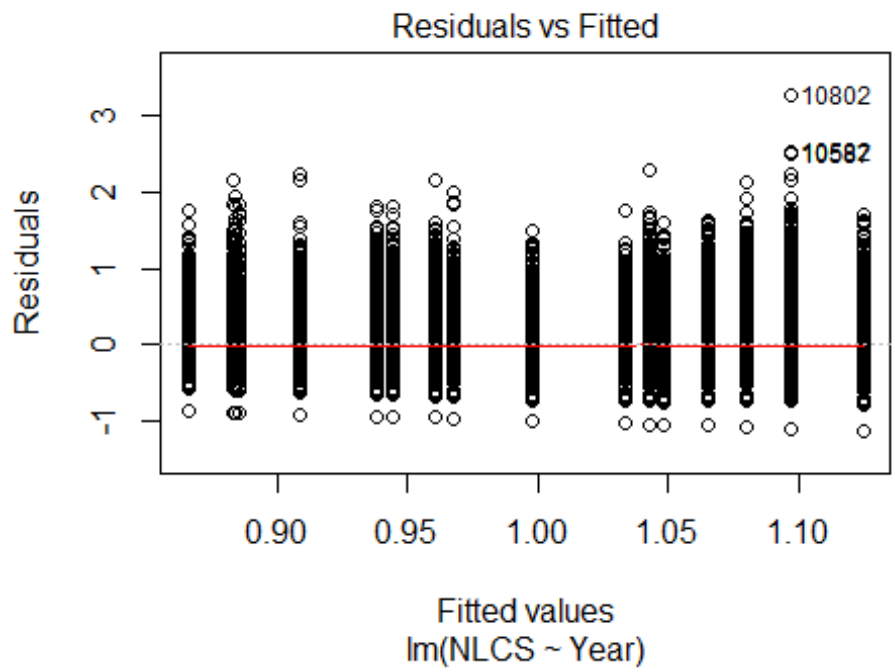
```
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##                   GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.008  1           1.004
## Year             1.008 16           1.000
```



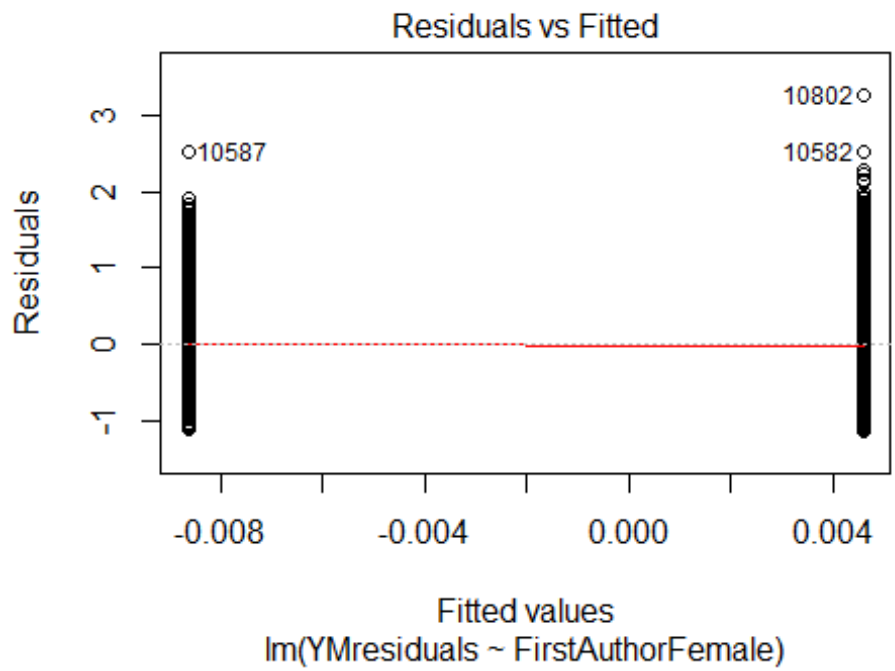

```
## [1] "List of  15 outliers with residuals above  2.5"
##          ScopusId  NLCS Year OneField Fields residuals
## 334    0030305209 3.653 1996     3312      2     2.612
## 1697   0000490567 3.695 1998     1208      2     2.620
## 1739   0032018622 3.646 1998     3312      2     2.511
## 3237   0034336793 3.778 2000     3312      2     2.688
## 3307   0034409890 3.593 2000     3312      2     2.503
## 7693  18044385112 3.777 2004     3320      1     2.647
## 9307  31044445688 3.879 2006     3312      2     2.774
## 11387 34249885738 3.757 2007     3312      2     2.687
## 14678 68949083793 3.825 2009     1208      2     2.740
## 14680 68949090683 4.540 2009     1208      2     3.455
```

```
## 14682 68949092620 3.727 2009     1208      2     2.642
## 14687 68949135917 3.636 2009     1208      2     2.551
## 16628 77957276053 3.638 2010     3312      2     2.541
## 19299 79960371692 4.157 2011     3320      1     3.067
## 20933 84866718537 4.477 2012     1208      2     3.425
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.19525 -0.58010 -0.00291  0.54019  3.45462
##
## Coefficients:
##                   Estimate Std. Error t value Pr(>|t|)
## (Intercept)        1.04196    0.03523   29.57  < 2e-16 ***
## LastAuthorFemale1  0.05614    0.01454    3.86  0.00011 ***
## Year1997          -0.00115    0.04843   -0.02  0.98106
## Year1998           0.03364    0.04965    0.68  0.49811
## Year1999           0.01423    0.04917    0.29  0.77226
## Year2000           0.04863    0.04592    1.06  0.28955
## Year2001           0.08237    0.04726    1.74  0.08135 .
## Year2002           0.06174    0.04451    1.39  0.16540
## Year2003           0.06284    0.04304    1.46  0.14427
## Year2004           0.07214    0.04379    1.65  0.09948 .
## Year2005           0.09714    0.04412    2.20  0.02768 *
## Year2006           0.06398    0.04286    1.49  0.13553
## Year2007          -0.01405    0.04331   -0.32  0.74567
## Year2008           0.02286    0.04188    0.55  0.58523
## Year2009           0.04342    0.04062    1.07  0.28506
## Year2010           0.05568    0.04011    1.39  0.16511
## Year2011           0.04844    0.04017    1.21  0.22783
## Year2012           0.01125    0.04094    0.27  0.78344
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.817
## Multiple R-squared:  0.0023, Adjusted R-squared:  0.00129
## Convergence in 10 IRWLS iterations
##
## Robustness weights:
##  1416 weights are ~= 1. The remaining 15325 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0342  0.8570  0.9490  0.9190  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
```

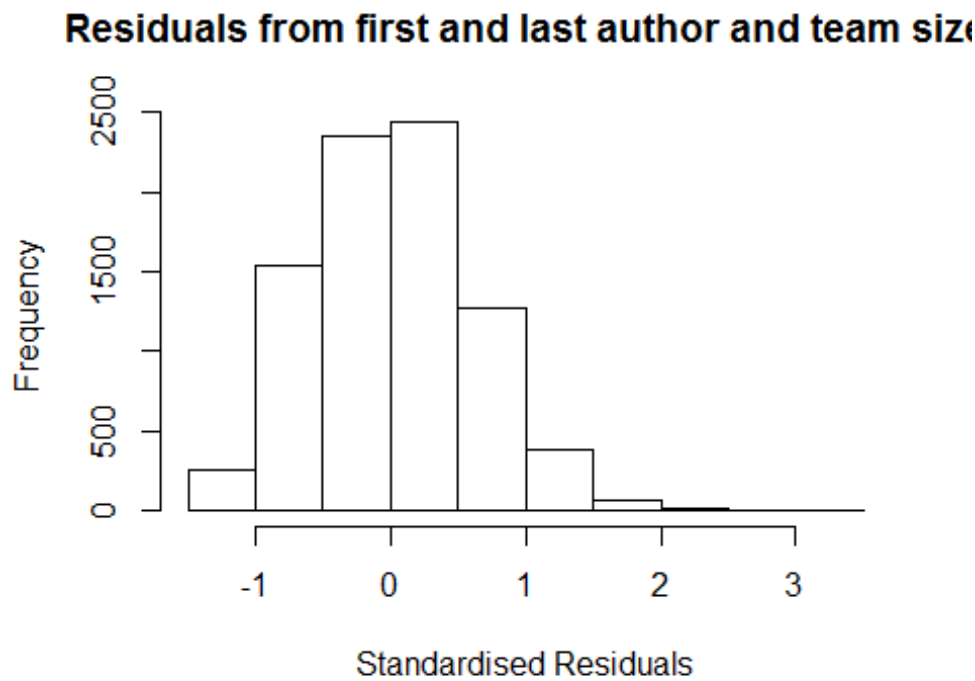
```
##          1.00e-07          1.00e-07          5.97e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                    psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Sample size for the above analysis:  16741"
## [1] ""
## [1] ""
## [1] "####################################"
## [1] "Analysis of AJSC 3321"
## [1] "####################################"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  386  367  410  448  559  590  583  396  467  498  568  627  735  892  920
## 2011 2012
##  886  909
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  347  317  343  368  475  475  509  359  416  443  492  554  629  748  780
## 2011 2012
##  734  784
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  337  306  329  353  454  454  483  348  399  425  455  520  599  698  736
## 2011 2012
##  683  734
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
##  Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 92, df = 16, p-value = 1e-12
```




```
##
##  Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 3.3, df = 1, p-value = 0.07
```

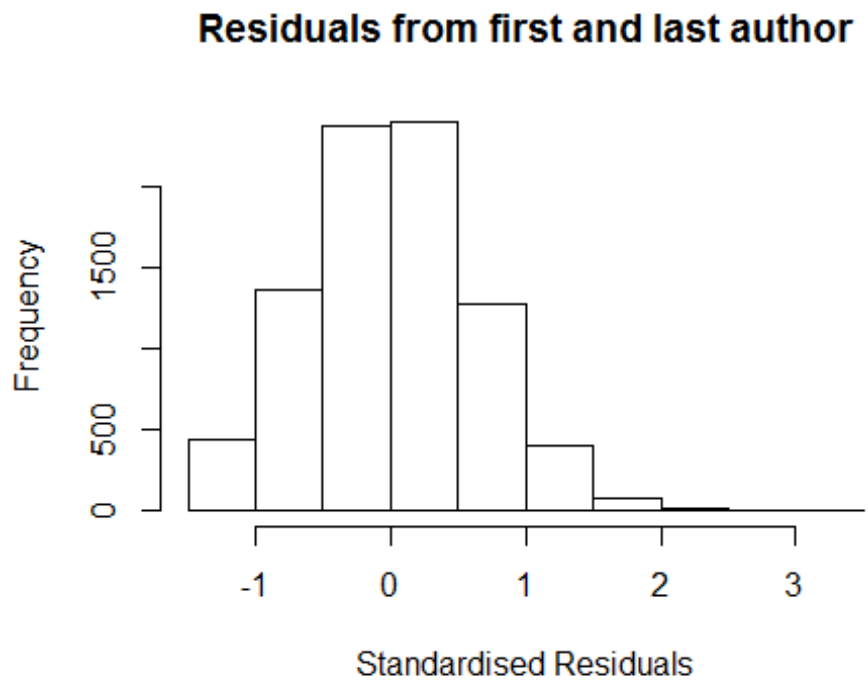



```
## [1] "Female first author team size 2018 geometric mean: 1.66742326763824"
## [1] "Male first author team size 2018 geometric mean: 1.60095127301474"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 70000, p-value = 0.3
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 1.64821830976359"
## [1] "Male last author team size 2018 geometric mean: 1.61416275494726"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 65000, p-value = 0.9
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.751  1           1.323
## LastAuthorFemale  1.751  1           1.323
## UniqueAuthors     1.060  4           1.007
## Year              1.057 16           1.002
```



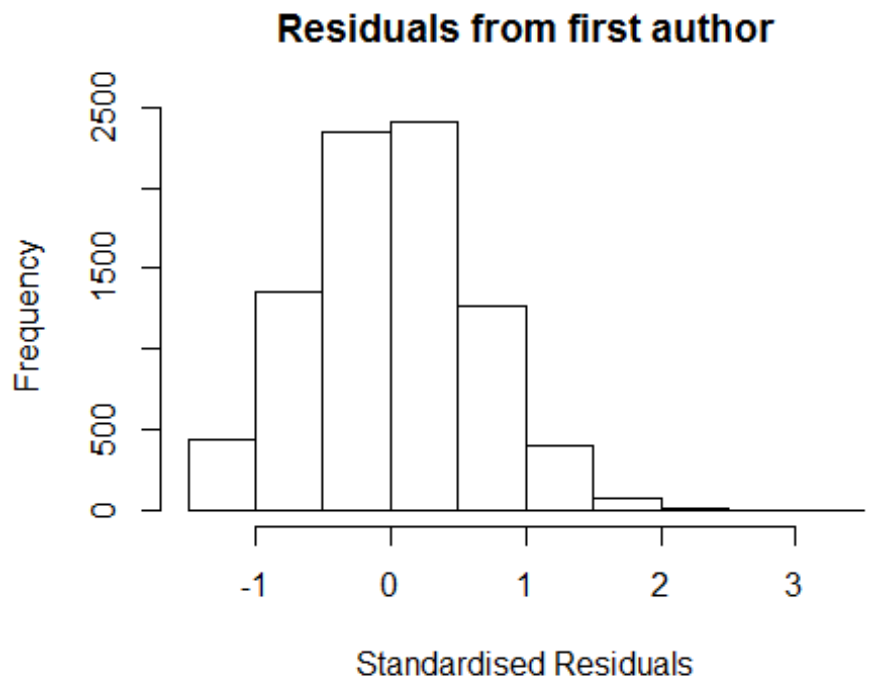

```
## [1] "List of  2 outliers with residuals above  2.5"
##          ScopusId  NLCS Year OneField Fields residuals
## 10587 84883394654 3.596 2012     1208      2     2.638
## 10802 84864376206 4.361 2012     1208      2     3.178
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
UniqueAuthors +
##     Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##     k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.2714 -0.4220  0.0026  0.4065  3.1779
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)         8.04e-01   3.33e-02   24.12  < 2e-16 ***
## FirstAuthorFemale1 -3.76e-02   1.90e-02   -1.98  0.04784 *
## LastAuthorFemale1   2.55e-02   1.90e-02    1.34  0.17899
## UniqueAuthors2      2.13e-01   1.59e-02   13.35  < 2e-16 ***
## UniqueAuthors3      2.33e-01   2.47e-02    9.44  < 2e-16 ***
## UniqueAuthors4      2.72e-01   4.21e-02    6.45  1.2e-10 ***
## UniqueAuthors5      2.54e-01   5.53e-02    4.60  4.3e-06 ***
## Year1997           -5.44e-05   4.73e-02    0.00  0.99908
## Year1998           -2.17e-02   4.92e-02   -0.44  0.65940
```

```
## Year1999           -8.56e-04   4.54e-02   -0.02  0.98497
## Year2000            6.30e-02   4.41e-02    1.43  0.15323
## Year2001            3.55e-02   4.29e-02    0.83  0.40803
## Year2002            3.71e-02   4.39e-02    0.85  0.39726
## Year2003            1.56e-02   4.63e-02    0.34  0.73653
## Year2004            1.29e-01   4.32e-02    2.99  0.00284 **
## Year2005            6.17e-02   4.39e-02    1.41  0.15992
## Year2006            1.01e-01   4.25e-02    2.37  0.01779 *
## Year2007            1.48e-01   4.10e-02    3.60  0.00031 ***
## Year2008            1.19e-01   4.14e-02    2.87  0.00414 **
## Year2009            1.52e-01   4.04e-02    3.78  0.00016 ***
## Year2010            2.09e-01   3.97e-02    5.27  1.4e-07 ***
## Year2011            1.58e-01   4.19e-02    3.76  0.00017 ***
## Year2012            1.67e-01   4.25e-02    3.92  9.0e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.621
## Multiple R-squared:  0.0473, Adjusted R-squared:  0.0448
## Convergence in 10 IRWLS iterations
##
## Robustness weights:
##  observation 8174 is an outlier with |weight| = 0 ( < 1.2e-05);
##  751 weights are ~= 1. The remaining 7561 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.032   0.861   0.950   0.912   0.985   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.20e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##                     GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.734  1           1.317
## LastAuthorFemale  1.733  1           1.316
## Year              1.018 16           1.001
```



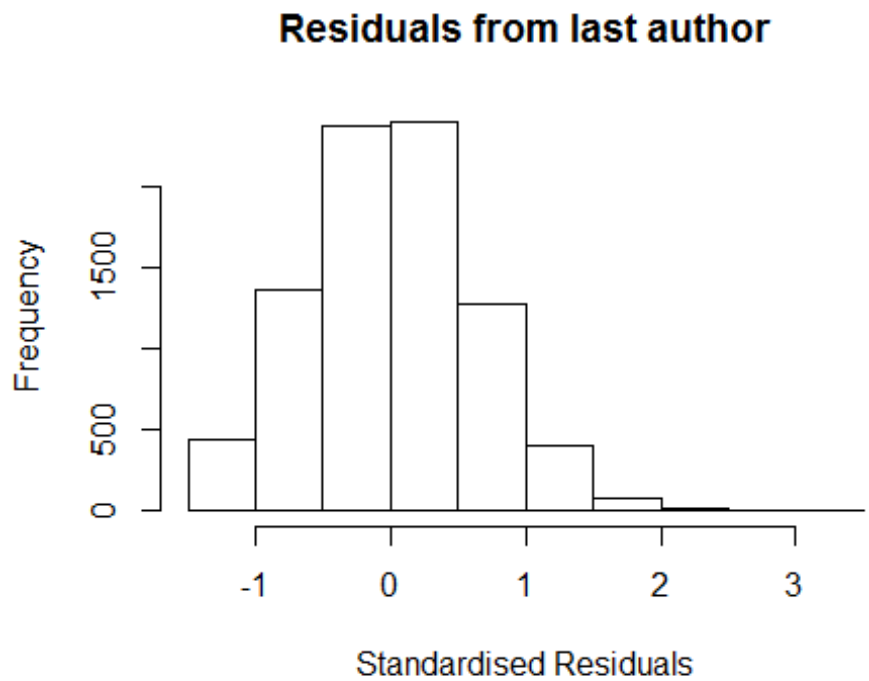

```
## [1] "List of  3 outliers with residuals above  2.5"
##          ScopusId  NLCS Year OneField Fields residuals
## 10582 84883352820 3.619 2012     1208      2     2.550
## 10587 84883394654 3.596 2012     1208      2     2.522
## 10802 84864376206 4.361 2012     1208      2     3.292
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.14154 -0.41811 -0.00106  0.42655  3.29196
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         0.87045    0.03387   25.70  < 2e-16 ***
## FirstAuthorFemale1 -0.02844    0.01926   -1.48  0.13996
## LastAuthorFemale1   0.03298    0.01925    1.71  0.08668 .
## Year1997           -0.00712    0.04796   -0.15  0.88199
## Year1998           -0.01256    0.05008   -0.25  0.80191
## Year1999           -0.01627    0.04636   -0.35  0.72565
## Year2000            0.06876    0.04502    1.53  0.12670
## Year2001            0.05906    0.04386    1.35  0.17816
## Year2002            0.04812    0.04517    1.07  0.28677
## Year2003            0.01550    0.04712    0.33  0.74229
```

```
## Year2004            0.15128    0.04412    3.43  0.00061 ***
## Year2005            0.08321    0.04489    1.85  0.06387 .
## Year2006            0.12052    0.04358    2.77  0.00570 **
## Year2007            0.16618    0.04169    3.99  6.8e-05 ***
## Year2008            0.14523    0.04251    3.42  0.00064 ***
## Year2009            0.18564    0.04122    4.50  6.8e-06 ***
## Year2010            0.23811    0.04050    5.88  4.3e-09 ***
## Year2011            0.18942    0.04292    4.41  1.0e-05 ***
## Year2012            0.19859    0.04375    4.54  5.7e-06 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.629
## Multiple R-squared:  0.0173, Adjusted R-squared:  0.0151
## Convergence in 10 IRWLS iterations
##
## Robustness weights:
##  observation 8174 is an outlier with |weight| = 0 ( < 1.2e-05);
##  711 weights are ~= 1. The remaining 7601 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0634  0.8700  0.9500  0.9120  0.9860  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.20e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##                   GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.01  1           1.005
## Year              1.01 16           1.000
```




```
## [1] "List of  3 outliers with residuals above  2.5"
##          ScopusId  NLCS Year OneField Fields residuals
## 10582 84883352820 3.619 2012     1208      2     2.550
## 10587 84883394654 3.596 2012     1208      2     2.522
## 10802 84864376206 4.361 2012     1208      2     3.292
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##       Min        1Q    Median        3Q       Max
## -1.113339 -0.419578  0.000682  0.426465  3.287567
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         0.87409    0.03376   25.89  < 2e-16 ***
## FirstAuthorFemale1 -0.00673    0.01469   -0.46  0.64673
## Year1997           -0.00732    0.04796   -0.15  0.87861
## Year1998           -0.01355    0.05004   -0.27  0.78657
## Year1999           -0.01700    0.04635   -0.37  0.71374
## Year2000            0.06785    0.04501    1.51  0.13172
## Year2001            0.05863    0.04383    1.34  0.18107
## Year2002            0.04822    0.04517    1.07  0.28576
## Year2003            0.01649    0.04713    0.35  0.72652
## Year2004            0.15144    0.04412    3.43  0.00060 ***
```

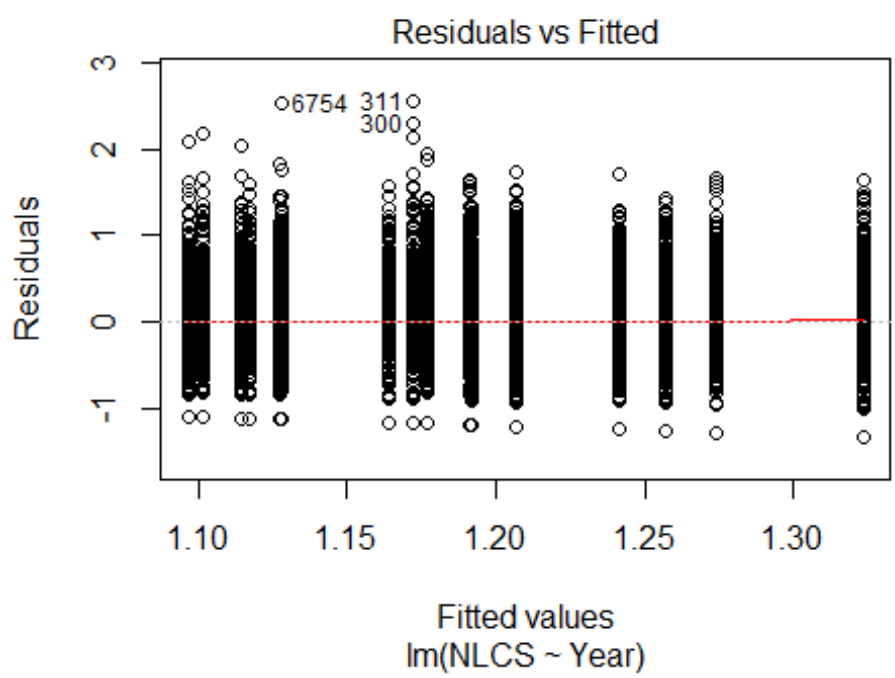
```
## Year2005            0.08357    0.04487    1.86  0.06255 .
## Year2006            0.12022    0.04358    2.76  0.00581 **
## Year2007            0.16657    0.04169    4.00  6.5e-05 ***
## Year2008            0.14571    0.04250    3.43  0.00061 ***
## Year2009            0.18617    0.04123    4.52  6.4e-06 ***
## Year2010            0.23925    0.04050    5.91  3.6e-09 ***
## Year2011            0.19023    0.04293    4.43  9.5e-06 ***
## Year2012            0.19934    0.04376    4.56  5.3e-06 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.629
## Multiple R-squared:  0.0169, Adjusted R-squared:  0.0149
## Convergence in 10 IRWLS iterations
##
## Robustness weights:
##  observation 8174 is an outlier with |weight| = 0 ( < 1.2e-05);
##  717 weights are ~= 1. The remaining 7595 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.065   0.870   0.950   0.912   0.986   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.20e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##                   GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.009  1           1.005
## Year             1.009 16           1.000
```



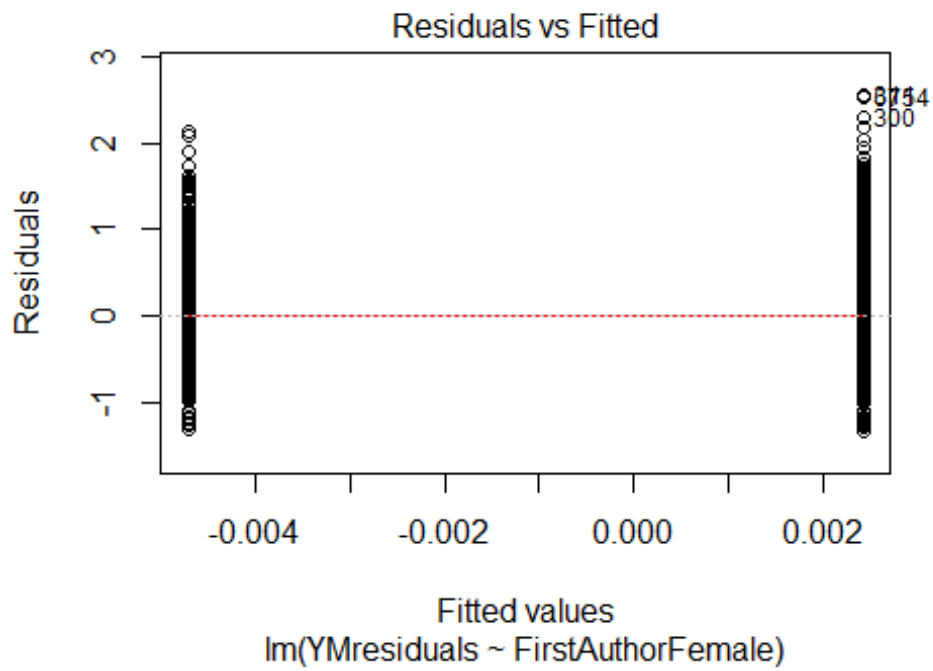

```
## [1] "List of  3 outliers with residuals above  2.5"
##          ScopusId  NLCS Year OneField Fields residuals
## 10582 84883352820 3.619 2012     1208      2     2.550
## 10587 84883394654 3.596 2012     1208      2     2.522
## 10802 84864376206 4.361 2012     1208      2     3.292
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##       Min        1Q    Median        3Q       Max
## -1.119322 -0.419685 -0.000499  0.425281  3.295828
##
## Coefficients:
##                   Estimate Std. Error t value Pr(>|t|)
## (Intercept)        0.86797    0.03385   25.64  < 2e-16 ***
## LastAuthorFemale1  0.01422    0.01471    0.97  0.33377
## Year1997          -0.00618    0.04797   -0.13  0.89748
## Year1998          -0.01248    0.05002   -0.25  0.80301
## Year1999          -0.01730    0.04639   -0.37  0.70915
## Year2000           0.06753    0.04503    1.50  0.13374
## Year2001           0.05820    0.04385    1.33  0.18451
## Year2002           0.04650    0.04518    1.03  0.30345
## Year2003           0.01504    0.04716    0.32  0.74980
## Year2004           0.15054    0.04415    3.41  0.00065 ***
```

```
## Year2005           0.08187    0.04491    1.82  0.06835 .
## Year2006           0.11899    0.04363    2.73  0.00640 **
## Year2007           0.16506    0.04173    3.96  7.7e-05 ***
## Year2008           0.14388    0.04252    3.38  0.00072 ***
## Year2009           0.18482    0.04126    4.48  7.6e-06 ***
## Year2010           0.23713    0.04052    5.85  5.0e-09 ***
## Year2011           0.18847    0.04294    4.39  1.1e-05 ***
## Year2012           0.19720    0.04375    4.51  6.7e-06 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.629
## Multiple R-squared:  0.017,  Adjusted R-squared:  0.015
## Convergence in 10 IRWLS iterations
##
## Robustness weights:
##  observation 8174 is an outlier with |weight| = 0 ( < 1.2e-05);
##  716 weights are ~= 1. The remaining 7596 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0622  0.8690  0.9500  0.9110  0.9860  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.20e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Sample size for the above analysis:  8313"
## [1] ""
## [1] ""
## [1] "####################################"
## [1] "Analysis of AJSC 3322"
## [1] "####################################"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  399  352  363  373  428  443  489  453  428  395  410  441  506  508  639
## 2011 2012
##  663  679
##
```

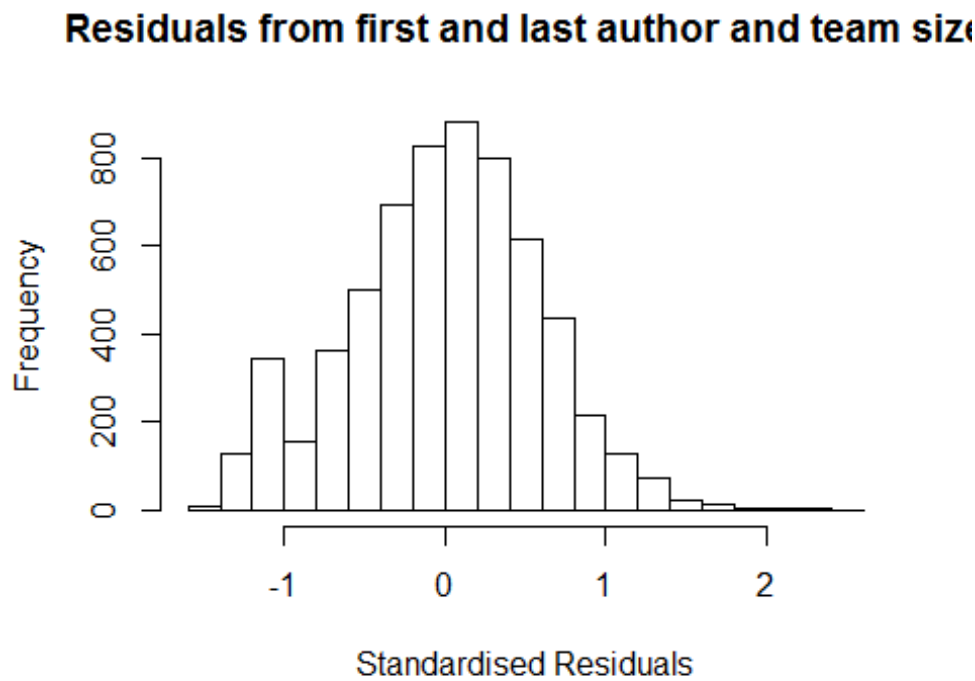
```
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  327  285  296  304  338  309  425  384  376  335  342  364  416  447  534
## 2011 2012
##  558  566
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  313  271  285  290  324  286  409  363  352  316  321  333  383  411  505
## 2011 2012
##  524  520
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
##  Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 18, df = 16, p-value = 0.3
```




```
##
##  Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 9, df = 1, p-value = 0.003
```

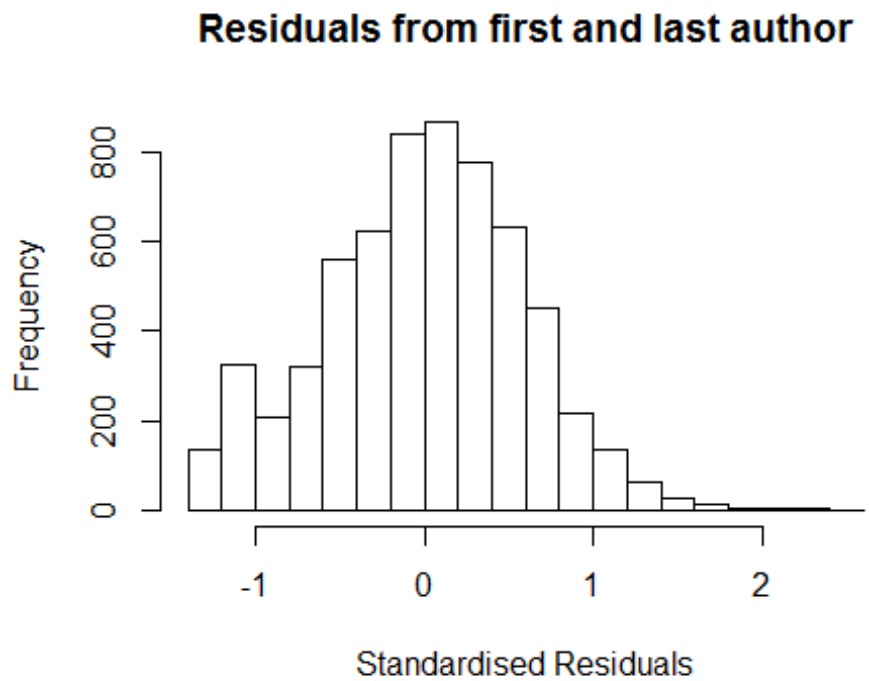



```
## [1] "Female first author team size 2018 geometric mean: 1.57682856153749"
## [1] "Male first author team size 2018 geometric mean: 1.70937011090819"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 34000, p-value = 0.08
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 1.57574579211996"
## [1] "Male last author team size 2018 geometric mean: 1.71166965212981"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 34000, p-value = 0.07
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.835  1           1.355
## LastAuthorFemale  1.817  1           1.348
## UniqueAuthors     1.078  4           1.009
## Year              1.094 16           1.003
```



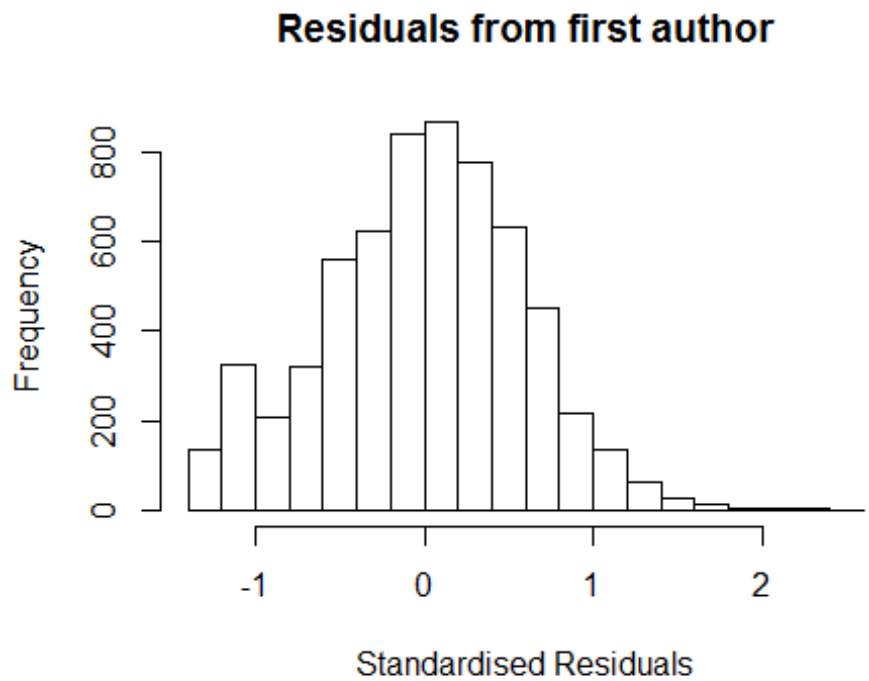

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
UniqueAuthors +
##     Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##     k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.4762 -0.3818  0.0212  0.3917  2.4311
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.092390   0.036823   29.67  < 2e-16 ***
## FirstAuthorFemale1  0.003807   0.021438    0.18  0.85905
## LastAuthorFemale1   0.000365   0.021460    0.02  0.98643
## UniqueAuthors2      0.169042   0.018241    9.27  < 2e-16 ***
## UniqueAuthors3      0.209690   0.026187    8.01  1.4e-15 ***
## UniqueAuthors4      0.156065   0.046408    3.36  0.00078 ***
## UniqueAuthors5      0.235445   0.051831    4.54  5.7e-06 ***
## Year1997           -0.067343   0.052743   -1.28  0.20171
## Year1998            0.039622   0.050046    0.79  0.42856
## Year1999           -0.084549   0.048944   -1.73  0.08413 .
```

```
## Year2000           -0.039044   0.049748   -0.78  0.43259
## Year2001           -0.043772   0.051337   -0.85  0.39389
## Year2002           -0.009550   0.046667   -0.20  0.83787
## Year2003            0.052084   0.048285    1.08  0.28077
## Year2004            0.046168   0.047236    0.98  0.32841
## Year2005            0.005743   0.049372    0.12  0.90740
## Year2006            0.092888   0.047255    1.97  0.04938 *
## Year2007            0.078686   0.049522    1.59  0.11214
## Year2008            0.173738   0.046921    3.70  0.00022 ***
## Year2009            0.089516   0.046794    1.91  0.05580 .
## Year2010           -0.040561   0.045380   -0.89  0.37146
## Year2011            0.020059   0.045042    0.45  0.65609
## Year2012           -0.011056   0.046574   -0.24  0.81237
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.573
## Multiple R-squared:  0.0343, Adjusted R-squared:  0.0309
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  533 weights are ~= 1. The remaining 5673 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.032   0.865   0.950   0.902   0.986   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.61e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.873  1           1.368
## LastAuthorFemale  1.855  1           1.362
## Year              1.038 16           1.001
```



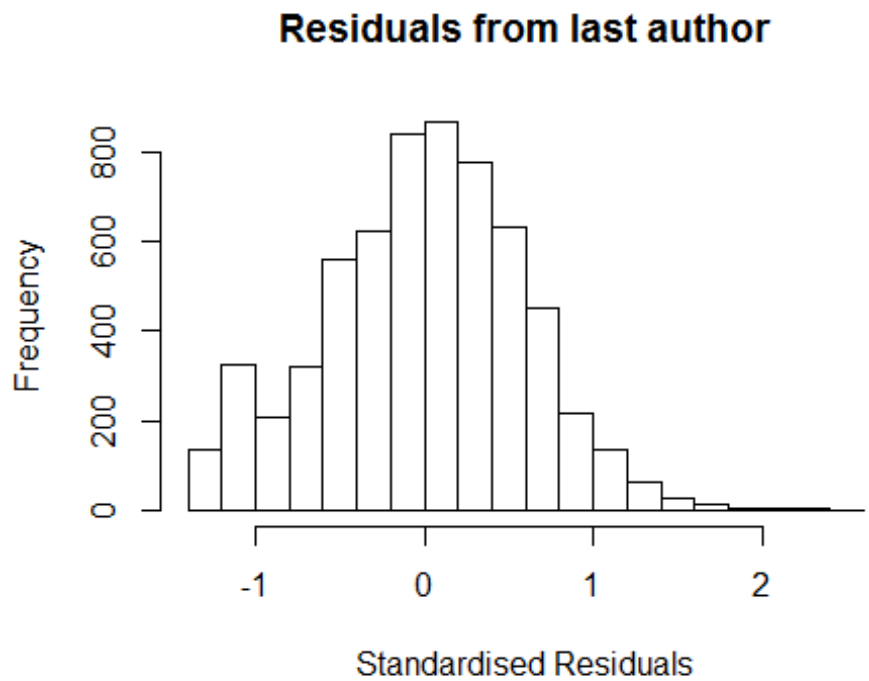

```
## [1] "List of  1 outliers with residuals above  2.5"
##         ScopusId  NLCS Year OneField Fields residuals
## 6754 77954069125 3.652 2010     3303      3     2.526
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.3363 -0.3983  0.0216  0.3992  2.5256
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.16e+00   3.63e-02   31.88  < 2e-16 ***
## FirstAuthorFemale1 -3.95e-03   2.20e-02   -0.18  0.85722
## LastAuthorFemale1   6.77e-05   2.20e-02    0.00  0.99754
## Year1997           -7.90e-02   5.32e-02   -1.49  0.13735
## Year1998            5.00e-02   5.04e-02    0.99  0.32109
## Year1999           -8.29e-02   4.88e-02   -1.70  0.08953 .
## Year2000           -3.56e-02   5.01e-02   -0.71  0.47664
## Year2001           -3.27e-02   5.16e-02   -0.63  0.52649
## Year2002           -1.90e-02   4.65e-02   -0.41  0.68256
## Year2003            5.31e-02   4.85e-02    1.09  0.27401
## Year2004            4.45e-02   4.72e-02    0.94  0.34577
## Year2005            1.58e-02   4.95e-02    0.32  0.74913
```

```
## Year2006            1.03e-01   4.73e-02    2.18  0.02908 *
## Year2007            8.52e-02   5.01e-02    1.70  0.08900 .
## Year2008            1.78e-01   4.71e-02    3.78  0.00016 ***
## Year2009            1.05e-01   4.69e-02    2.24  0.02535 *
## Year2010           -3.19e-02   4.58e-02   -0.70  0.48594
## Year2011            4.05e-02   4.52e-02    0.90  0.37055
## Year2012            1.65e-02   4.64e-02    0.36  0.72219
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.579
## Multiple R-squared:  0.012,  Adjusted R-squared:  0.00915
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  569 weights are ~= 1. The remaining 5637 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0181  0.8630  0.9490  0.9020  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.61e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.031  1           1.016
## Year              1.031 16           1.001
```




```
## [1] "List of  1 outliers with residuals above  2.5"
##         ScopusId  NLCS Year OneField Fields residuals
## 6754 77954069125 3.652 2010     3303      3     2.526
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.3363 -0.3983  0.0216  0.3992  2.5256
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.15831    0.03631   31.90  < 2e-16 ***
## FirstAuthorFemale1 -0.00391    0.01631   -0.24  0.81062
## Year1997           -0.07904    0.05319   -1.49  0.13733
## Year1998            0.04999    0.05037    0.99  0.32102
## Year1999           -0.08292    0.04883   -1.70  0.08954 .
## Year2000           -0.03564    0.05008   -0.71  0.47671
## Year2001           -0.03267    0.05159   -0.63  0.52661
## Year2002           -0.01903    0.04654   -0.41  0.68270
## Year2003            0.05310    0.04852    1.09  0.27384
## Year2004            0.04453    0.04723    0.94  0.34574
## Year2005            0.01584    0.04951    0.32  0.74904
## Year2006            0.10332    0.04732    2.18  0.02905 *
```

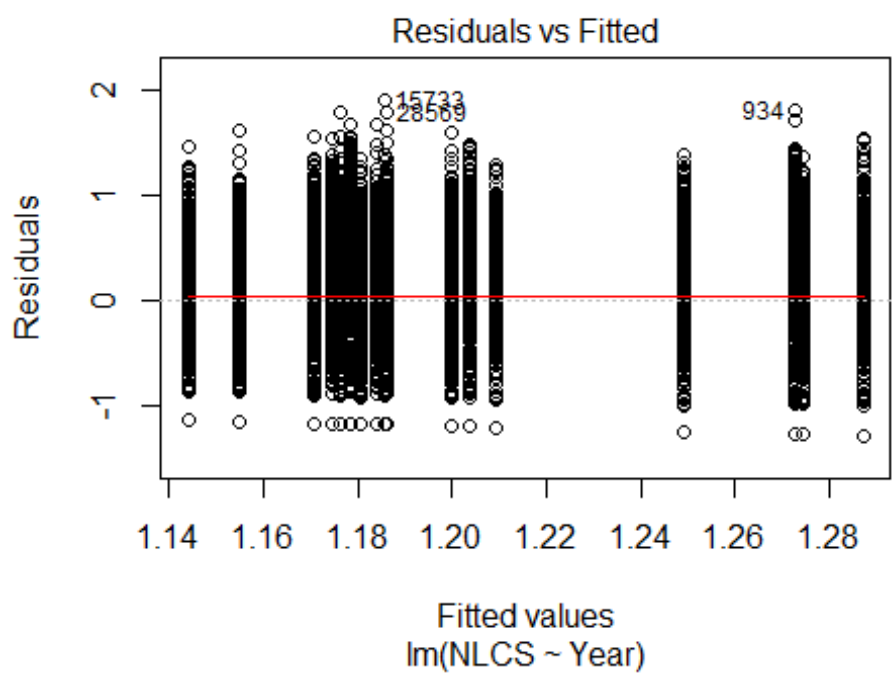
```
## Year2007            0.08525    0.05011    1.70  0.08891 .
## Year2008            0.17795    0.04708    3.78  0.00016 ***
## Year2009            0.10486    0.04685    2.24  0.02526 *
## Year2010           -0.03191    0.04580   -0.70  0.48604
## Year2011            0.04050    0.04519    0.90  0.37018
## Year2012            0.01651    0.04638    0.36  0.72190
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.579
## Multiple R-squared:  0.012,  Adjusted R-squared:  0.00931
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  569 weights are ~= 1. The remaining 5637 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.018   0.863   0.949   0.902   0.985   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.61e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##                   GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.022  1           1.011
## Year             1.022 16           1.001
```



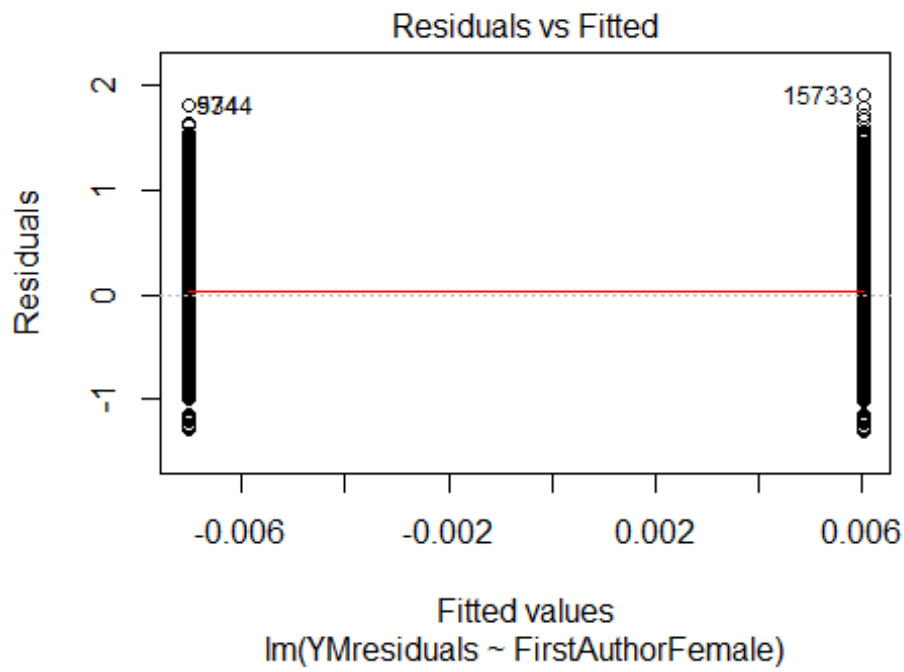

```
## [1] "List of  1 outliers with residuals above  2.5"
##         ScopusId  NLCS Year OneField Fields residuals
## 6754 77954069125 3.652 2010     3303      3     2.526
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.3358 -0.3992  0.0211  0.3988  2.5262
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.15799    0.03624   31.95  < 2e-16 ***
## LastAuthorFemale1 -0.00261    0.01632   -0.16  0.87281
## Year1997          -0.07919    0.05319   -1.49  0.13663
## Year1998           0.05002    0.05038    0.99  0.32077
## Year1999          -0.08287    0.04883   -1.70  0.08971 .
## Year2000          -0.03558    0.05008   -0.71  0.47753
## Year2001          -0.03281    0.05160   -0.64  0.52487
## Year2002          -0.01921    0.04653   -0.41  0.67976
## Year2003           0.05302    0.04854    1.09  0.27474
## Year2004           0.04447    0.04723    0.94  0.34648
## Year2005           0.01555    0.04947    0.31  0.75329
## Year2006           0.10325    0.04733    2.18  0.02920 *
```

```
## Year2007           0.08516    0.05013    1.70  0.08939 .
## Year2008           0.17783    0.04712    3.77  0.00016 ***
## Year2009           0.10476    0.04689    2.23  0.02552 *
## Year2010          -0.03222    0.04579   -0.70  0.48167
## Year2011           0.04028    0.04521    0.89  0.37297
## Year2012           0.01621    0.04644    0.35  0.72705
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.579
## Multiple R-squared:  0.012,  Adjusted R-squared:  0.00931
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  567 weights are ~= 1. The remaining 5639 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0179  0.8630  0.9490  0.9020  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.61e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Sample size for the above analysis:  6206"
## [1] ""
## [1] ""
## [1] "####################################"
## [1] "Analysis of AJSC 3400"
## [1] "####################################"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 1629 1310 1569 1553 1583 1633 1555 1336 1517 1544 1837 1754 1812 1853 1780
## 2011 2012
## 1887 1843
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 1012  820  993  989  967  938 1105  951 1121 1111 1309 1266 1234 1265 1246
## 2011 2012
```

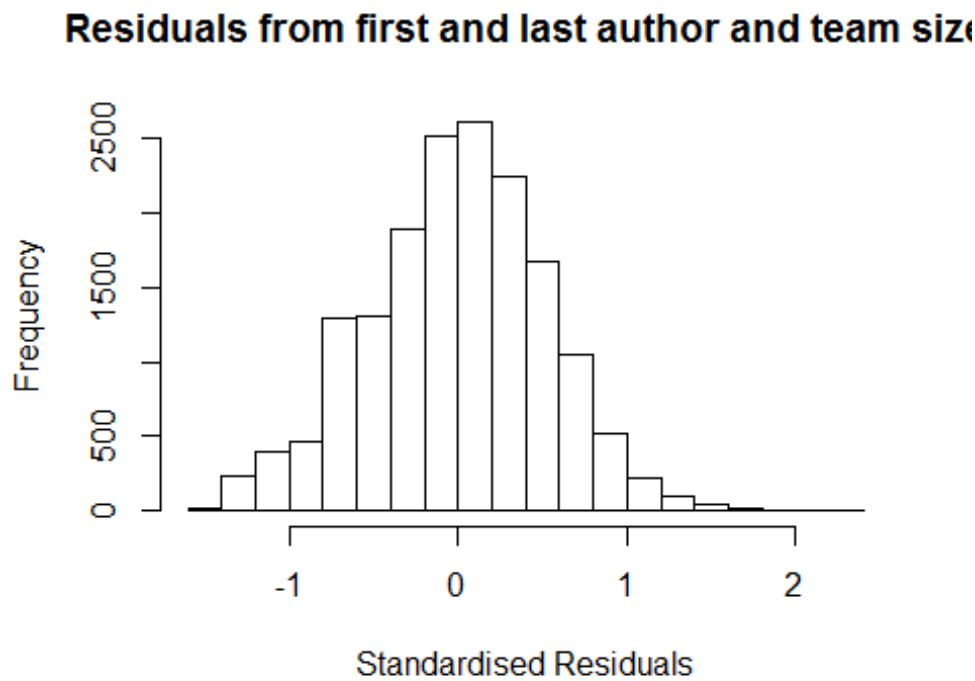
```
## 1295 1251
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  896  731  896  876  853  828  988  835 1007  973 1136 1118 1041 1121 1075
## 2011 2012
## 1128 1102
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
##  Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 280, df = 16, p-value <2e-16
```




```
##
##  Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 0.039, df = 1, p-value = 0.8
```

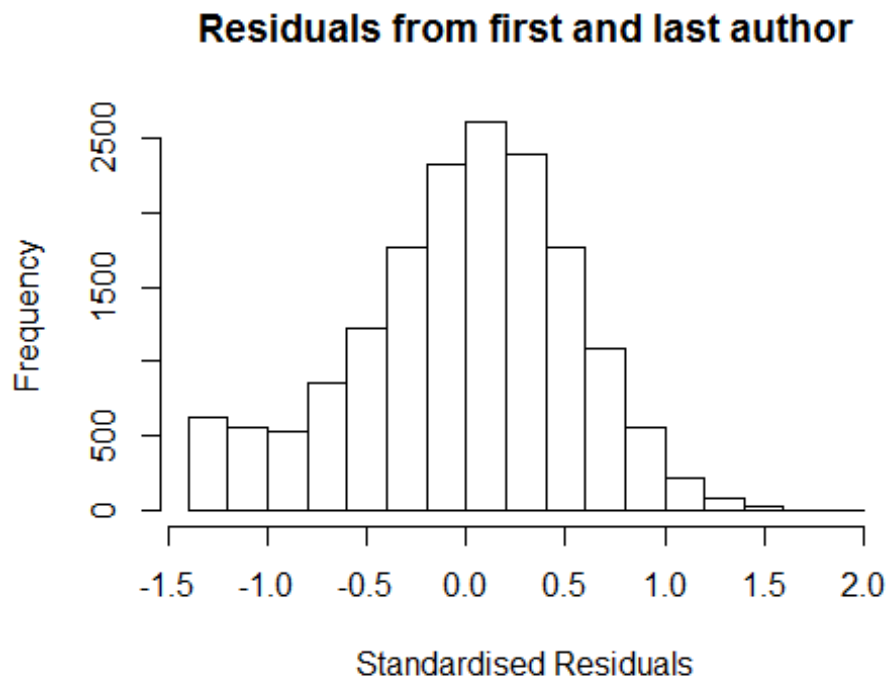



```
## [1] "Female first author team size 2018 geometric mean: 4.43327729132147"
## [1] "Male first author team size 2018 geometric mean: 4.07319880135915"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 160000, p-value = 0.1
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 4.23921377690888"
## [1] "Male last author team size 2018 geometric mean: 4.33898521869423"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 150000, p-value = 0.2
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.026  1           1.013
## LastAuthorFemale  1.024  1           1.012
## UniqueAuthors     1.062  4           1.008
## Year              1.075 16           1.002
```



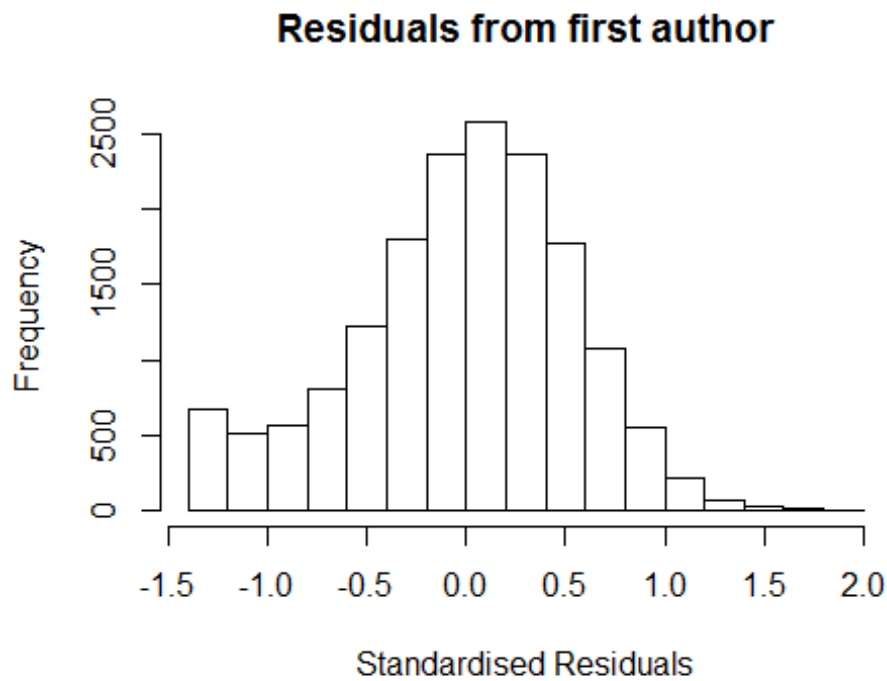

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
UniqueAuthors +
##     Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##     k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.4852 -0.3500  0.0143  0.3495  2.3364
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         0.80362    0.02788   28.83  < 2e-16 ***
## FirstAuthorFemale1 -0.01152    0.00843   -1.37   0.1721
## LastAuthorFemale1  -0.02237    0.00910   -2.46   0.0139 *
## UniqueAuthors2      0.39831    0.02133   18.68  < 2e-16 ***
## UniqueAuthors3      0.53639    0.01966   27.29  < 2e-16 ***
## UniqueAuthors4      0.59054    0.01944   30.38  < 2e-16 ***
## UniqueAuthors5      0.66753    0.01827   36.53  < 2e-16 ***
## Year1997            0.01406    0.03255    0.43   0.6658
## Year1998           -0.04012    0.02993   -1.34   0.1802
## Year1999           -0.06241    0.02927   -2.13   0.0330 *
```

```
## Year2000           -0.03495    0.02897   -1.21   0.2277
## Year2001            0.00431    0.02867    0.15   0.8805
## Year2002           -0.08031    0.02816   -2.85   0.0044 **
## Year2003           -0.08733    0.02785   -3.14   0.0017 **
## Year2004           -0.12203    0.02694   -4.53  5.9e-06 ***
## Year2005           -0.12738    0.02723   -4.68  2.9e-06 ***
## Year2006           -0.15149    0.02595   -5.84  5.4e-09 ***
## Year2007           -0.14276    0.02632   -5.42  5.9e-08 ***
## Year2008           -0.14147    0.02754   -5.14  2.8e-07 ***
## Year2009           -0.13125    0.02659   -4.94  8.0e-07 ***
## Year2010           -0.16454    0.02669   -6.16  7.2e-10 ***
## Year2011           -0.15701    0.02744   -5.72  1.1e-08 ***
## Year2012           -0.15607    0.02831   -5.51  3.6e-08 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.505
## Multiple R-squared:  0.142,  Adjusted R-squared:  0.141
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
##  1440 weights are ~= 1. The remaining 15164 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0006  0.8580  0.9470  0.9010  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          6.02e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.019  1           1.009
## LastAuthorFemale  1.009  1           1.005
## Year              1.025 16           1.001
```



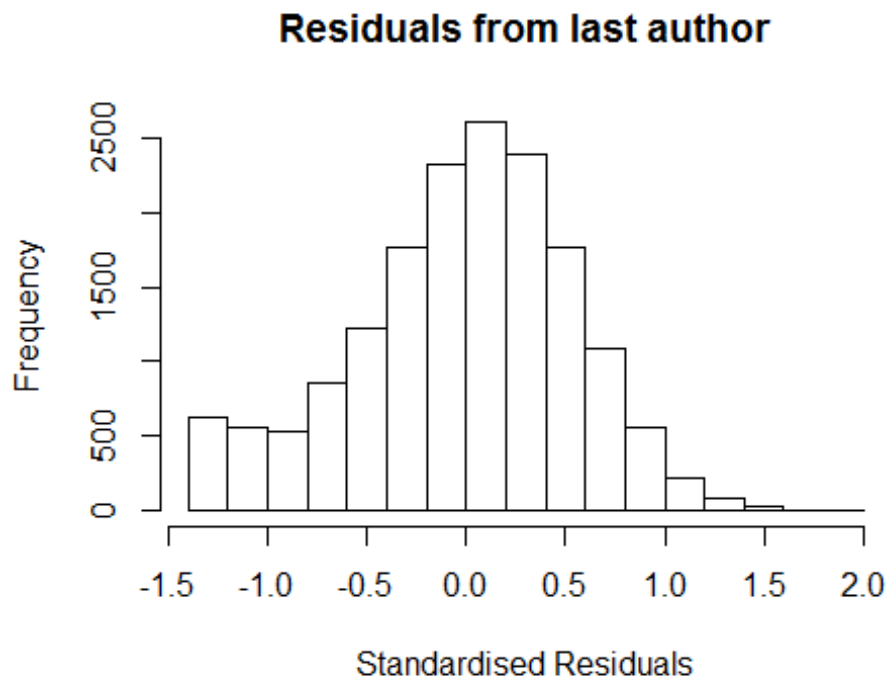

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.3266 -0.3555  0.0295  0.3590  1.8652
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.30357    0.02417   53.94  < 2e-16 ***
## FirstAuthorFemale1 -0.00239    0.00892   -0.27  0.78913
## LastAuthorFemale1  -0.05632    0.00975   -5.78  7.8e-09 ***
## Year1997            0.02298    0.03594    0.64  0.52254
## Year1998           -0.06027    0.03368   -1.79  0.07357 .
## Year1999           -0.07318    0.03205   -2.28  0.02242 *
## Year2000           -0.02540    0.03118   -0.81  0.41540
## Year2001            0.01218    0.03103    0.39  0.69465
## Year2002           -0.07735    0.02992   -2.59  0.00974 **
## Year2003           -0.04263    0.02993   -1.42  0.15441
## Year2004           -0.09094    0.02910   -3.12  0.00178 **
## Year2005           -0.07777    0.02920   -2.66  0.00774 **
```

```
## Year2006           -0.12035    0.02786   -4.32  1.6e-05 ***
## Year2007           -0.11442    0.02827   -4.05  5.2e-05 ***
## Year2008           -0.10813    0.02956   -3.66  0.00026 ***
## Year2009           -0.08469    0.02852   -2.97  0.00299 **
## Year2010           -0.11390    0.02855   -3.99  6.6e-05 ***
## Year2011           -0.09154    0.02903   -3.15  0.00162 **
## Year2012           -0.07750    0.02999   -2.58  0.00978 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.523
## Multiple R-squared:  0.00827,    Adjusted R-squared:  0.00719
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
##  1342 weights are ~= 1. The remaining 15262 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.177   0.859   0.949   0.896   0.986   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          6.02e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.019  1           1.009
## Year              1.019 16           1.001
```




```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.3152 -0.3540  0.0302  0.3589  1.8784
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.29051    0.02406   53.63  < 2e-16 ***
## FirstAuthorFemale1 -0.00970    0.00898   -1.08  0.27996
## Year1997            0.02473    0.03587    0.69  0.49068
## Year1998           -0.05942    0.03367   -1.76  0.07764 .
## Year1999           -0.07426    0.03199   -2.32  0.02028 *
## Year2000           -0.02593    0.03119   -0.83  0.40570
## Year2001            0.01057    0.03106    0.34  0.73369
## Year2002           -0.07991    0.02995   -2.67  0.00764 **
## Year2003           -0.04572    0.02995   -1.53  0.12697
## Year2004           -0.09238    0.02911   -3.17  0.00151 **
## Year2005           -0.07795    0.02920   -2.67  0.00761 **
## Year2006           -0.12242    0.02789   -4.39  1.1e-05 ***
```

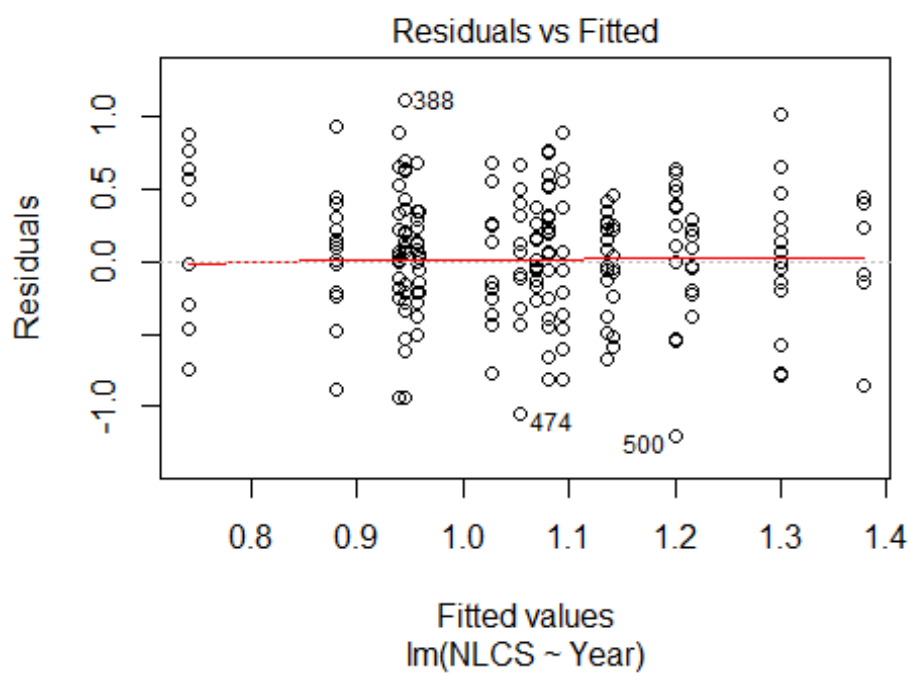
```
## Year2007           -0.11820    0.02826   -4.18  2.9e-05 ***
## Year2008           -0.11175    0.02958   -3.78  0.00016 ***
## Year2009           -0.08662    0.02852   -3.04  0.00239 **
## Year2010           -0.11600    0.02860   -4.06  5.0e-05 ***
## Year2011           -0.09492    0.02902   -3.27  0.00107 **
## Year2012           -0.08104    0.02999   -2.70  0.00690 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.523
## Multiple R-squared:  0.00604,    Adjusted R-squared:  0.00502
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
##  1422 weights are ~= 1. The remaining 15182 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.170   0.858   0.949   0.895   0.985   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          6.02e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##                   GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.008  1           1.004
## Year             1.008 16           1.000
```




```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.3257 -0.3565  0.0292  0.3584  1.8662
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.30277    0.02392   54.47  < 2e-16 ***
## LastAuthorFemale1 -0.05667    0.00979   -5.79  7.3e-09 ***
## Year1997            0.02293    0.03594    0.64  0.52357
## Year1998           -0.06020    0.03367   -1.79  0.07381 .
## Year1999           -0.07322    0.03205   -2.28  0.02235 *
## Year2000           -0.02554    0.03118   -0.82  0.41269
## Year2001            0.01204    0.03103    0.39  0.69816
## Year2002           -0.07742    0.02992   -2.59  0.00968 **
## Year2003           -0.04280    0.02993   -1.43  0.15280
## Year2004           -0.09111    0.02909   -3.13  0.00174 **
## Year2005           -0.07799    0.02919   -2.67  0.00755 **
## Year2006           -0.12057    0.02785   -4.33  1.5e-05 ***
```

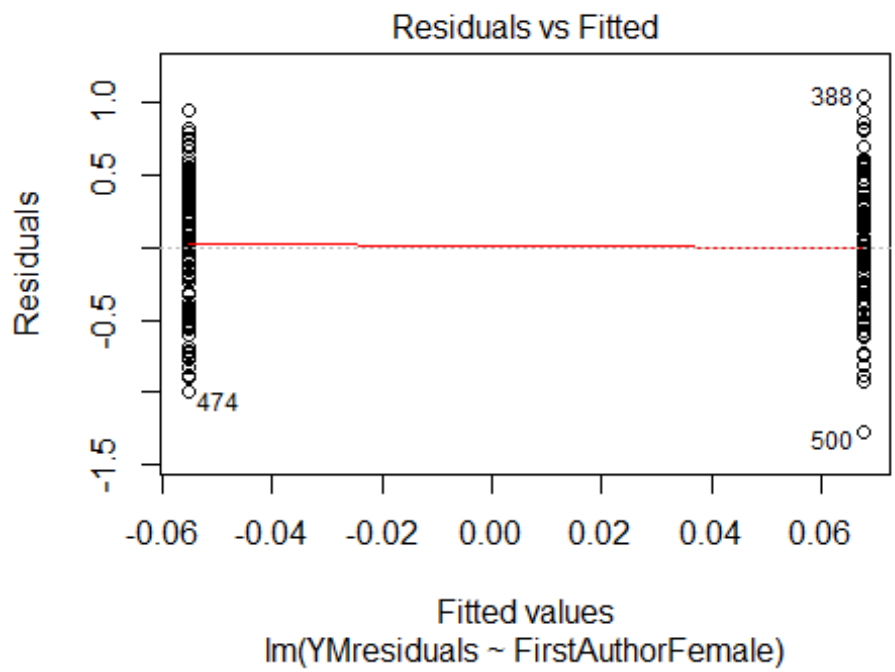
```
## Year2007          -0.11461    0.02828   -4.05  5.1e-05 ***
## Year2008          -0.10837    0.02955   -3.67  0.00025 ***
## Year2009          -0.08497    0.02851   -2.98  0.00288 **
## Year2010          -0.11423    0.02852   -4.01  6.2e-05 ***
## Year2011          -0.09184    0.02902   -3.16  0.00156 **
## Year2012          -0.07793    0.02994   -2.60  0.00925 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.523
## Multiple R-squared:  0.00826,    Adjusted R-squared:  0.00724
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
##  1343 weights are ~= 1. The remaining 15261 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.176   0.859   0.949   0.896   0.986   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          6.02e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Sample size for the above analysis:  16604"
## [1] ""
## [1] ""
## [1] "####################################"
## [1] "Analysis of AJSC 3401"
## [1] "####################################"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##   15   24   20   24   18   26   27   23   27   15   29   21   29   51   41
## 2011 2012
##   31   50
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    6   17   13   15   12   15   14   10   11    6   12   16   11   22   21
## 2011 2012
```

```
##   13   17
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    4   14   13   13   10   12    9    6   10    6   10   16   10   19   18
## 2011 2012
##   11   15
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
##  Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 36, df = 16, p-value = 0.003
```

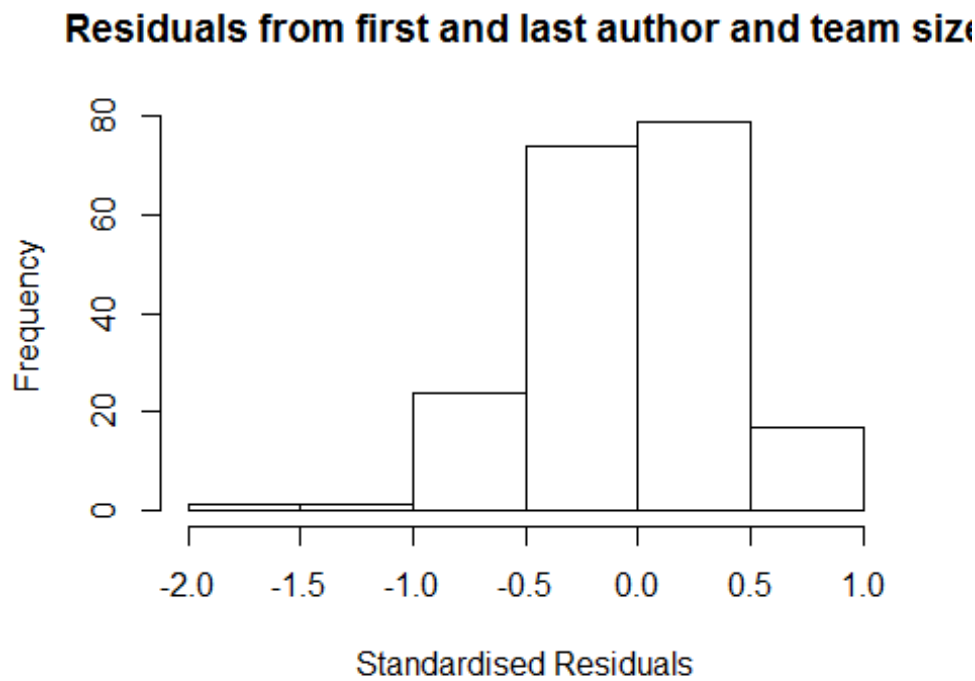



```
##
##  Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 0.0088, df = 1, p-value = 0.9

## [1] "Female first author team size 2018 geometric mean: 3.27906133744434"
## [1] "Male first author team size 2018 geometric mean: 2.7565637131945"

## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties
```

```
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 160, p-value = 0.5
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 3.02148909269044"
## [1] "Male last author team size 2018 geometric mean: 3.05326821786485"

## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties
```

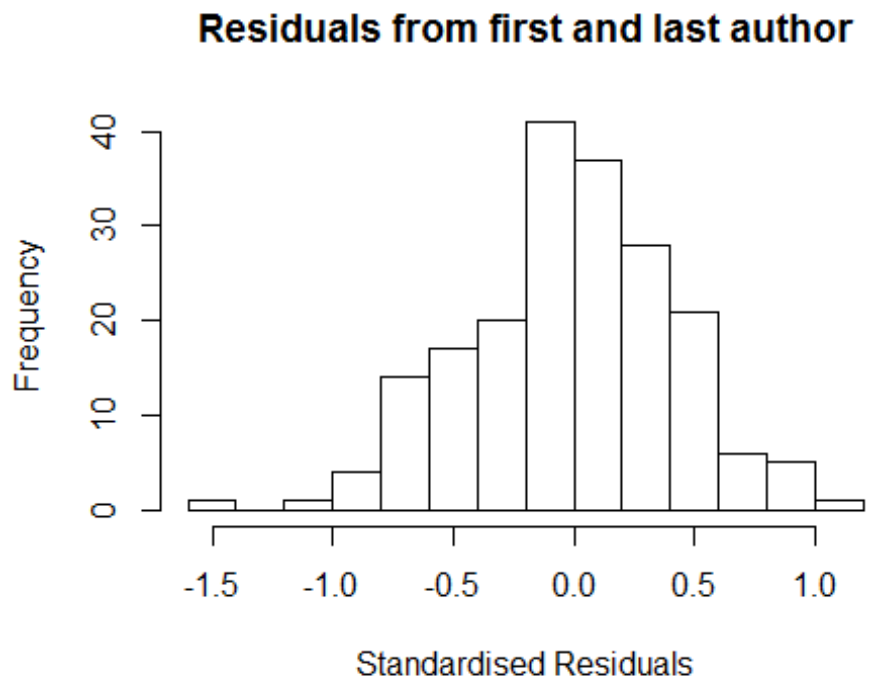



```
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 140, p-value = 0.8
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##                      GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale   3.312  1           1.820
## LastAuthorFemale    4.650  1           2.156
## UniqueAuthors      13.032  4           1.378
## Year               31.588 16           1.114
```



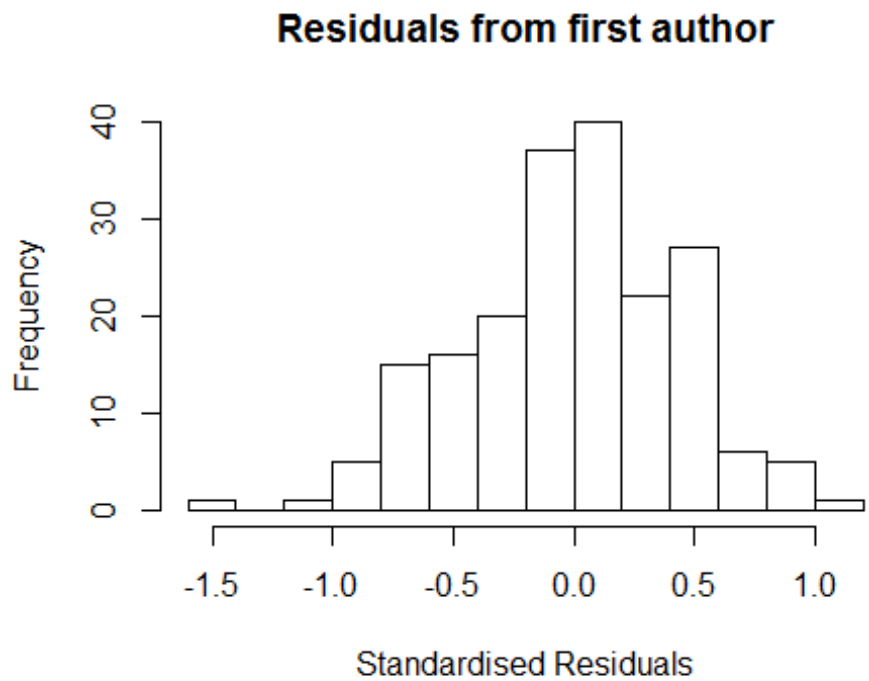

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
UniqueAuthors +
##     Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##     k.max = 1000))
##  \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.55803 -0.27254 -0.00528  0.25204  0.98644
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)          1.3456     0.1907    7.06    4e-11 ***
## FirstAuthorFemale1   0.1607     0.0953    1.69  0.09342 .
## LastAuthorFemale1    0.1016     0.1087    0.94  0.35100
## UniqueAuthors2       0.1518     0.1037    1.46  0.14500
## UniqueAuthors3       0.2517     0.0941    2.67  0.00821 **
## UniqueAuthors4       0.0178     0.1238    0.14  0.88575
## UniqueAuthors5       0.2044     0.1460    1.40  0.16332
## Year1997            -0.7559     0.2260   -3.35  0.00101 **
## Year1998            -0.3860     0.1862   -2.07  0.03965 *
## Year1999            -0.6115     0.1961   -3.12  0.00213 **
```

```
## Year2000            -0.8260     0.4414   -1.87  0.06298 .
## Year2001            -0.6024     0.1861   -3.24  0.00145 **
## Year2002            -0.5677     0.1744   -3.26  0.00136 **
## Year2003            -0.1329     0.1817   -0.73  0.46553
## Year2004            -0.3836     0.1970   -1.95  0.05319 .
## Year2005            -0.5903     0.1739   -3.39  0.00085 ***
## Year2006            -0.4969     0.2030   -2.45  0.01537 *
## Year2007            -0.2916     0.1970   -1.48  0.14066
## Year2008            -0.4885     0.2381   -2.05  0.04169 *
## Year2009            -0.6628     0.2310   -2.87  0.00462 **
## Year2010            -0.5165     0.2164   -2.39  0.01805 *
## Year2011            -0.5277     0.2113   -2.50  0.01346 *
## Year2012            -0.2542     0.2080   -1.22  0.22326
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.382
## Multiple R-squared:  0.244,  Adjusted R-squared:  0.148
## Convergence in 27 IRWLS iterations
##
## Robustness weights:
##  16 weights are ~= 1. The remaining 180 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0582  0.8590  0.9510  0.8880  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          5.10e-04          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 2.619  1           1.618
## LastAuthorFemale  2.834  1           1.683
## Year              2.970 16           1.035
```



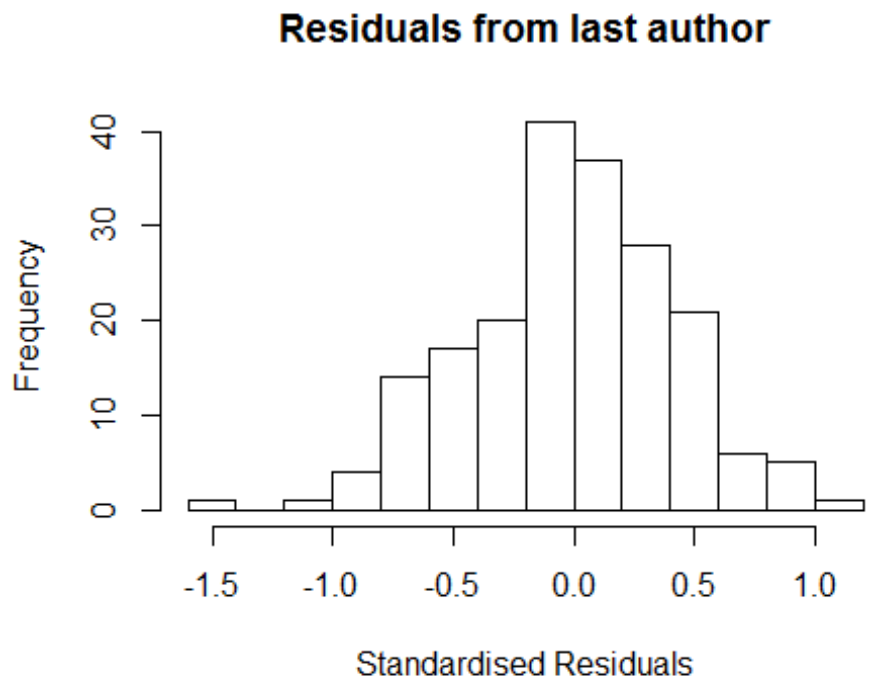

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.45599 -0.23754 -0.00181  0.28565  1.00385
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)           1.5078     0.1460   10.33  < 2e-16 ***
## FirstAuthorFemale1    0.1390     0.0896    1.55  0.12286
## LastAuthorFemale1     0.1010     0.0925    1.09  0.27637
## Year1997             -0.8469     0.1837   -4.61  7.7e-06 ***
## Year1998             -0.4585     0.1677   -2.73  0.00689 **
## Year1999             -0.6289     0.1818   -3.46  0.00068 ***
## Year2000             -0.9125     0.3173   -2.88  0.00452 **
## Year2001             -0.6336     0.1733   -3.66  0.00034 ***
## Year2002             -0.5571     0.1530   -3.64  0.00035 ***
## Year2003             -0.2149     0.1614   -1.33  0.18498
## Year2004             -0.4307     0.1773   -2.43  0.01615 *
## Year2005             -0.6677     0.1586   -4.21  4.1e-05 ***
```

```
## Year2006            -0.5610     0.1859   -3.02  0.00293 **
## Year2007            -0.3336     0.1762   -1.89  0.05999 .
## Year2008            -0.5313     0.2298   -2.31  0.02195 *
## Year2009            -0.6662     0.2090   -3.19  0.00169 **
## Year2010            -0.5417     0.2087   -2.60  0.01023 *
## Year2011            -0.5274     0.1814   -2.91  0.00410 **
## Year2012            -0.2918     0.1944   -1.50  0.13521
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.43
## Multiple R-squared:  0.187,  Adjusted R-squared:  0.105
## Convergence in 16 IRWLS iterations
##
## Robustness weights:
##  11 weights are ~= 1. The remaining 185 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.228   0.869   0.960   0.909   0.990   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          5.10e-04          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.847  1           1.359
## Year              1.847 16           1.019
```




```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.4450 -0.2711  0.0109  0.2801  1.0318
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)           1.5386     0.1610    9.55  < 2e-16 ***
## FirstAuthorFemale1    0.1855     0.0762    2.43  0.01589 *
## Year1997             -0.8226     0.2058   -4.00  9.4e-05 ***
## Year1998             -0.4415     0.1844   -2.39  0.01771 *
## Year1999             -0.6555     0.1919   -3.42  0.00079 ***
## Year2000             -0.9055     0.3072   -2.95  0.00363 **
## Year2001             -0.6557     0.1886   -3.48  0.00064 ***
## Year2002             -0.5596     0.1737   -3.22  0.00151 **
## Year2003             -0.2367     0.1742   -1.36  0.17599
## Year2004             -0.4544     0.1897   -2.40  0.01762 *
## Year2005             -0.6723     0.1734   -3.88  0.00015 ***
## Year2006             -0.5783     0.1988   -2.91  0.00408 **
```

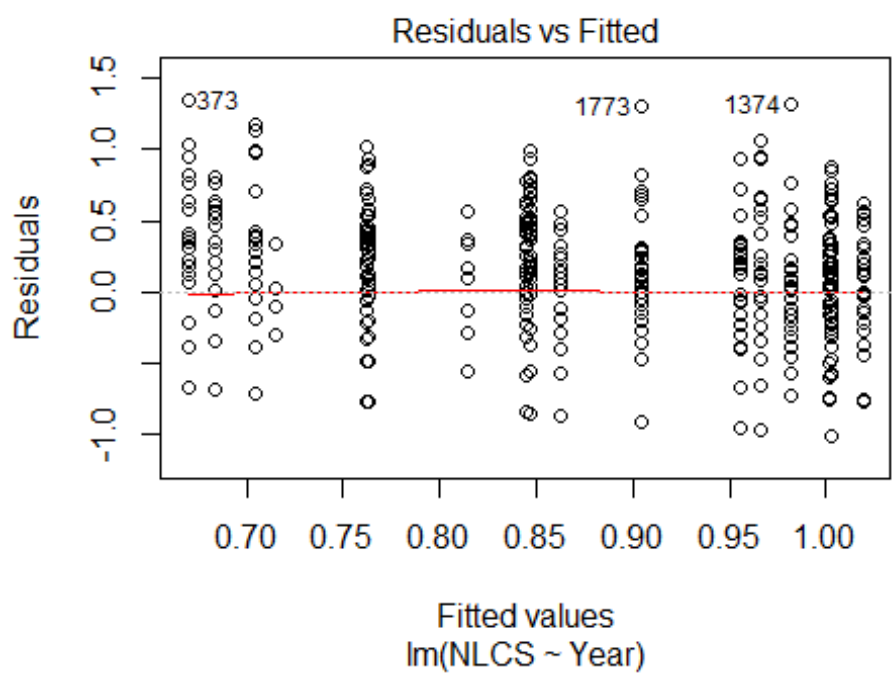
```
## Year2007            -0.3545     0.1921   -1.84  0.06670 .
## Year2008            -0.5336     0.2380   -2.24  0.02617 *
## Year2009            -0.6980     0.2174   -3.21  0.00157 **
## Year2010            -0.5566     0.2247   -2.48  0.01416 *
## Year2011            -0.5264     0.1995   -2.64  0.00906 **
## Year2012            -0.2792     0.2124   -1.31  0.19045
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.423
## Multiple R-squared:  0.183,  Adjusted R-squared:  0.105
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
##  13 weights are ~= 1. The remaining 183 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.220   0.859   0.955   0.905   0.990   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          5.10e-04          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##                   GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 2.051  1           1.432
## Year             2.051 16           1.023
```




```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.4213 -0.2512  0.0334  0.2882  1.0935
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)          1.5398     0.1084   14.21  < 2e-16 ***
## LastAuthorFemale1    0.1719     0.0777    2.21  0.02828 *
## Year1997            -0.8339     0.1495   -5.58  9.0e-08 ***
## Year1998            -0.4910     0.1320   -3.72  0.00027 ***
## Year1999            -0.6371     0.1517   -4.20  4.2e-05 ***
## Year2000            -0.9217     0.3175   -2.90  0.00416 **
## Year2001            -0.6143     0.1423   -4.32  2.6e-05 ***
## Year2002            -0.5724     0.1150   -4.98  1.5e-06 ***
## Year2003            -0.2355     0.1255   -1.88  0.06209 .
## Year2004            -0.4318     0.1510   -2.86  0.00475 **
## Year2005            -0.6659     0.1301   -5.12  8.0e-07 ***
## Year2006            -0.5565     0.1712   -3.25  0.00138 **
```

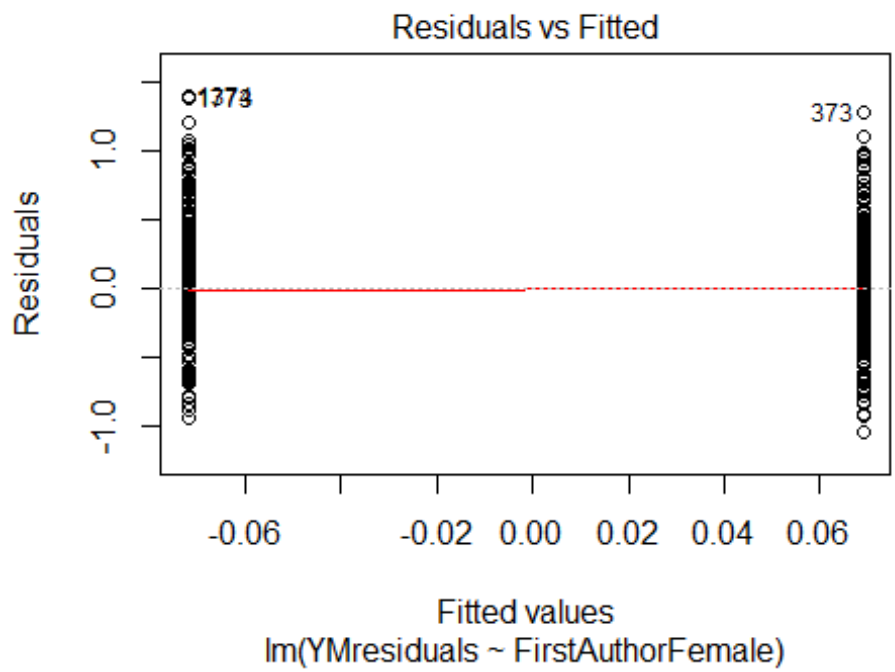
```
## Year2007           -0.3163     0.1512   -2.09  0.03783 *
## Year2008           -0.5478     0.2432   -2.25  0.02551 *
## Year2009           -0.6323     0.1871   -3.38  0.00089 ***
## Year2010           -0.5305     0.1901   -2.79  0.00583 **
## Year2011           -0.5472     0.1463   -3.74  0.00025 ***
## Year2012           -0.2904     0.1599   -1.82  0.07114 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.419
## Multiple R-squared:  0.172,  Adjusted R-squared:  0.0934
## Convergence in 15 IRWLS iterations
##
## Robustness weights:
##  8 weights are ~= 1. The remaining 188 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.227   0.853   0.958   0.905   0.991   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          5.10e-04          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Sample size for the above analysis:  196"
## [1] ""
## [1] ""
## [1] "####################################"
## [1] "Analysis of AJSC 3402"
## [1] "####################################"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  219  179  167  139  132  107  125   96   71   98  160  130  134  166  173
## 2011 2012
##  144  163
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##   57   47   31   38   10    5   33   31   28   41   62   39   37   29   48
## 2011 2012
```

```
##   36   51
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##   44   37   20   24   10    4   28   29   26   39   50   37   33   23   43
## 2011 2012
##   32   44
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
##  Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 28, df = 16, p-value = 0.03
```

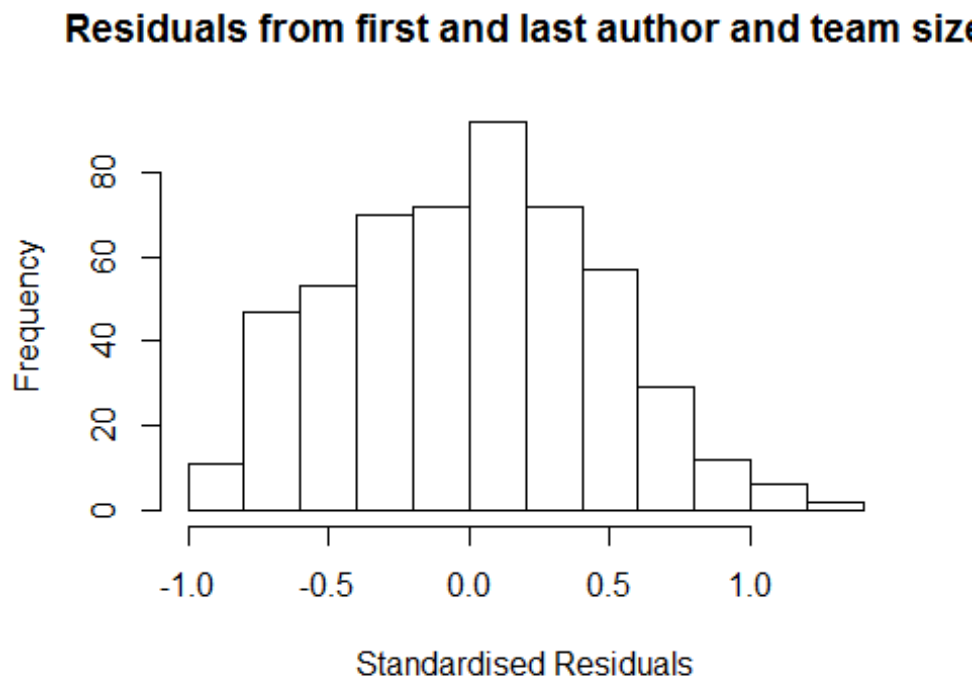



```
##
##  Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 2.1, df = 1, p-value = 0.1

## [1] "Female first author team size 2018 geometric mean: 4.19928605479686"
## [1] "Male first author team size 2018 geometric mean: 4.24519472666281"

## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties
```

```
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 240, p-value = 0.8
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 3.48273355601102"
## [1] "Male last author team size 2018 geometric mean: 4.89080014829878"

## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties
```

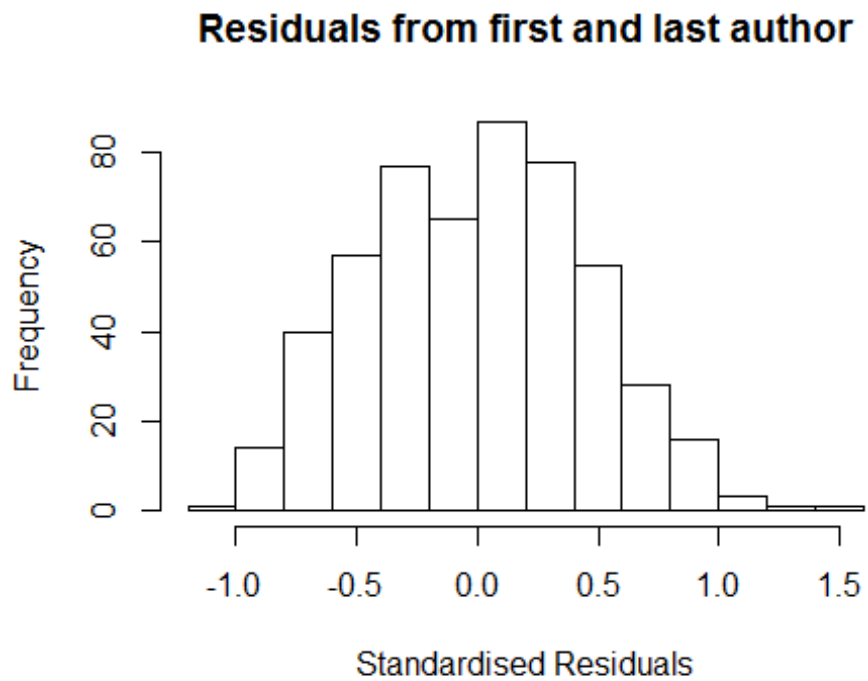



```
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 200, p-value = 0.03
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##                     GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.339  1           1.157
## LastAuthorFemale  1.636  1           1.279
## UniqueAuthors     1.873  4           1.082
## Year              2.287 16           1.026
```



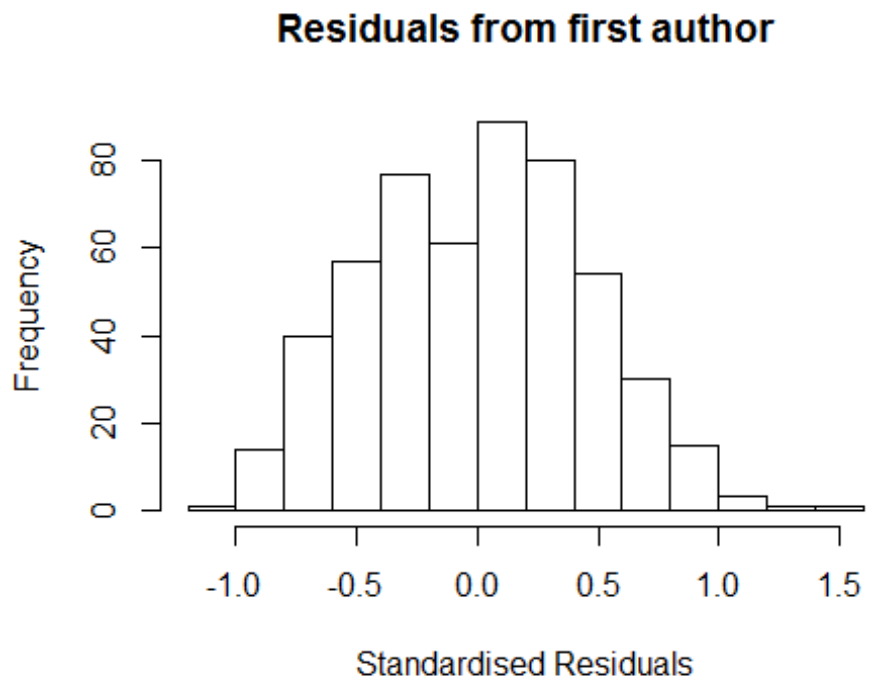

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
UniqueAuthors +
##     Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##     k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -0.9773 -0.3475  0.0228  0.3089  1.3796
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)          0.6465     0.1106    5.85  9.1e-09 ***
## FirstAuthorFemale1   0.1192     0.0459    2.60   0.0096 **
## LastAuthorFemale1    0.0112     0.0502    0.22   0.8236
## UniqueAuthors2       0.0884     0.0698    1.27   0.2060
## UniqueAuthors3       0.1223     0.0644    1.90   0.0582 .
## UniqueAuthors4       0.1415     0.0640    2.21   0.0274 *
## UniqueAuthors5       0.2706     0.0622    4.35  1.6e-05 ***
## Year1997            -0.2221     0.1518   -1.46   0.1439
## Year1998            -0.0699     0.1562   -0.45   0.6549
## Year1999            -0.1421     0.1446   -0.98   0.3265
```

```
## Year2000            -0.0258     0.1416   -0.18   0.8554
## Year2001            -0.2109     0.1467   -1.44   0.1511
## Year2002             0.1921     0.1284    1.50   0.1352
## Year2003             0.2202     0.1266    1.74   0.0826 .
## Year2004             0.1925     0.1275    1.51   0.1318
## Year2005             0.1594     0.1248    1.28   0.2022
## Year2006             0.2293     0.1312    1.75   0.0812 .
## Year2007             0.0996     0.1379    0.72   0.4708
## Year2008            -0.0216     0.1470   -0.15   0.8833
## Year2009             0.0698     0.1264    0.55   0.5810
## Year2010            -0.1711     0.1272   -1.34   0.1792
## Year2011             0.0701     0.1477    0.47   0.6351
## Year2012            -0.2213     0.1340   -1.65   0.0994 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.478
## Multiple R-squared:  0.153,  Adjusted R-squared:  0.116
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
##  31 weights are ~= 1. The remaining 492 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.385   0.881   0.952   0.919   0.986   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.91e-04          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.261  1           1.123
## LastAuthorFemale  1.436  1           1.199
## Year              1.383 16           1.010
```



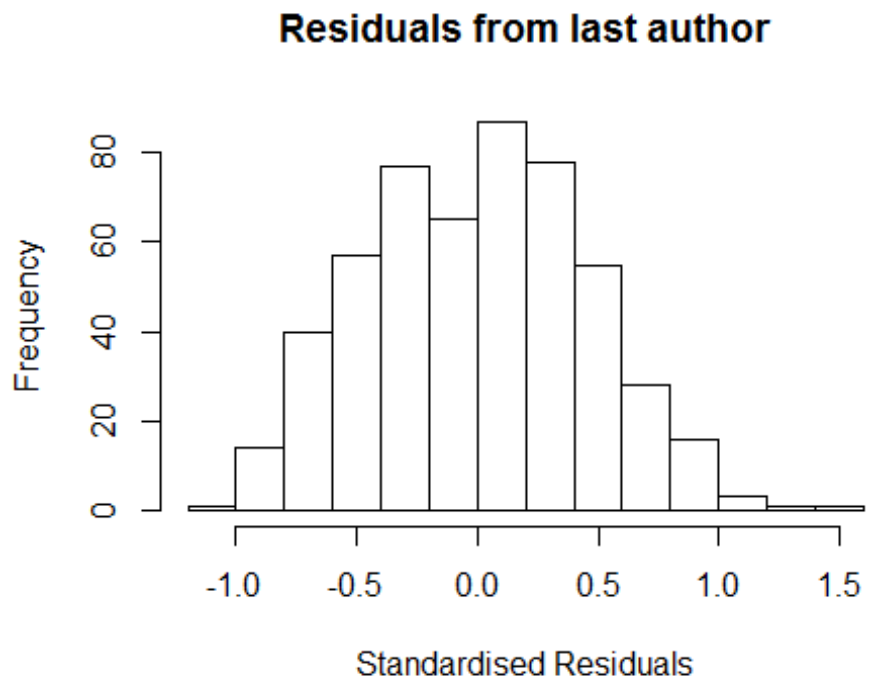

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.0175 -0.3492  0.0297  0.3314  1.4114
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)           0.7178     0.1109    6.47  2.3e-10 ***
## FirstAuthorFemale1    0.1487     0.0459    3.24   0.0013 **
## LastAuthorFemale1    -0.0174     0.0482   -0.36   0.7178
## Year1997             -0.2174     0.1562   -1.39   0.1646
## Year1998             -0.0471     0.1540   -0.31   0.7597
## Year1999             -0.1186     0.1479   -0.80   0.4229
## Year2000              0.0415     0.1538    0.27   0.7875
## Year2001             -0.1938     0.1321   -1.47   0.1428
## Year2002              0.1931     0.1355    1.42   0.1549
## Year2003              0.2622     0.1303    2.01   0.0447 *
## Year2004              0.2613     0.1337    1.95   0.0513 .
## Year2005              0.1698     0.1293    1.31   0.1899
```

```
## Year2006             0.2361     0.1311    1.80   0.0722 .
## Year2007             0.1045     0.1413    0.74   0.4600
## Year2008             0.0105     0.1457    0.07   0.9428
## Year2009             0.1017     0.1269    0.80   0.4232
## Year2010            -0.1516     0.1285   -1.18   0.2388
## Year2011             0.1510     0.1455    1.04   0.2997
## Year2012            -0.1298     0.1388   -0.94   0.3502
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.481
## Multiple R-squared:  0.121,  Adjusted R-squared:  0.0894
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
##  37 weights are ~= 1. The remaining 486 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.370   0.877   0.948   0.916   0.982   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.91e-04          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.118  1           1.057
## Year              1.118 16           1.003
```




```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.0086 -0.3533  0.0314  0.3238  1.4166
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)          0.71746    0.11134    6.44  2.7e-10 ***
## FirstAuthorFemale1   0.14149    0.04326    3.27   0.0011 **
## Year1997            -0.22167    0.15563   -1.42   0.1550
## Year1998            -0.04902    0.15341   -0.32   0.7494
## Year1999            -0.12576    0.14519   -0.87   0.3868
## Year2000             0.03941    0.15327    0.26   0.7972
## Year2001            -0.19250    0.13181   -1.46   0.1448
## Year2002             0.19285    0.13581    1.42   0.1562
## Year2003             0.25677    0.12854    2.00   0.0463 *
## Year2004             0.26115    0.13379    1.95   0.0515 .
## Year2005             0.16495    0.12691    1.30   0.1943
## Year2006             0.23231    0.12963    1.79   0.0737 .
```

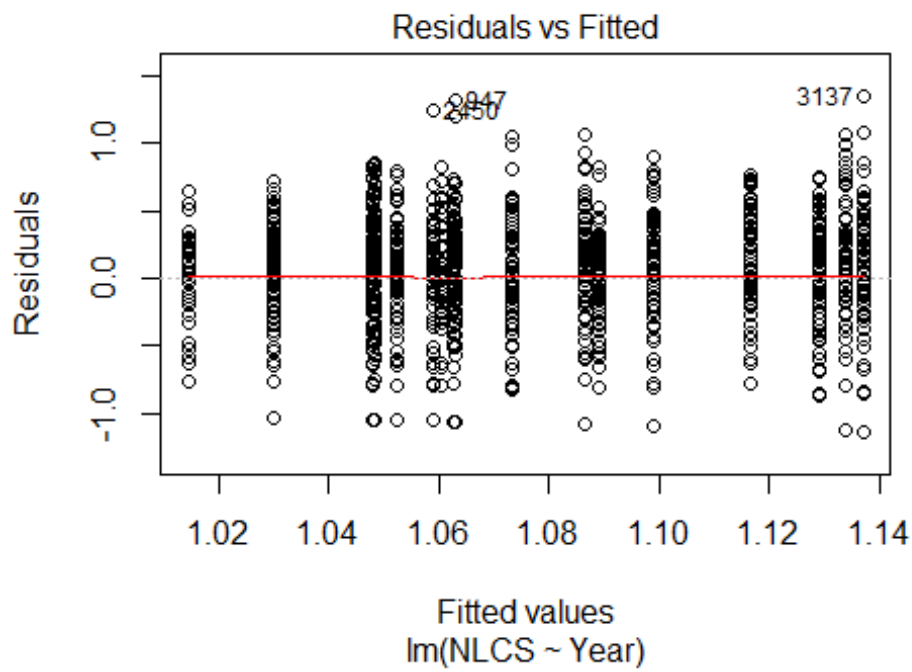
```
## Year2007            0.10145    0.13951    0.73   0.4675
## Year2008            0.00702    0.14480    0.05   0.9614
## Year2009            0.10061    0.12758    0.79   0.4307
## Year2010           -0.15338    0.12802   -1.20   0.2315
## Year2011            0.14962    0.14547    1.03   0.3042
## Year2012           -0.13499    0.13564   -1.00   0.3201
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.479
## Multiple R-squared:  0.121,  Adjusted R-squared:  0.0912
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
##  35 weights are ~= 1. The remaining 488 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.362   0.877   0.948   0.916   0.982   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.91e-04          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##                   GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.265  1           1.125
## Year             1.265 16           1.007
```




```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.0172 -0.3297  0.0275  0.3121  1.3679
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)        0.766244   0.104459    7.34  8.9e-13 ***
## LastAuthorFemale1  0.050032   0.045873    1.09    0.276
## Year1997          -0.217337   0.150696   -1.44    0.150
## Year1998          -0.055273   0.154903   -0.36    0.721
## Year1999          -0.140521   0.147286   -0.95    0.341
## Year2000           0.031700   0.155369    0.20    0.838
## Year2001          -0.148468   0.122336   -1.21    0.225
## Year2002           0.189362   0.134914    1.40    0.161
## Year2003           0.249469   0.127894    1.95    0.052 .
## Year2004           0.267652   0.128871    2.08    0.038 *
## Year2005           0.164906   0.123286    1.34    0.182
## Year2006           0.251001   0.127916    1.96    0.050 .
```

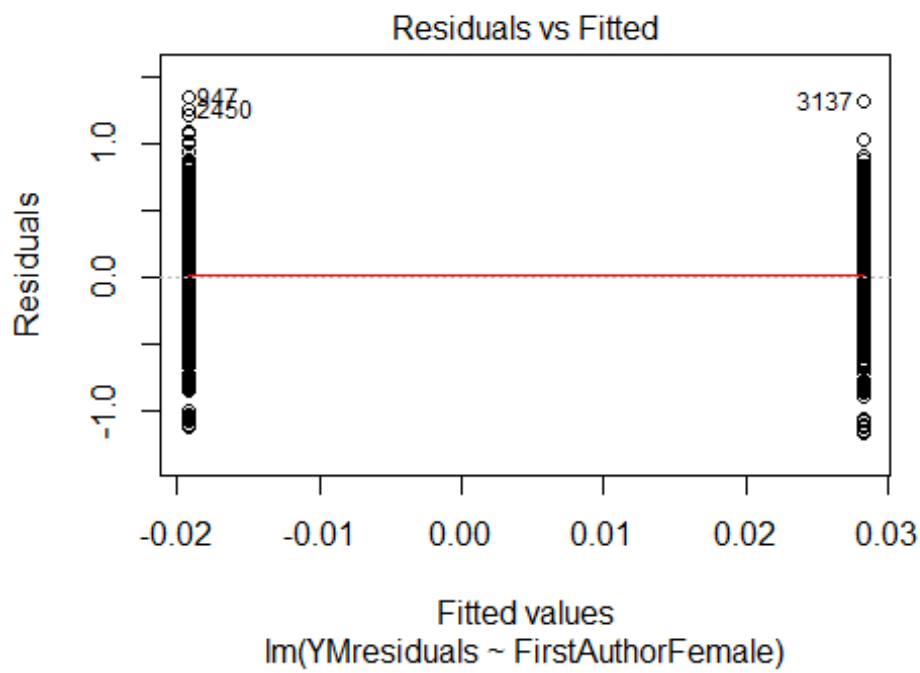
```
## Year2007           0.122756   0.136886    0.90    0.370
## Year2008           0.000236   0.144408    0.00    0.999
## Year2009           0.112726   0.126945    0.89    0.375
## Year2010          -0.153732   0.125206   -1.23    0.220
## Year2011           0.163282   0.145762    1.12    0.263
## Year2012          -0.139535   0.136096   -1.03    0.306
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.482
## Multiple R-squared:  0.105,  Adjusted R-squared:  0.0745
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
##  38 weights are ~= 1. The remaining 485 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.401   0.866   0.950   0.915   0.984   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.91e-04          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Sample size for the above analysis:  523"
## [1] ""
## [1] ""
## [1] "####################################"
## [1] "Analysis of AJSC 3403"
## [1] "####################################"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  305  254  235  199  243  226  256  258  231  215  213  243  217  253  261
## 2011 2012
##  258  244
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##   73   75   76   58   50   38   94  102   87   63   89   71   80   88   94
## 2011 2012
```

```
##   87   97
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##   62   62   62   48   38   34   78   81   80   55   80   61   62   74   74
## 2011 2012
##   78   88
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
##  Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 36, df = 16, p-value = 0.003
```

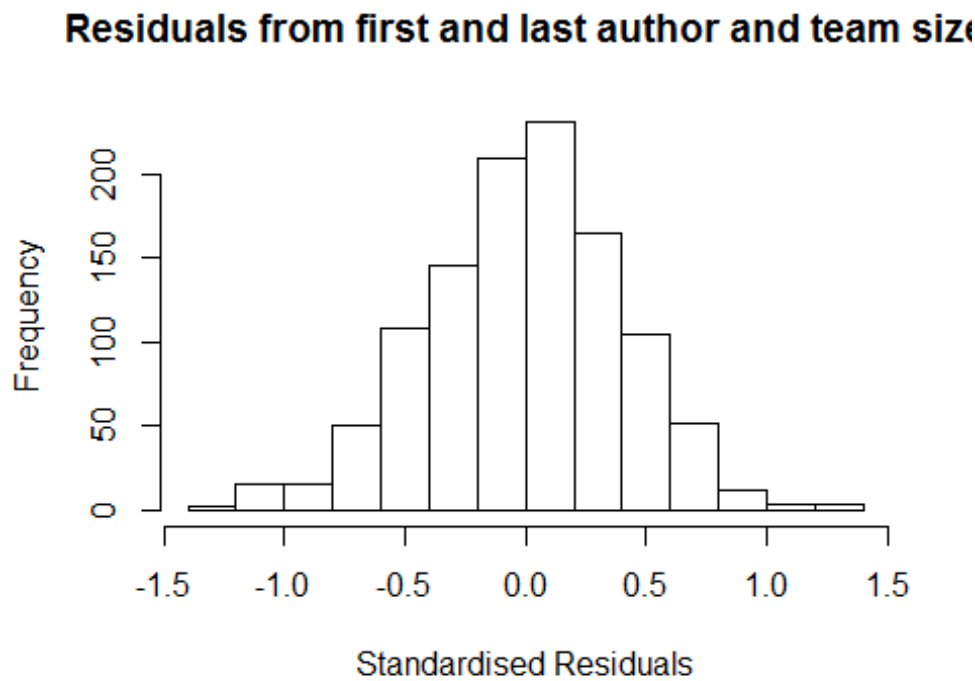



```
##
##  Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 0.099, df = 1, p-value = 0.8

## [1] "Female first author team size 2018 geometric mean: 3.5866251447166"
## [1] "Male first author team size 2018 geometric mean: 3.26433548298085"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 760, p-value = 0.6
```

```
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 2.999871443351"
## [1] "Male last author team size 2018 geometric mean: 4.07591667087584"

## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties
```

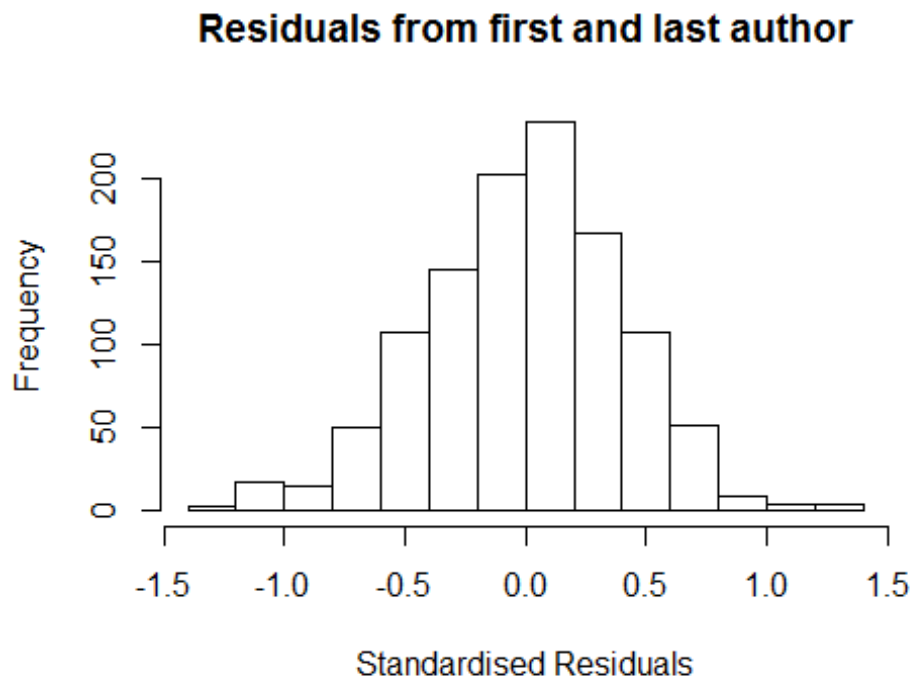



```
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 620, p-value = 0.03
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##                     GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.121  1           1.059
## LastAuthorFemale  1.074  1           1.036
## UniqueAuthors     1.356  4           1.039
## Year              1.472 16           1.012
```



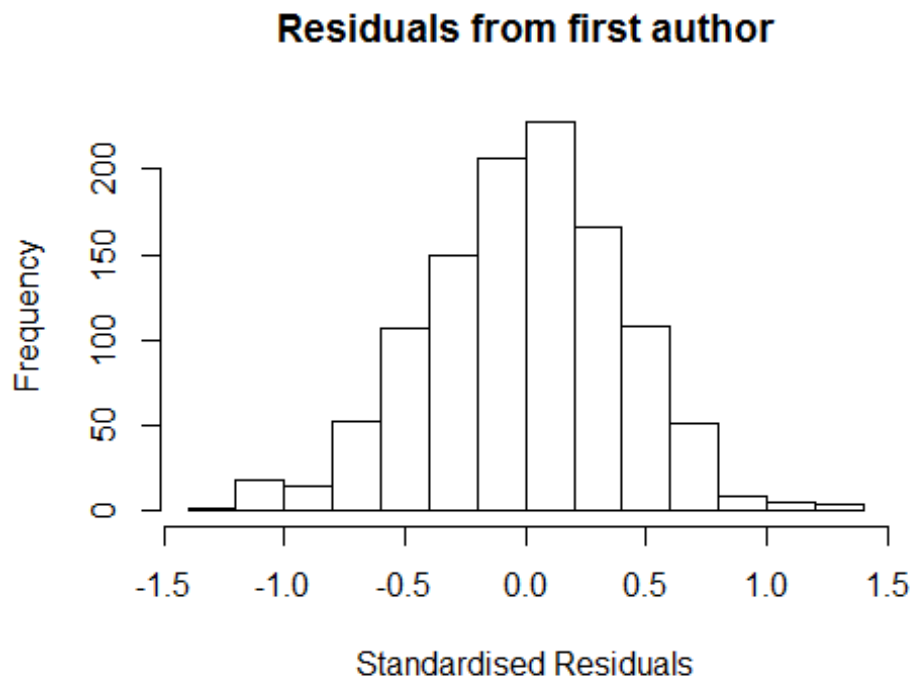

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
UniqueAuthors +
##     Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##     k.max = 1000))
##  \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.25839 -0.26297  0.00965  0.26137  1.39628
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.026725   0.065712   15.62  < 2e-16 ***
## FirstAuthorFemale1  0.049090   0.026056    1.88  0.05983 .
## LastAuthorFemale1   0.102409   0.028697    3.57  0.00037 ***
## UniqueAuthors2      0.027648   0.060089    0.46  0.64552
## UniqueAuthors3      0.015579   0.057292    0.27  0.78573
## UniqueAuthors4      0.031745   0.056681    0.56  0.57555
## UniqueAuthors5      0.026582   0.055450    0.48  0.63176
## Year1997           -0.046190   0.072713   -0.64  0.52540
## Year1998           -0.000681   0.078190   -0.01  0.99306
## Year1999           -0.047580   0.081559   -0.58  0.55975
```

```
## Year2000           -0.059386   0.075560   -0.79  0.43207
## Year2001           -0.087950   0.076726   -1.15  0.25193
## Year2002           -0.012448   0.070019   -0.18  0.85893
## Year2003           -0.041930   0.063630   -0.66  0.51006
## Year2004            0.005621   0.059766    0.09  0.92509
## Year2005           -0.019752   0.076075   -0.26  0.79519
## Year2006            0.048805   0.071293    0.68  0.49376
## Year2007            0.080164   0.082959    0.97  0.33410
## Year2008           -0.000963   0.072586   -0.01  0.98942
## Year2009            0.036647   0.072081    0.51  0.61126
## Year2010            0.030268   0.068941    0.44  0.66072
## Year2011            0.010885   0.076817    0.14  0.88734
## Year2012            0.041646   0.067677    0.62  0.53844
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.398
## Multiple R-squared:  0.0307, Adjusted R-squared:  0.0112
## Convergence in 15 IRWLS iterations
##
## Robustness weights:
##  117 weights are ~= 1. The remaining 1000 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.193   0.865   0.949   0.901   0.984   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          8.95e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.097  1           1.048
## LastAuthorFemale  1.066  1           1.032
## Year              1.112 16           1.003
```



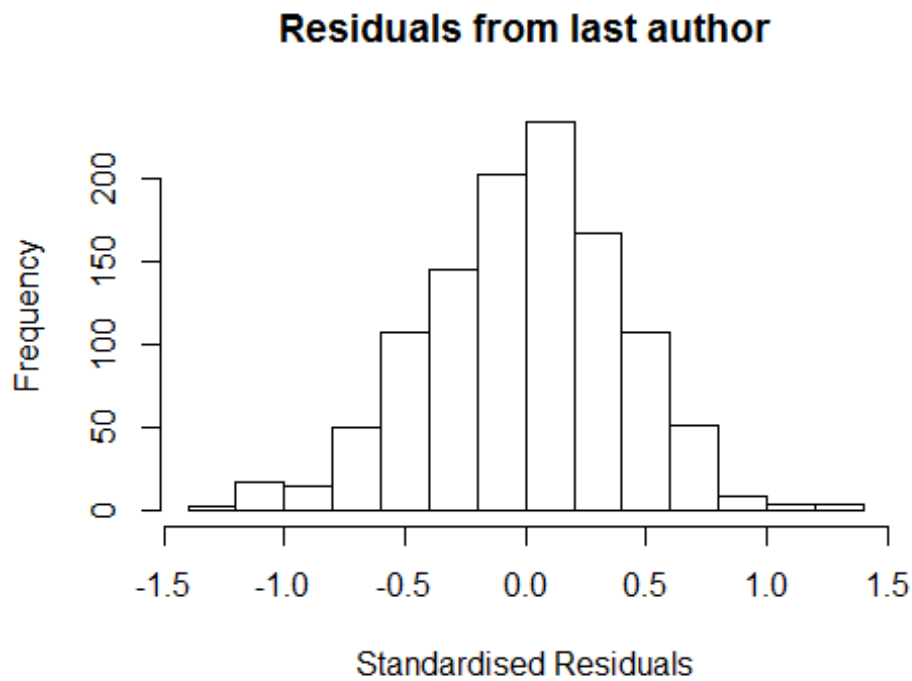

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.2824 -0.2641  0.0119  0.2596  1.3879
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)          1.04694    0.05064   20.67  < 2e-16 ***
## FirstAuthorFemale1   0.05082    0.02583    1.97  0.04938 *
## LastAuthorFemale1    0.10075    0.02853    3.53  0.00043 ***
## Year1997            -0.04279    0.07197   -0.59  0.55226
## Year1998             0.00256    0.07787    0.03  0.97380
## Year1999            -0.04386    0.08137   -0.54  0.58997
## Year2000            -0.05647    0.07535   -0.75  0.45371
## Year2001            -0.08605    0.07644   -1.13  0.26051
## Year2002            -0.01017    0.06916   -0.15  0.88315
## Year2003            -0.03839    0.06316   -0.61  0.54339
## Year2004             0.00767    0.05935    0.13  0.89717
## Year2005            -0.01657    0.07541   -0.22  0.82608
```

```
## Year2006            0.04770    0.07079    0.67  0.50058
## Year2007            0.08393    0.08179    1.03  0.30502
## Year2008           -0.00353    0.07194   -0.05  0.96091
## Year2009            0.04040    0.07121    0.57  0.57060
## Year2010            0.03268    0.06828    0.48  0.63225
## Year2011            0.01306    0.07549    0.17  0.86264
## Year2012            0.04330    0.06722    0.64  0.51956
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.399
## Multiple R-squared:  0.0301, Adjusted R-squared:  0.0142
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
##  114 weights are ~= 1. The remaining 1003 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.202   0.867   0.949   0.902   0.985   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          8.95e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.069  1           1.034
## Year              1.069 16           1.002
```




```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.2293 -0.2626  0.0107  0.2694  1.3760
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)          1.05569    0.05003   21.10   <2e-16 ***
## FirstAuthorFemale1   0.07308    0.02582    2.83   0.0047 **
## Year1997            -0.02425    0.07184   -0.34   0.7358
## Year1998             0.01030    0.07725    0.13   0.8939
## Year1999            -0.04066    0.08165   -0.50   0.6186
## Year2000            -0.05266    0.07744   -0.68   0.4966
## Year2001            -0.06731    0.07594   -0.89   0.3756
## Year2002             0.00103    0.06869    0.01   0.9881
## Year2003            -0.02628    0.06300   -0.42   0.6766
## Year2004             0.01190    0.05883    0.20   0.8398
## Year2005             0.00446    0.07364    0.06   0.9517
## Year2006             0.05927    0.07050    0.84   0.4007
```

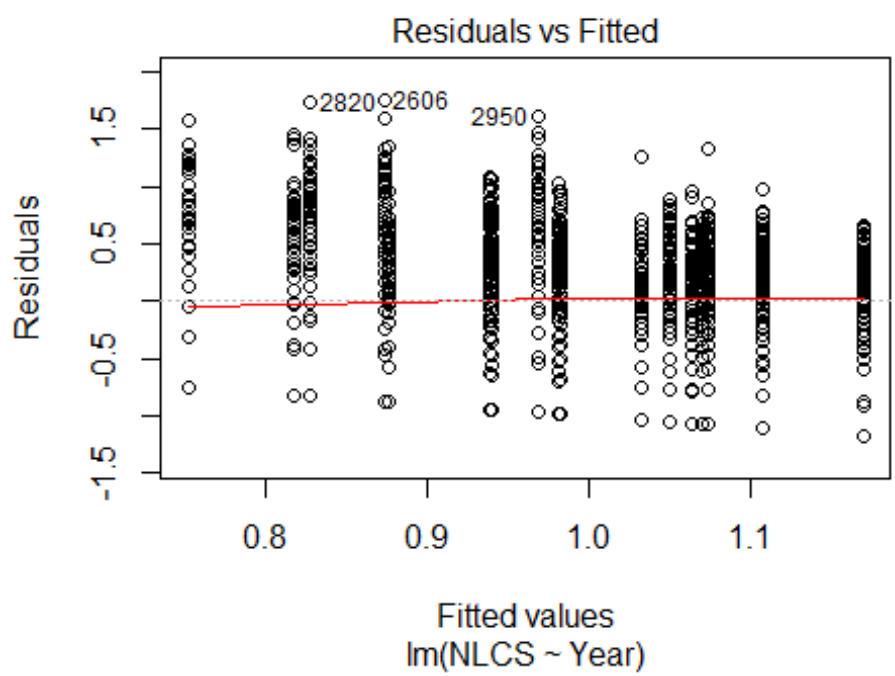
```
## Year2007            0.10057    0.08147    1.23   0.2173
## Year2008            0.00977    0.07148    0.14   0.8913
## Year2009            0.04001    0.07044    0.57   0.5701
## Year2010            0.04104    0.06805    0.60   0.5465
## Year2011            0.02596    0.07585    0.34   0.7323
## Year2012            0.05865    0.06669    0.88   0.3793
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.402
## Multiple R-squared:  0.0185, Adjusted R-squared:  0.0033
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
##  98 weights are ~= 1. The remaining 1019 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.217   0.870   0.952   0.904   0.986   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          8.95e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##                   GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.039  1           1.019
## Year             1.039 16           1.001
```



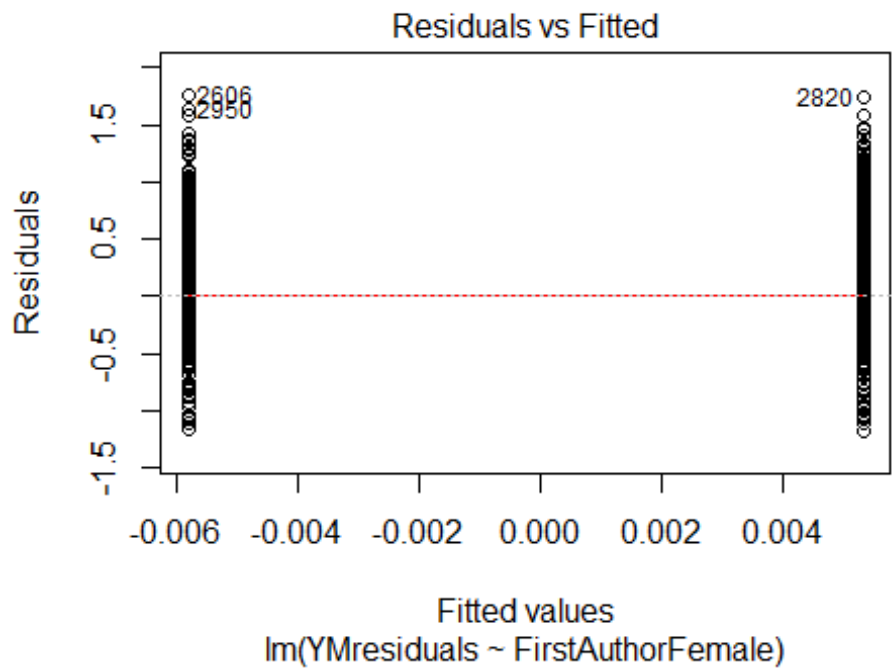

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.2691 -0.2667  0.0148  0.2623  1.3659
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.06193    0.04908   21.64  < 2e-16 ***
## LastAuthorFemale1   0.11413    0.02842    4.02  6.3e-05 ***
## Year1997           -0.04919    0.07200   -0.68     0.49
## Year1998            0.00257    0.07747    0.03     0.97
## Year1999           -0.03684    0.08076   -0.46     0.65
## Year2000           -0.05928    0.07471   -0.79     0.43
## Year2001           -0.09039    0.07620   -1.19     0.24
## Year2002           -0.00503    0.06896   -0.07     0.94
## Year2003           -0.03747    0.06278   -0.60     0.55
## Year2004            0.01080    0.05888    0.18     0.85
## Year2005           -0.02161    0.07477   -0.29     0.77
## Year2006            0.05091    0.07098    0.72     0.47
```

```
## Year2007           0.09300    0.08154    1.14     0.25
## Year2008          -0.00280    0.07207   -0.04     0.97
## Year2009           0.04403    0.07138    0.62     0.54
## Year2010           0.03742    0.06773    0.55     0.58
## Year2011           0.01322    0.07495    0.18     0.86
## Year2012           0.04732    0.06679    0.71     0.48
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.402
## Multiple R-squared:  0.0264, Adjusted R-squared:  0.0113
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  110 weights are ~= 1. The remaining 1007 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.224   0.874   0.951   0.903   0.985   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          8.95e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Sample size for the above analysis:  1117"
## [1] ""
## [1] ""
## [1] "####################################"
## [1] "Analysis of AJSC 3404"
## [1] "####################################"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  243  196  175  185  197  189  192  162  168  131  134  137  184  213  204
## 2011 2012
##  242  202
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  134  123  107   79   88   76  130  107  131   94   98   80  113  138  155
## 2011 2012
```

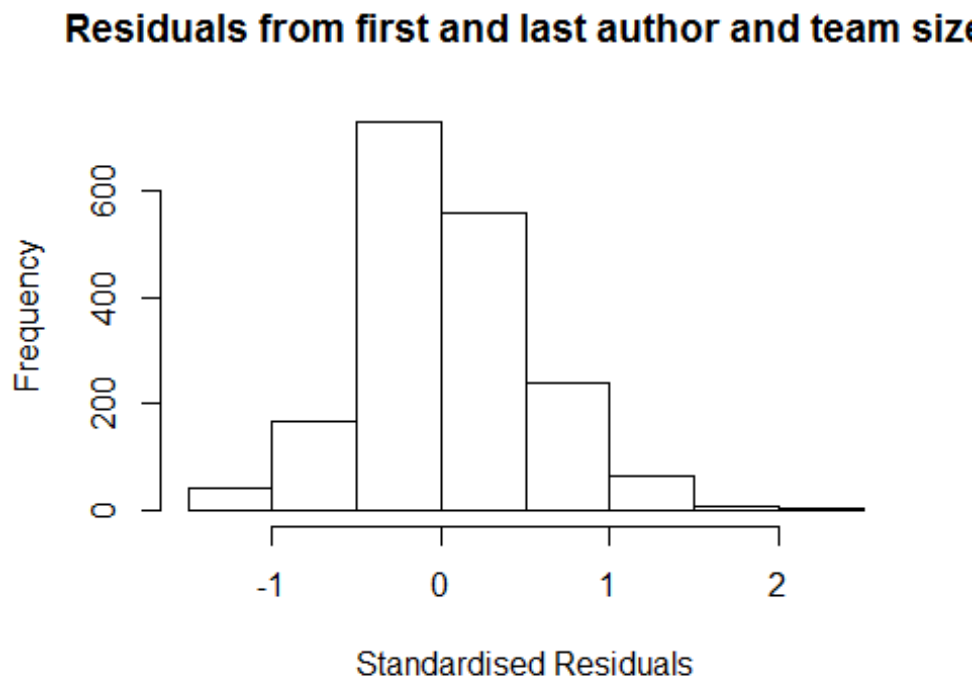
```
##  181  129
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  125  113   93   69   82   64  118   95  123   86   91   76  107  131  144
## 2011 2012
##  170  123
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
##  Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 180, df = 16, p-value <2e-16
```




```
##
##  Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 1.6, df = 1, p-value = 0.2
```

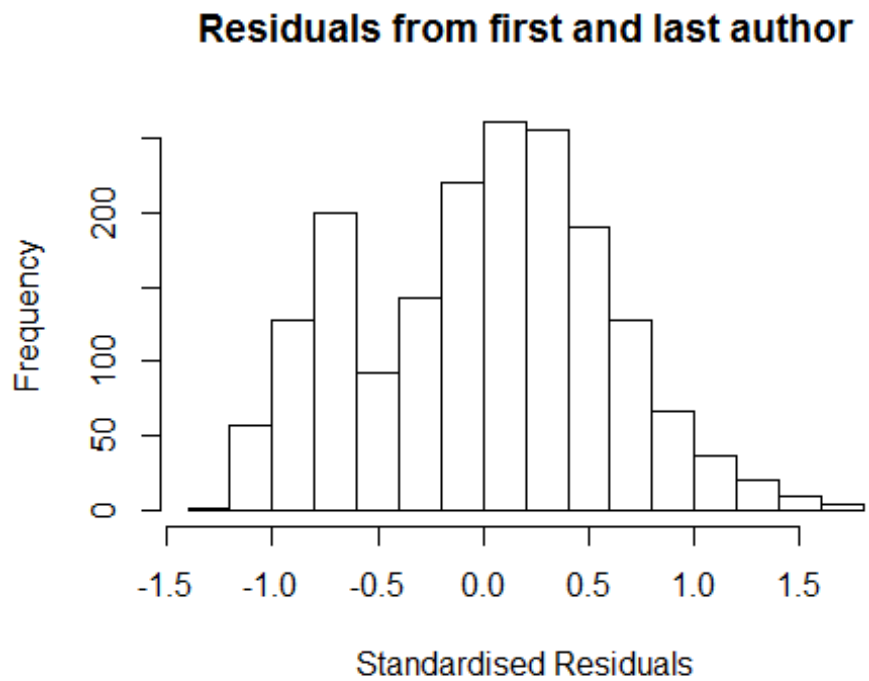



```
## [1] "Female first author team size 2018 geometric mean: 3.63257083630712"
## [1] "Male first author team size 2018 geometric mean: 3.0644413075689"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 1200, p-value = 0.4
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 3.10734903166882"
## [1] "Male last author team size 2018 geometric mean: 3.85884113534239"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 1100, p-value = 0.1
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.149  1           1.072
## LastAuthorFemale  1.163  1           1.079
## UniqueAuthors     1.492  4           1.051
## Year              1.588 16           1.015
```



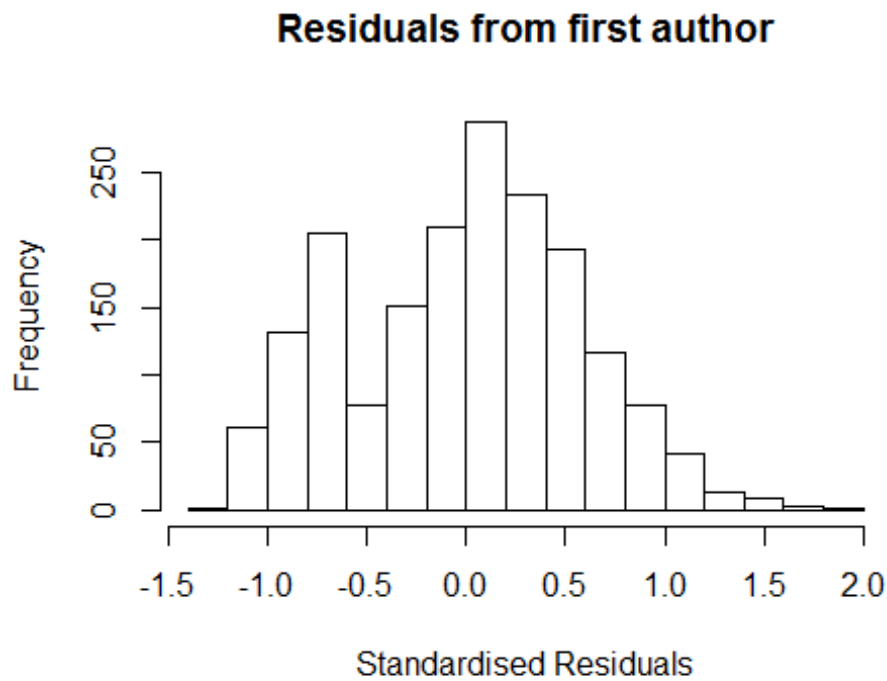

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
UniqueAuthors +
##     Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##     k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.4611 -0.3487 -0.0163  0.3560  2.0447
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         0.61666    0.05843   10.55  < 2e-16 ***
## FirstAuthorFemale1  0.00286    0.02642    0.11  0.91368
## LastAuthorFemale1  -0.03055    0.02753   -1.11  0.26729
## UniqueAuthors2      0.43603    0.04927    8.85  < 2e-16 ***
## UniqueAuthors3      0.59326    0.04415   13.44  < 2e-16 ***
## UniqueAuthors4      0.72437    0.04317   16.78  < 2e-16 ***
## UniqueAuthors5      0.77988    0.04131   18.88  < 2e-16 ***
## Year1997            0.08068    0.06625    1.22  0.22344
## Year1998           -0.14834    0.07509   -1.98  0.04838 *
## Year1999           -0.13287    0.07449   -1.78  0.07465 .
```

```
## Year2000           -0.04055    0.06820   -0.59  0.55216
## Year2001           -0.09091    0.07089   -1.28  0.19984
## Year2002            0.06458    0.06519    0.99  0.32204
## Year2003            0.04972    0.06818    0.73  0.46595
## Year2004           -0.00569    0.06817   -0.08  0.93350
## Year2005           -0.03071    0.06904   -0.44  0.65652
## Year2006           -0.05173    0.07245   -0.71  0.47529
## Year2007           -0.18896    0.07780   -2.43  0.01525 *
## Year2008           -0.17179    0.08437   -2.04  0.04187 *
## Year2009           -0.17268    0.08099   -2.13  0.03314 *
## Year2010           -0.26793    0.07332   -3.65  0.00027 ***
## Year2011           -0.14076    0.06860   -2.05  0.04033 *
## Year2012           -0.32299    0.06962   -4.64  3.7e-06 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.432
## Multiple R-squared:  0.335,  Adjusted R-squared:  0.326
## Convergence in 16 IRWLS iterations
##
## Robustness weights:
##  2 observations c(1285,1709) are outliers with |weight| = 0 ( < 5.5e-05);
##  127 weights are ~= 1. The remaining 1681 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0339  0.8390  0.9310  0.8740  0.9800  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          5.52e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.104  1           1.051
## LastAuthorFemale  1.062  1           1.031
## Year              1.122 16           1.004
```



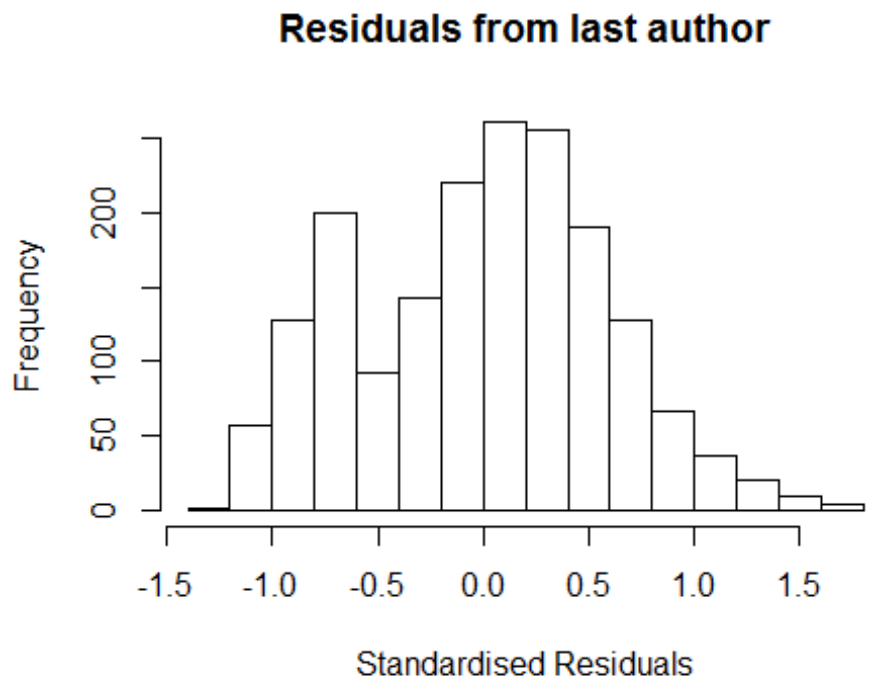

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.2129 -0.4550  0.0521  0.4008  1.7762
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)          0.95541    0.05372   17.79  < 2e-16 ***
## FirstAuthorFemale1   0.06039    0.03036    1.99   0.0468 *
## LastAuthorFemale1   -0.12372    0.03151   -3.93  8.9e-05 ***
## Year1997             0.16894    0.06793    2.49   0.0130 *
## Year1998            -0.00891    0.08281   -0.11   0.9143
## Year1999             0.04952    0.08772    0.56   0.5724
## Year2000             0.17487    0.07321    2.39   0.0170 *
## Year2001             0.07949    0.08817    0.90   0.3674
## Year2002             0.17741    0.06474    2.74   0.0062 **
## Year2003             0.25747    0.06193    4.16  3.4e-05 ***
## Year2004             0.14432    0.06558    2.20   0.0279 *
## Year2005             0.10527    0.06512    1.62   0.1062
```

```
## Year2006            0.12806    0.07123    1.80   0.0724 .
## Year2007           -0.06247    0.09347   -0.67   0.5040
## Year2008           -0.17296    0.09707   -1.78   0.0749 .
## Year2009           -0.10863    0.09498   -1.14   0.2528
## Year2010           -0.18431    0.09632   -1.91   0.0558 .
## Year2011            0.02642    0.09168    0.29   0.7733
## Year2012           -0.22469    0.09256   -2.43   0.0153 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.54
## Multiple R-squared:  0.0637, Adjusted R-squared:  0.0543
## Convergence in 15 IRWLS iterations
##
## Robustness weights:
##  147 weights are ~= 1. The remaining 1663 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.256   0.827   0.934   0.892   0.984   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          5.52e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.065  1           1.032
## Year              1.065 16           1.002
```




```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.2086 -0.4247  0.0573  0.4016  1.8236
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)          0.93108    0.05308   17.54  < 2e-16 ***
## FirstAuthorFemale1   0.02217    0.03052    0.73   0.4678
## Year1997             0.16643    0.06788    2.45   0.0143 *
## Year1998            -0.00677    0.08326   -0.08   0.9352
## Year1999             0.04939    0.08606    0.57   0.5661
## Year2000             0.15580    0.07230    2.15   0.0313 *
## Year2001             0.08138    0.08998    0.90   0.3659
## Year2002             0.17925    0.06465    2.77   0.0056 **
## Year2003             0.25530    0.06192    4.12  3.9e-05 ***
## Year2004             0.14453    0.06577    2.20   0.0281 *
## Year2005             0.10317    0.06532    1.58   0.1144
## Year2006             0.11794    0.07151    1.65   0.0993 .
```

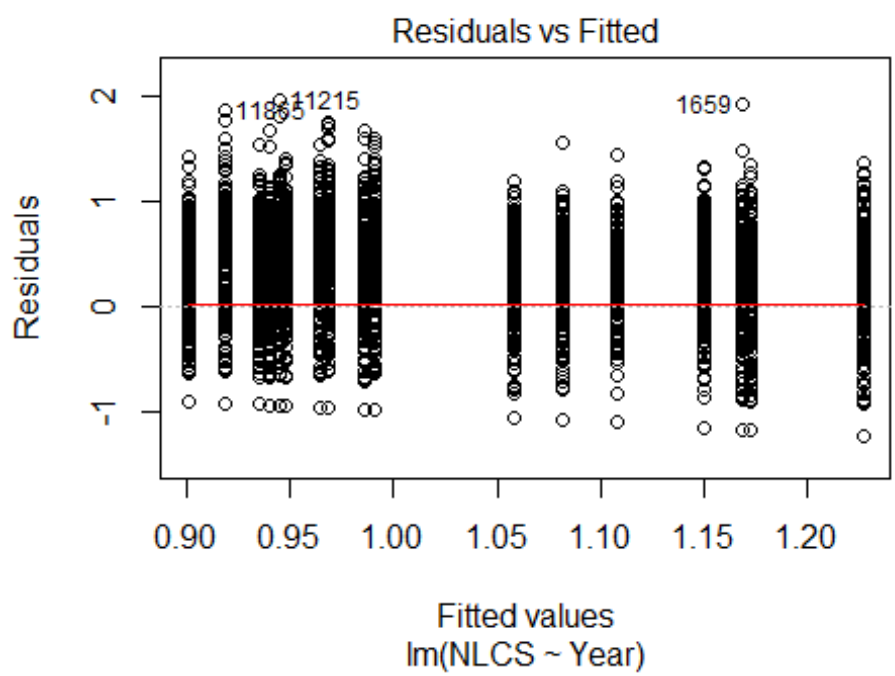
```
## Year2007           -0.05312    0.09313   -0.57   0.5685
## Year2008           -0.18056    0.09874   -1.83   0.0676 .
## Year2009           -0.13165    0.09838   -1.34   0.1810
## Year2010           -0.18795    0.09789   -1.92   0.0550 .
## Year2011            0.02534    0.09337    0.27   0.7861
## Year2012           -0.23252    0.09473   -2.45   0.0142 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.534
## Multiple R-squared:  0.0559, Adjusted R-squared:  0.0469
## Convergence in 16 IRWLS iterations
##
## Robustness weights:
##  133 weights are ~= 1. The remaining 1677 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.219   0.821   0.936   0.890   0.984   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          5.52e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##                   GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.037  1           1.018
## Year             1.037 16           1.001
```



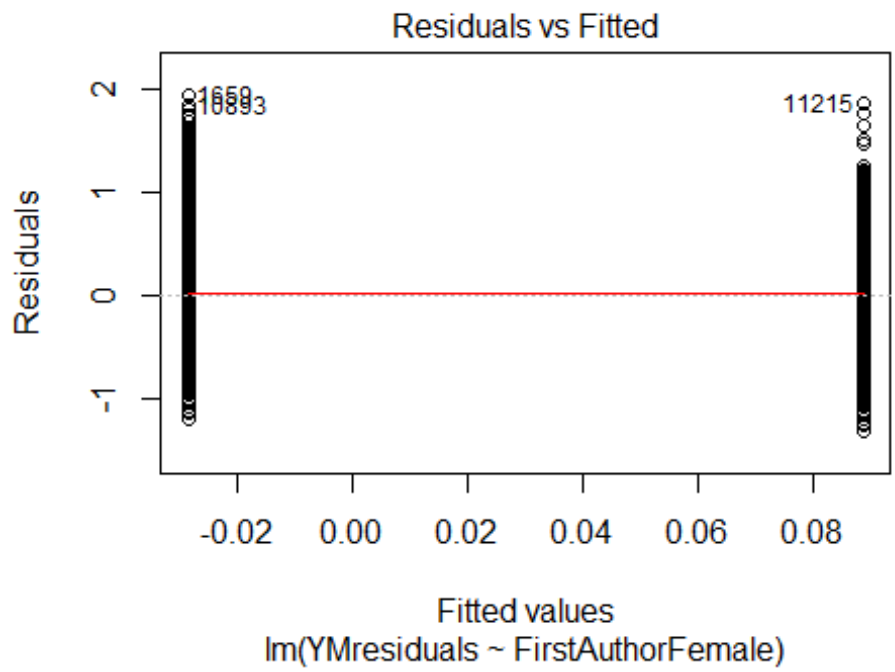

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.2377 -0.4454  0.0583  0.3986  1.7683
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         0.9750     0.0521   18.72  < 2e-16 ***
## LastAuthorFemale1  -0.1042     0.0315   -3.31  0.00096 ***
## Year1997            0.1654     0.0675    2.45  0.01431 *
## Year1998           -0.0020     0.0830   -0.02  0.98074
## Year1999            0.0537     0.0874    0.61  0.53879
## Year2000            0.1801     0.0733    2.46  0.01413 *
## Year2001            0.0834     0.0880    0.95  0.34342
## Year2002            0.1791     0.0646    2.77  0.00564 **
## Year2003            0.2627     0.0621    4.23  2.5e-05 ***
## Year2004            0.1482     0.0655    2.26  0.02367 *
## Year2005            0.1043     0.0645    1.62  0.10622
## Year2006            0.1359     0.0715    1.90  0.05759 .
```

```
## Year2007           -0.0563     0.0931   -0.60  0.54533
## Year2008           -0.1657     0.0967   -1.71  0.08674 .
## Year2009           -0.1050     0.0966   -1.09  0.27725
## Year2010           -0.1743     0.0966   -1.80  0.07143 .
## Year2011            0.0388     0.0916    0.42  0.67197
## Year2012           -0.2197     0.0931   -2.36  0.01839 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.539
## Multiple R-squared:  0.0618, Adjusted R-squared:  0.0529
## Convergence in 15 IRWLS iterations
##
## Robustness weights:
##  132 weights are ~= 1. The remaining 1678 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.259   0.829   0.934   0.892   0.983   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          5.52e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Sample size for the above analysis:  1810"
## [1] ""
## [1] ""
## [1] "####################################"
## [1] "Analysis of AJSC 3500"
## [1] "####################################"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  815  818  768  752  817  768  796  595  634  563  726  798  835  767  721
## 2011 2012
##  737  631
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  370  425  391  366  363  389  567  399  454  385  526  584  601  527  510
## 2011 2012
```

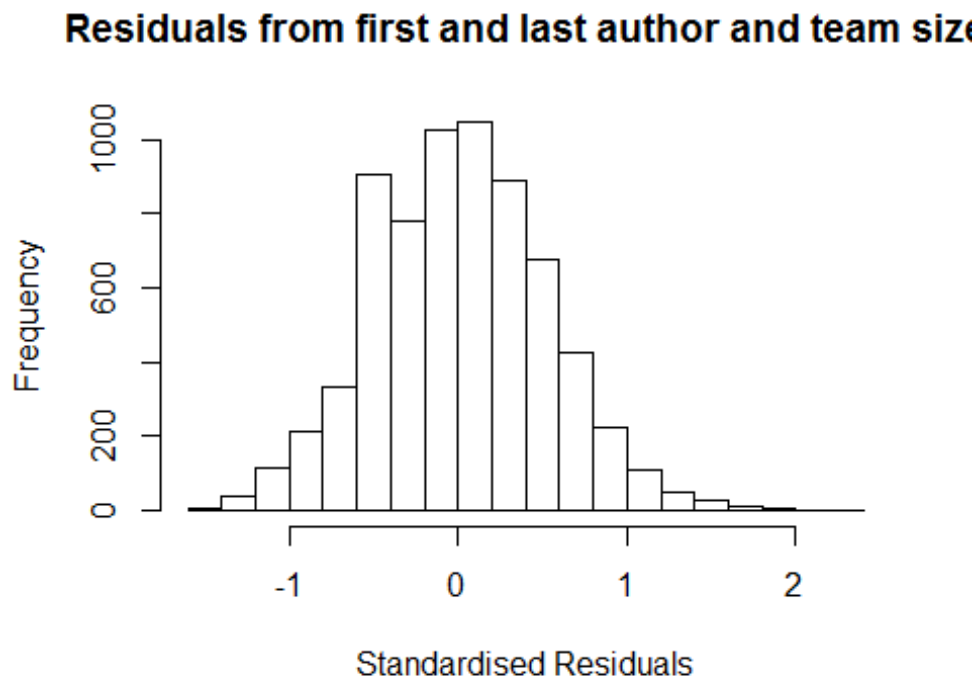
```
##  496  427
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  328  382  349  316  318  338  505  349  394  338  469  514  527  474  451
## 2011 2012
##  437  379
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
##  Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 100, df = 16, p-value = 3e-14
```




```
##
##  Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 21, df = 1, p-value = 6e-06
```

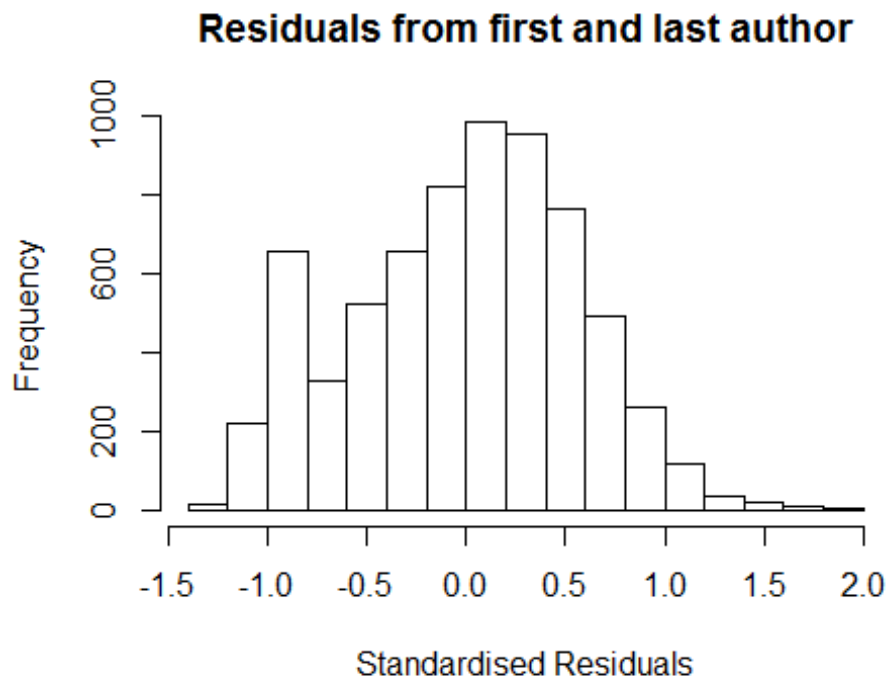



```
## [1] "Female first author team size 2018 geometric mean: 3.77925038162588"
## [1] "Male first author team size 2018 geometric mean: 2.59035620047027"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 16000, p-value = 8e-06
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 3.53437769819426"
## [1] "Male last author team size 2018 geometric mean: 2.73271796525043"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 14000, p-value = 0.006
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.120  1           1.058
## LastAuthorFemale  1.084  1           1.041
## UniqueAuthors     1.097  4           1.012
## Year              1.109 16           1.003
```



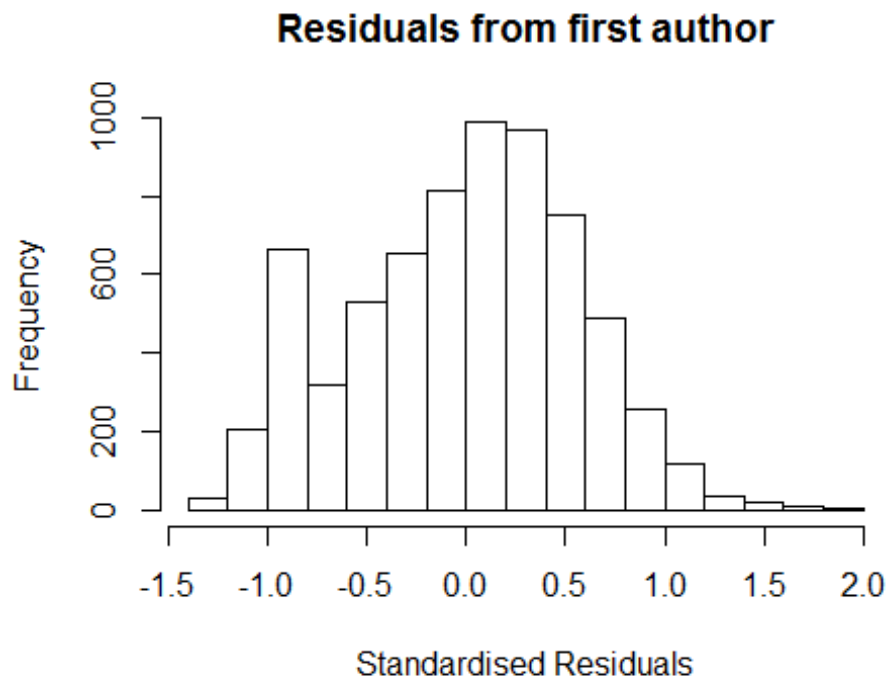

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
UniqueAuthors +
##     Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##     k.max = 1000))
##  \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.40195 -0.35851  0.00425  0.35414  2.24849
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)          0.8837     0.0315   28.04  < 2e-16 ***
## FirstAuthorFemale1   0.0411     0.0151    2.72   0.0066 **
## LastAuthorFemale1    0.0053     0.0166    0.32   0.7496
## UniqueAuthors2       0.3706     0.0217   17.06  < 2e-16 ***
## UniqueAuthors3       0.5164     0.0204   25.34  < 2e-16 ***
## UniqueAuthors4       0.5650     0.0203   27.83  < 2e-16 ***
## UniqueAuthors5       0.6451     0.0192   33.69  < 2e-16 ***
## Year1997            -0.0541     0.0426   -1.27   0.2036
## Year1998            -0.0467     0.0417   -1.12   0.2625
## Year1999            -0.1212     0.0423   -2.87   0.0042 **
```

```
## Year2000            -0.1641     0.0415   -3.95  7.9e-05 ***
## Year2001            -0.1288     0.0417   -3.09   0.0020 **
## Year2002            -0.3254     0.0375   -8.67  < 2e-16 ***
## Year2003            -0.3942     0.0392  -10.06  < 2e-16 ***
## Year2004            -0.4062     0.0378  -10.76  < 2e-16 ***
## Year2005            -0.2638     0.0397   -6.65  3.1e-11 ***
## Year2006            -0.3547     0.0367   -9.68  < 2e-16 ***
## Year2007            -0.3287     0.0365   -9.01  < 2e-16 ***
## Year2008            -0.3610     0.0369   -9.79  < 2e-16 ***
## Year2009            -0.3718     0.0392   -9.48  < 2e-16 ***
## Year2010            -0.3279     0.0386   -8.48  < 2e-16 ***
## Year2011            -0.3334     0.0403   -8.28  < 2e-16 ***
## Year2012            -0.3305     0.0436   -7.57  4.1e-14 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.501
## Multiple R-squared:  0.23,   Adjusted R-squared:  0.227
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
##  576 weights are ~= 1. The remaining 6292 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0067  0.8750  0.9440  0.9020  0.9840  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.46e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.067  1           1.033
## LastAuthorFemale  1.055  1           1.027
## Year              1.036 16           1.001
```




```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##    Min     1Q Median     3Q    Max
## -1.341 -0.418  0.046  0.398  1.931
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)           1.1992     0.0303   39.52  < 2e-16 ***
## FirstAuthorFemale1    0.1265     0.0163    7.78  8.3e-15 ***
## LastAuthorFemale1     0.0153     0.0182    0.84  0.40150
## Year1997             -0.0525     0.0421   -1.24  0.21328
## Year1998             -0.0450     0.0412   -1.09  0.27461
## Year1999             -0.1519     0.0434   -3.50  0.00047 ***
## Year2000             -0.1148     0.0421   -2.72  0.00645 **
## Year2001             -0.0978     0.0436   -2.24  0.02512 *
## Year2002             -0.2960     0.0409   -7.24  4.9e-13 ***
## Year2003             -0.3187     0.0429   -7.43  1.2e-13 ***
## Year2004             -0.3592     0.0412   -8.71  < 2e-16 ***
## Year2005             -0.1694     0.0414   -4.10  4.3e-05 ***
```

```
## Year2006            -0.3072     0.0394   -7.79  7.6e-15 ***
## Year2007            -0.2735     0.0392   -6.98  3.2e-12 ***
## Year2008            -0.3040     0.0402   -7.56  4.6e-14 ***
## Year2009            -0.3512     0.0431   -8.14  4.5e-16 ***
## Year2010            -0.3131     0.0420   -7.46  9.8e-14 ***
## Year2011            -0.3259     0.0459   -7.11  1.3e-12 ***
## Year2012            -0.2779     0.0497   -5.60  2.3e-08 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.579
## Multiple R-squared:  0.0443, Adjusted R-squared:  0.0418
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  577 weights are ~= 1. The remaining 6291 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.244   0.864   0.945   0.911   0.984   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.46e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.025  1           1.012
## Year              1.025 16           1.001
```

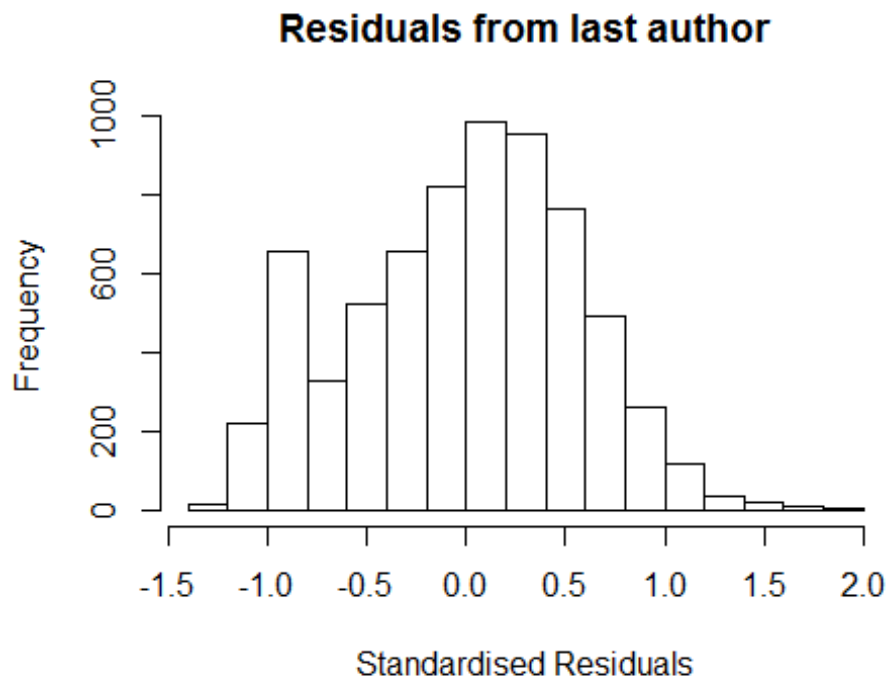



```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.3320 -0.4164  0.0471  0.3983  1.9291
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)           1.2011     0.0302   39.80  < 2e-16 ***
## FirstAuthorFemale1    0.1309     0.0160    8.17  3.8e-16 ***
## Year1997             -0.0522     0.0421   -1.24  0.21537
## Year1998             -0.0446     0.0411   -1.08  0.27824
## Year1999             -0.1527     0.0434   -3.52  0.00043 ***
## Year2000             -0.1146     0.0421   -2.72  0.00647 **
## Year2001             -0.0973     0.0436   -2.23  0.02568 *
## Year2002             -0.2969     0.0408   -7.27  4.0e-13 ***
## Year2003             -0.3191     0.0429   -7.45  1.1e-13 ***
## Year2004             -0.3597     0.0412   -8.74  < 2e-16 ***
## Year2005             -0.1699     0.0414   -4.11  4.1e-05 ***
## Year2006             -0.3073     0.0394   -7.80  7.2e-15 ***
```

```
## Year2007            -0.2733     0.0391   -6.98  3.1e-12 ***
## Year2008            -0.3039     0.0402   -7.56  4.5e-14 ***
## Year2009            -0.3510     0.0431   -8.14  4.7e-16 ***
## Year2010            -0.3128     0.0420   -7.45  1.0e-13 ***
## Year2011            -0.3258     0.0459   -7.10  1.3e-12 ***
## Year2012            -0.2771     0.0497   -5.58  2.5e-08 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.579
## Multiple R-squared:  0.0441, Adjusted R-squared:  0.0418
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  579 weights are ~= 1. The remaining 6289 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.245   0.863   0.946   0.911   0.984   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.46e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##                   GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.013  1           1.007
## Year             1.013 16           1.000

## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.2759 -0.4084  0.0436  0.3983  1.9153
```

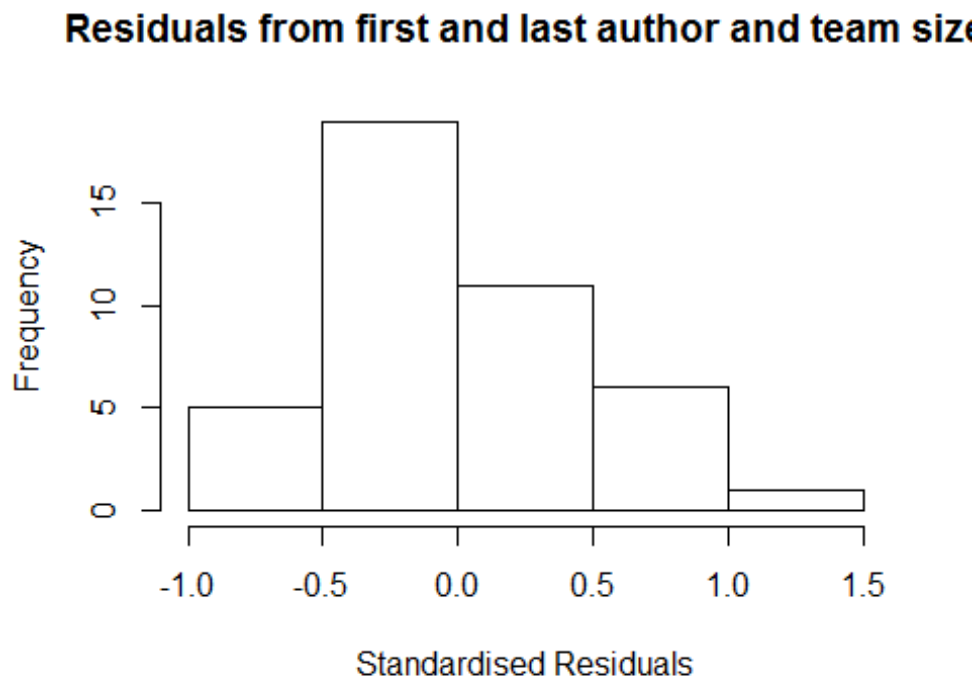
```
##
## Coefficients:
##                   Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.2175     0.0303   40.18  < 2e-16 ***
## LastAuthorFemale1   0.0584     0.0176    3.32  0.00092 ***
## Year1997           -0.0548     0.0421   -1.30  0.19244
## Year1998           -0.0449     0.0412   -1.09  0.27675
## Year1999           -0.1481     0.0433   -3.42  0.00064 ***
## Year2000           -0.1098     0.0422   -2.60  0.00935 **
## Year2001           -0.0950     0.0435   -2.18  0.02896 *
## Year2002           -0.2930     0.0409   -7.16  8.9e-13 ***
## Year2003           -0.3196     0.0431   -7.41  1.4e-13 ***
## Year2004           -0.3573     0.0413   -8.65  < 2e-16 ***
## Year2005           -0.1647     0.0416   -3.96  7.6e-05 ***
## Year2006           -0.2999     0.0397   -7.56  4.6e-14 ***
## Year2007           -0.2732     0.0394   -6.94  4.3e-12 ***
## Year2008           -0.2951     0.0404   -7.31  2.9e-13 ***
## Year2009           -0.3439     0.0434   -7.92  2.8e-15 ***
## Year2010           -0.3044     0.0422   -7.21  6.3e-13 ***
## Year2011           -0.3083     0.0463   -6.65  3.1e-11 ***
## Year2012           -0.2695     0.0500   -5.39  7.1e-08 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.581
## Multiple R-squared:  0.0366, Adjusted R-squared:  0.0342
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  562 weights are ~= 1. The remaining 6306 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.254   0.867   0.947   0.911   0.984   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.46e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Sample size for the above analysis:  6868"
## [1] ""
```

```
## [1] ""
## [1] "####################################"
## [1] "Analysis of AJSC 3501"
## [1] "####################################"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2003 2004 2005 2006 2008 2009 2010 2011 2012
##    1    7    5    1    1    1    1    2    2    4    4    9    9   15   45
##
## 1996 1997 1998 1999 2000 2001 2003 2004 2005 2006 2008 2009 2010 2011 2012
##    0    0    0    0    0    1    0    1    0    1    0    6    5   10   27
##
## 1996 1997 1998 1999 2000 2001 2003 2004 2005 2006 2008 2009 2010 2011 2012
##    0    0    0    0    0    1    0    1    0    0    0    6    5    8   21
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
## [1] "Female first author team size 2018 geometric mean: 3.94449811622082"
## [1] "Male first author team size 2018 geometric mean: 4.49467012304143"

## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties

##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 110, p-value = 0.6
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 4.88941496808028"
## [1] "Male last author team size 2018 geometric mean: 3.59657026181035"

## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties
```

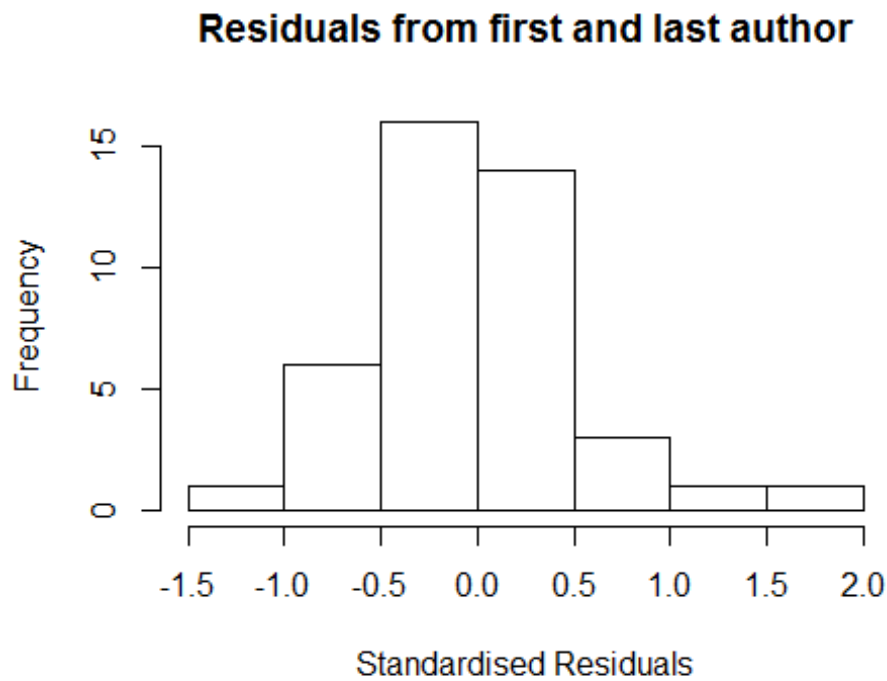



```
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 170, p-value = 0.09
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##                        GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 6.041e+13  1       7.772e+06
## LastAuthorFemale  9.331e+13  1       9.660e+06
## UniqueAuthors     5.872e+15  4       9.356e+01
## Year              2.915e+16  5       4.431e+01

## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
UniqueAuthors +
##     Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##     k.max = 1000))
##  \--> method = "MM"
## Residuals:
```

```
##      Min       1Q   Median       3Q      Max
## -0.96621 -0.24066 -0.00702  0.22831  1.41340
##
## Coefficients:
##                     Estimate Std. Error   t value Pr(>|t|)
## (Intercept)         1.86e+00   1.87e-08  9.97e+07  < 2e-16 ***
## FirstAuthorFemale1  3.28e-01   1.78e-01  1.84e+00   0.0758 .
## LastAuthorFemale1  -1.82e-01   2.21e-01 -8.20e-01   0.4163
## UniqueAuthors2      6.52e-01   4.70e-01  1.39e+00   0.1756
## UniqueAuthors3      4.45e-01   4.23e-01  1.05e+00   0.3015
## UniqueAuthors4      5.55e-01   4.90e-01  1.13e+00   0.2663
## UniqueAuthors5      8.04e-01   4.14e-01  1.94e+00   0.0616 .
## Year2004           -2.66e+00   5.62e-01 -4.74e+00  4.9e-05 ***
## Year2009           -2.10e+00   4.45e-01 -4.71e+00  5.3e-05 ***
## Year2010           -8.54e-01   4.27e-01 -2.00e+00   0.0547 .
## Year2011           -1.17e+00   5.05e-01 -2.32e+00   0.0273 *
## Year2012           -1.45e+00   4.41e-01 -3.30e+00   0.0025 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.45
## Multiple R-squared:  0.539,  Adjusted R-squared:  0.371
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
##  8 weights are ~= 1. The remaining 34 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.302   0.855   0.948   0.894   0.980   0.998
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          2.38e-03          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"

## Warning in lf.cov(init, x = x): .vcov.avar1: negative diag(<vcov>) fixed
## up; consider 'cov=".vcov.w."' instead
```



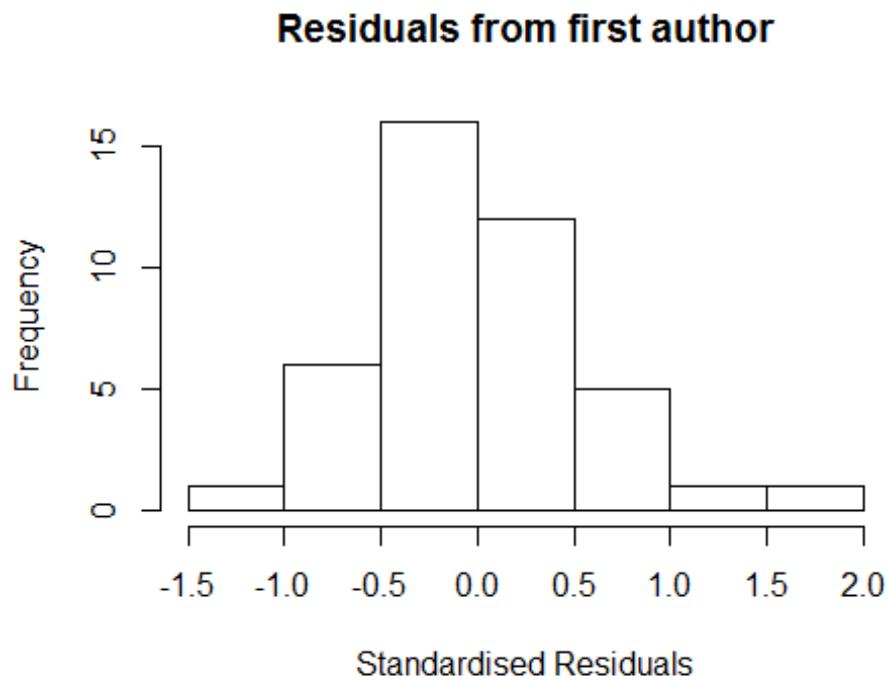

```
##                         GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 3.000e+14  1       1.732e+07
## LastAuthorFemale  5.994e+14  1       2.448e+07
## Year              1.711e+15  5       3.337e+01
```



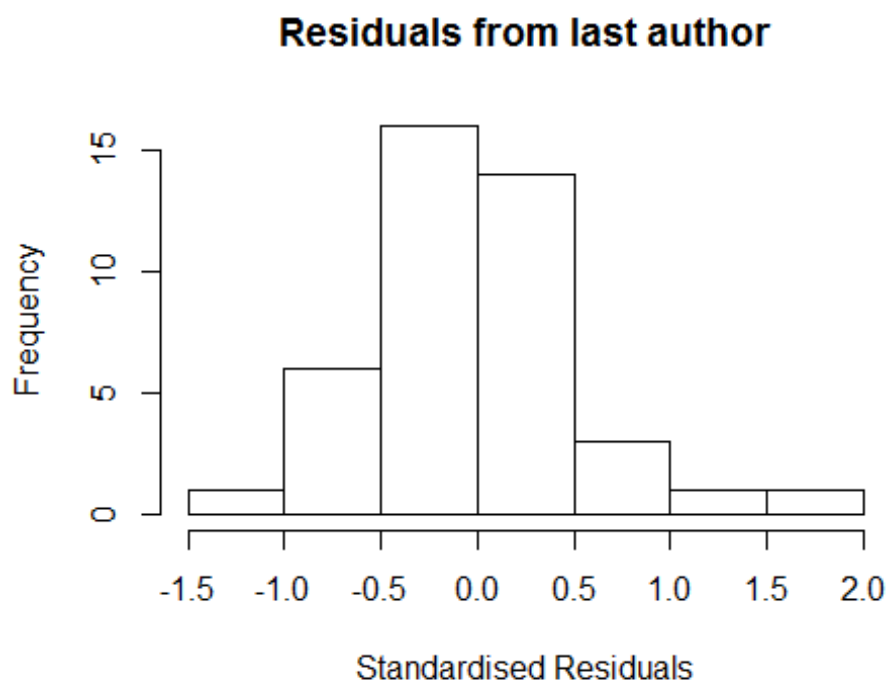

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.25661 -0.27030 -0.00231  0.34727  1.60439
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)           1.864      0.000     Inf  < 2e-16 ***
## FirstAuthorFemale1    0.326      0.195    1.67    0.104
## LastAuthorFemale1    -0.194      0.276   -0.70    0.486
## Year2004             -1.995      0.378   -5.28  7.5e-06 ***
## Year2009             -1.488      0.445   -3.34    0.002 **
## Year2010             -0.194      0.136   -1.43    0.162
## Year2011             -0.607      0.266   -2.29    0.029 *
## Year2012             -0.842      0.140   -6.01  8.4e-07 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.542
```

```
## Multiple R-squared:  0.425,  Adjusted R-squared:  0.306
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
##  6 weights are ~= 1. The remaining 36 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.362   0.911   0.943   0.914   0.989   0.998
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          2.38e-03          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##                        GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 9.898e+13  1       9.949e+06
## Year              9.898e+13  5       2.509e+01

## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.2124 -0.2381 -0.0212  0.3107  1.6467
##
## Coefficients:
##                     Estimate Std. Error   t value Pr(>|t|)
## (Intercept)         1.86e+00   2.17e-08  8.60e+07  < 2e-16 ***
## FirstAuthorFemale1  3.51e-01   1.84e-01  1.90e+00    0.065 .
## Year2004           -2.22e+00   1.84e-01 -1.20e+01  5.7e-14 ***
## Year2009           -1.67e+00   3.09e-01 -5.42e+00  4.5e-06 ***
## Year2010           -1.94e-01   1.36e-01 -1.43e+00    0.163
## Year2011           -6.52e-01   2.47e-01 -2.64e+00    0.012 *
## Year2012           -8.85e-01   1.16e-01 -7.63e+00  6.0e-09 ***
```

```
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.529
## Multiple R-squared:  0.428,  Adjusted R-squared:  0.33
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
##  6 weights are ~= 1. The remaining 36 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.312   0.901   0.953   0.908   0.989   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          2.38e-03          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"

## Warning in lf.cov(init, x = x): .vcov.avar1: negative diag(<vcov>) fixed
## up; consider 'cov=".vcov.w."' instead
```



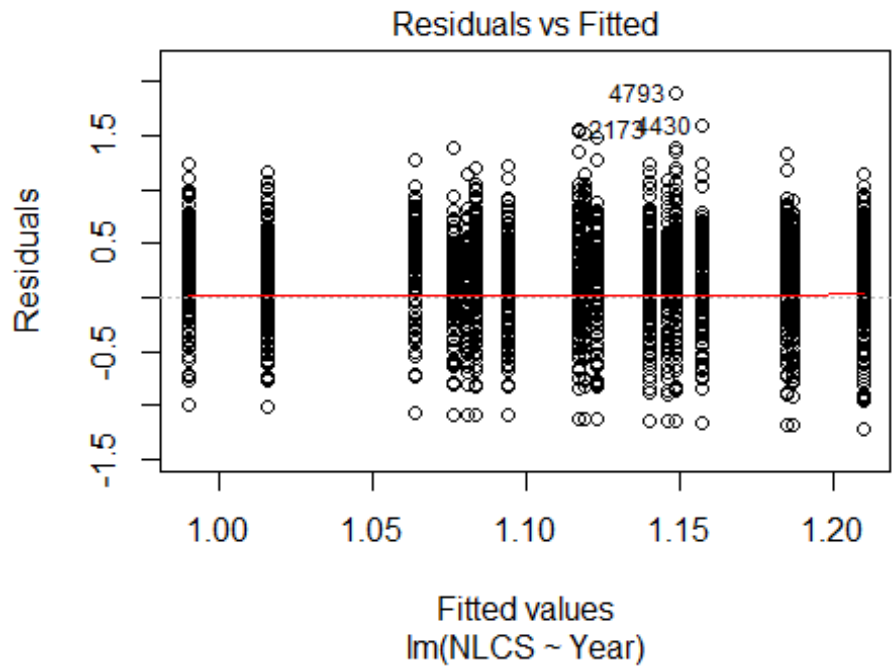

```
##                        GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 7.444e+13  1       8.628e+06
## Year             7.444e+13  5       2.439e+01
```



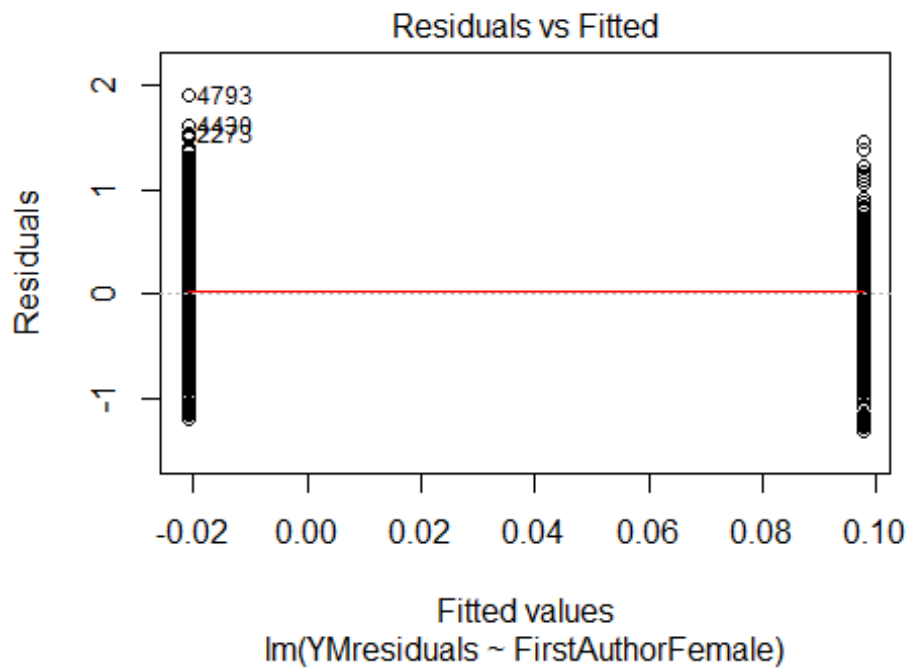
```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##       Min        1Q    Median        3Q       Max
## -1.46e+00 -3.92e-01  3.22e-15  3.64e-01  1.48e+00
##
## Coefficients:
##                   Estimate Std. Error t value Pr(>|t|)
## (Intercept)          1.864      0.000     Inf  < 2e-16 ***
## LastAuthorFemale1   -0.275      0.289   -0.95   0.3486
## Year2004            -1.589      0.289   -5.49  3.6e-06 ***
## Year2009            -1.165      0.415   -2.81   0.0081 **
## Year2010            -0.194      0.136   -1.43   0.1616
## Year2011            -0.400      0.172   -2.32   0.0261 *
## Year2012            -0.718      0.137   -5.25  7.5e-06 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.545
## Multiple R-squared:  0.388,  Adjusted R-squared:  0.283
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
##  8 weights are ~= 1. The remaining 34 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.441   0.902   0.946   0.903   0.978   0.998
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          2.38e-03          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                   "SM"
## seed : int(0)
## [1] "Sample size for the above analysis:  42"
```

```
## [1] ""
## [1] ""
## [1] "##################################"
## [1] "Analysis of AJSC 3503"
## [1] "##################################"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 2003 2005 2008 2010 2012
##    1    2    1    1    1
##
## 2003 2005 2008 2010 2012
##    1    0    1    0    0
##
## 2003 2005 2008 2010 2012
##    1    0    1    0    0
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
## [1] "Female first author team size 2018 geometric mean: 4"
## [1] "Male first author team size 2018 geometric mean: NaN"
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
## [1] "Regression 3: First author gender, Year as factors"
## [1] "Regression 4: Last author gender, Year as factors"
## [1] "Sample size for the above analysis:  2"
## [1] ""
## [1] ""
## [1] "##################################"
## [1] "Analysis of AJSC 3504"
## [1] "##################################"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  389  382  343  347  368  415  368  292  180  233  148  213  232  285  268
## 2011 2012
##  279  315
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  260  279  199  173  237  200  287  209  142  172  126  167  181  202  198
## 2011 2012
##  210  225
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  224  233  178  151  215  177  253  181  119  149  110  143  159  166  172
## 2011 2012
##  172  184
## [1] "Heteroscedasticity checks, confirming that there are problems with
```

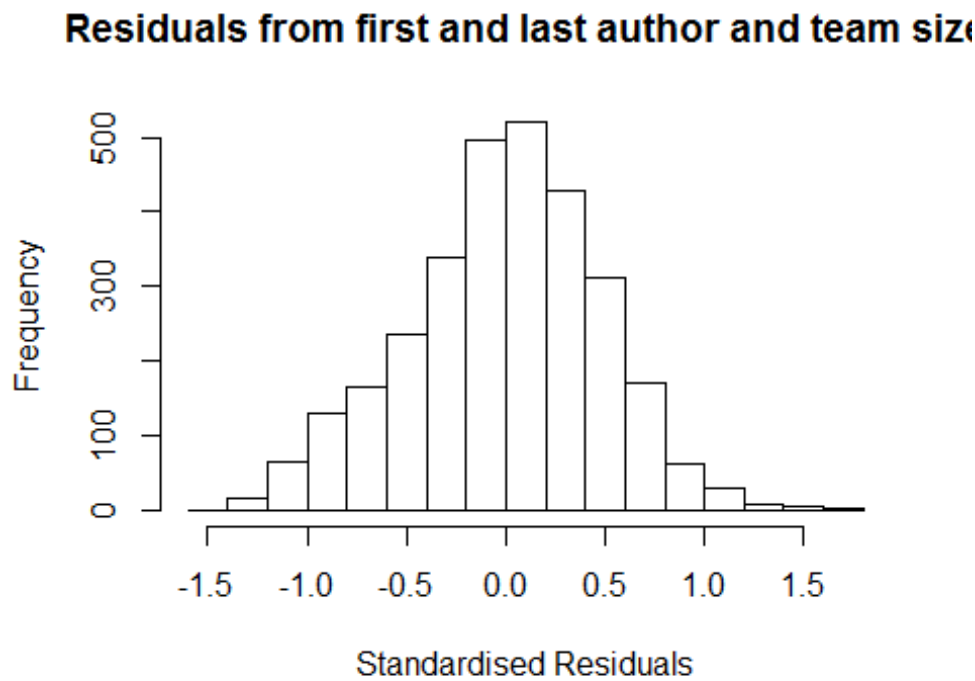
```
these"
##
##  Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 30, df = 16, p-value = 0.02
```




```
##
##  Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 4.4, df = 1, p-value = 0.04
```

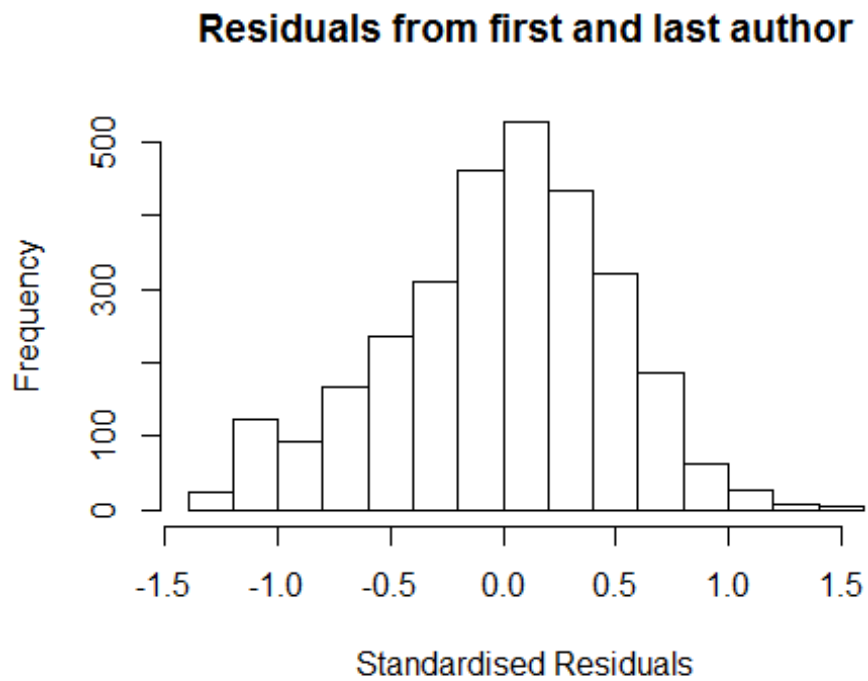



```
## [1] "Female first author team size 2018 geometric mean: 4.58919336434253"
## [1] "Male first author team size 2018 geometric mean: 4.03703146088941"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 9100, p-value = 0.3
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 4.50251119587957"
## [1] "Male last author team size 2018 geometric mean: 4.085870607764"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 8200, p-value = 0.3
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.125  1           1.061
## LastAuthorFemale  1.088  1           1.043
## UniqueAuthors     1.223  4           1.025
## Year              1.231 16           1.007
```



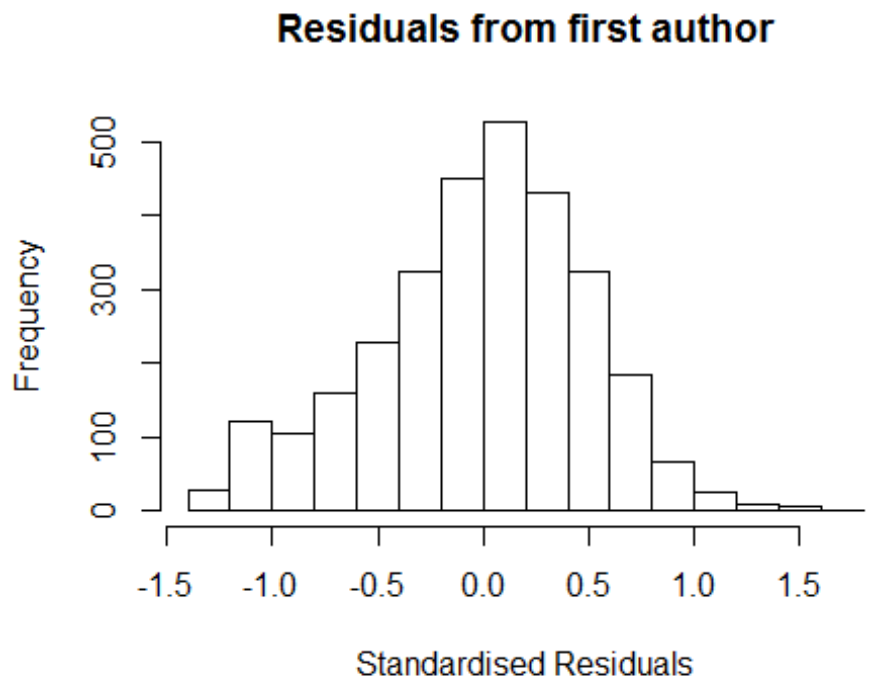

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
UniqueAuthors +
##     Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##     k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.4127 -0.3159  0.0154  0.3259  1.7394
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         0.94720    0.04272   22.17  < 2e-16 ***
## FirstAuthorFemale1  0.06191    0.02341    2.64  0.00824 **
## LastAuthorFemale1   0.05207    0.02715    1.92  0.05522 .
## UniqueAuthors2      0.09615    0.03688    2.61  0.00918 **
## UniqueAuthors3      0.22761    0.03461    6.58  5.7e-11 ***
## UniqueAuthors4      0.28813    0.03415    8.44  < 2e-16 ***
## UniqueAuthors5      0.41908    0.03360   12.47  < 2e-16 ***
## Year1997            0.05619    0.04643    1.21  0.22621
## Year1998            0.04835    0.04593    1.05  0.29259
## Year1999            0.02604    0.05166    0.50  0.61423
```

```
## Year2000           -0.05747    0.05282   -1.09  0.27670
## Year2001           -0.04061    0.05419   -0.75  0.45369
## Year2002           -0.17927    0.04645   -3.86  0.00012 ***
## Year2003           -0.16249    0.04774   -3.40  0.00067 ***
## Year2004           -0.10988    0.05203   -2.11  0.03478 *
## Year2005           -0.11673    0.05409   -2.16  0.03099 *
## Year2006           -0.14574    0.05283   -2.76  0.00584 **
## Year2007           -0.02450    0.05274   -0.46  0.64224
## Year2008           -0.07512    0.04981   -1.51  0.13166
## Year2009           -0.00563    0.05505   -0.10  0.91860
## Year2010           -0.08426    0.04749   -1.77  0.07616 .
## Year2011           -0.15128    0.04989   -3.03  0.00245 **
## Year2012           -0.11152    0.05292   -2.11  0.03519 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.47
## Multiple R-squared:  0.103,  Adjusted R-squared:  0.0963
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
##  276 weights are ~= 1. The remaining 2710 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.141   0.860   0.948   0.900   0.986   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          3.35e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.104  1           1.051
## LastAuthorFemale  1.091  1           1.045
## Year              1.063 16           1.002
```



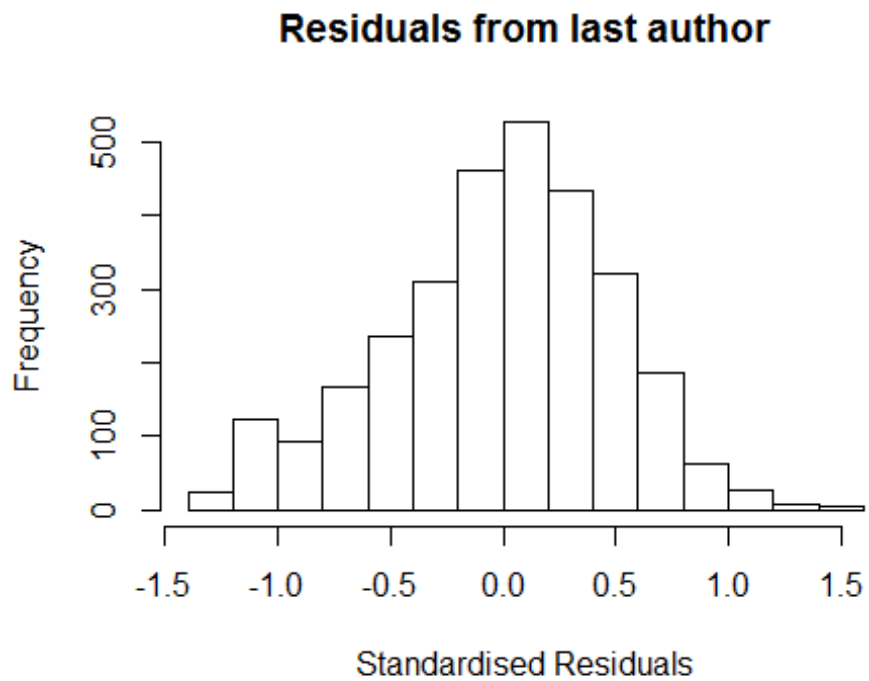

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.3922 -0.3352  0.0331  0.3317  1.5688
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)          1.13762    0.03533   32.20  < 2e-16 ***
## FirstAuthorFemale1   0.11177    0.02369    4.72  2.5e-06 ***
## LastAuthorFemale1    0.07712    0.02781    2.77  0.00559 **
## Year1997             0.06565    0.04779    1.37  0.16963
## Year1998             0.05194    0.04797    1.08  0.27901
## Year1999            -0.00670    0.05438   -0.12  0.90197
## Year2000            -0.05827    0.05417   -1.08  0.28213
## Year2001            -0.06041    0.05574   -1.08  0.27853
## Year2002            -0.17990    0.05132   -3.51  0.00046 ***
## Year2003            -0.17248    0.05004   -3.45  0.00057 ***
## Year2004            -0.08830    0.05483   -1.61  0.10739
## Year2005            -0.08220    0.05711   -1.44  0.15014
```

```
## Year2006           -0.10138    0.05755   -1.76  0.07824 .
## Year2007            0.02564    0.05308    0.48  0.62909
## Year2008            0.00201    0.05052    0.04  0.96826
## Year2009            0.04220    0.05527    0.76  0.44524
## Year2010           -0.04358    0.05023   -0.87  0.38567
## Year2011           -0.09513    0.05259   -1.81  0.07055 .
## Year2012           -0.05719    0.05401   -1.06  0.28974
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.481
## Multiple R-squared:  0.0316, Adjusted R-squared:  0.0257
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  287 weights are ~= 1. The remaining 2699 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.266   0.860   0.946   0.897   0.985   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          3.35e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.036  1           1.018
## Year              1.036 16           1.001
```




```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.3378 -0.3364  0.0291  0.3280  1.6043
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.14445    0.03504   32.66  < 2e-16 ***
## FirstAuthorFemale1  0.12647    0.02304    5.49  4.4e-08 ***
## Year1997            0.06688    0.04774    1.40  0.16139
## Year1998            0.05025    0.04792    1.05  0.29448
## Year1999           -0.00816    0.05422   -0.15  0.88044
## Year2000           -0.05851    0.05406   -1.08  0.27914
## Year2001           -0.05962    0.05576   -1.07  0.28505
## Year2002           -0.18316    0.05137   -3.57  0.00037 ***
## Year2003           -0.16911    0.04983   -3.39  0.00070 ***
## Year2004           -0.08725    0.05482   -1.59  0.11157
## Year2005           -0.07684    0.05703   -1.35  0.17800
## Year2006           -0.10200    0.05746   -1.78  0.07598 .
```

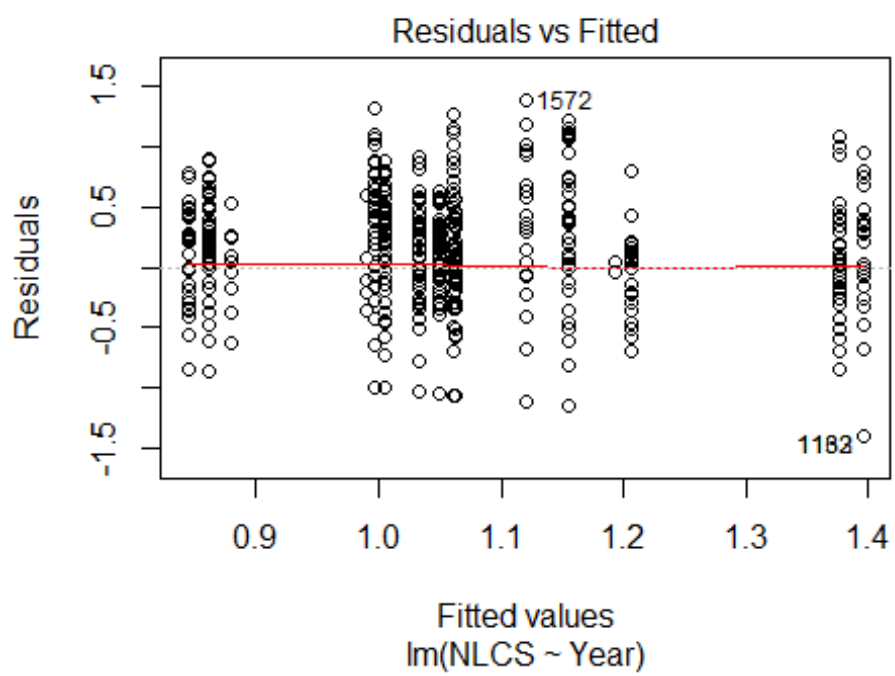
```
## Year2007            0.02655    0.05288    0.50  0.61558
## Year2008            0.00524    0.05066    0.10  0.91766
## Year2009            0.04236    0.05519    0.77  0.44287
## Year2010           -0.04234    0.05030   -0.84  0.39997
## Year2011           -0.08971    0.05224   -1.72  0.08604 .
## Year2012           -0.05407    0.05399   -1.00  0.31666
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.481
## Multiple R-squared:  0.0291, Adjusted R-squared:  0.0235
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  262 weights are ~= 1. The remaining 2724 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.244   0.860   0.949   0.898   0.986   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          3.35e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##                  GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.02  1           1.010
## Year             1.02 16           1.001
```




```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.3170 -0.3461  0.0353  0.3322  1.5804
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)        1.15e+00   3.52e-02   32.70  < 2e-16 ***
## LastAuthorFemale1  1.04e-01   2.72e-02    3.83  0.00013 ***
## Year1997           6.31e-02   4.78e-02    1.32  0.18727
## Year1998           5.11e-02   4.82e-02    1.06  0.28898
## Year1999           9.79e-05   5.41e-02    0.00  0.99856
## Year2000          -5.26e-02   5.47e-02   -0.96  0.33682
## Year2001          -5.31e-02   5.61e-02   -0.95  0.34387
## Year2002          -1.70e-01   5.13e-02   -3.32  0.00092 ***
## Year2003          -1.73e-01   5.06e-02   -3.43  0.00062 ***
## Year2004          -8.30e-02   5.52e-02   -1.50  0.13251
## Year2005          -7.44e-02   5.72e-02   -1.30  0.19365
## Year2006          -9.16e-02   5.75e-02   -1.59  0.11148
```

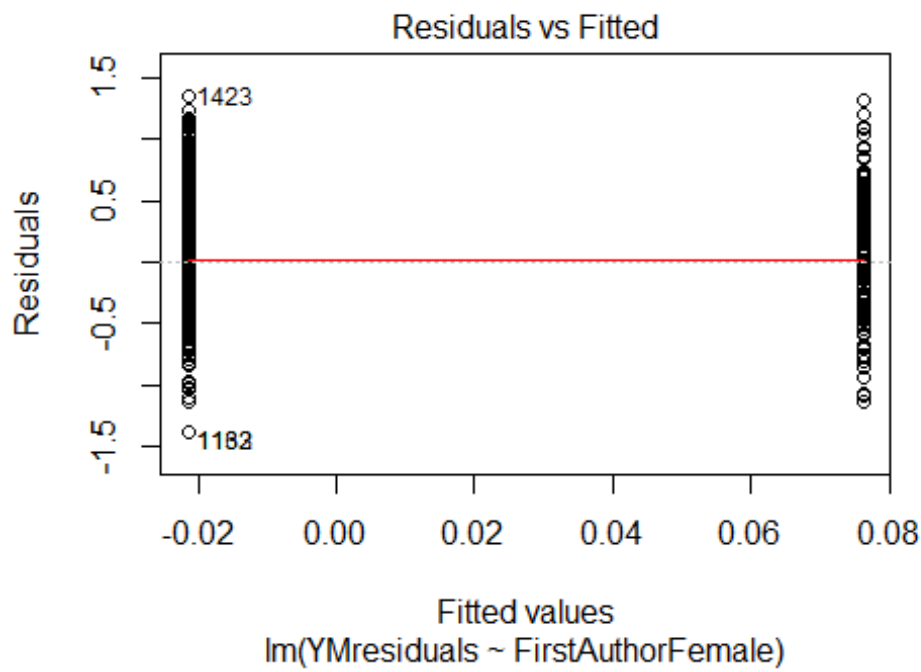
```
## Year2007           2.63e-02   5.32e-02    0.50  0.62037
## Year2008           9.40e-03   5.07e-02    0.19  0.85278
## Year2009           4.70e-02   5.48e-02    0.86  0.39116
## Year2010          -3.32e-02   5.04e-02   -0.66  0.50918
## Year2011          -9.10e-02   5.28e-02   -1.72  0.08512 .
## Year2012          -4.92e-02   5.43e-02   -0.91  0.36500
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.484
## Multiple R-squared:  0.0247, Adjusted R-squared:  0.0191
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  275 weights are ~= 1. The remaining 2711 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.264   0.860   0.947   0.898   0.985   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          3.35e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Sample size for the above analysis:  2986"
## [1] ""
## [1] ""
## [1] "####################################"
## [1] "Analysis of AJSC 3505"
## [1] "####################################"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##   81  104   94  113  117   82   90   84   49   56   33   49   40   81   75
## 2011 2012
##   99   65
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    2    5   14   10   72   54   70   53   35   39   25   41   29   61   54
## 2011 2012
```

```
##   78   50
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    2    5   14   10   65   44   61   40   32   28   21   37   22   57   45
## 2011 2012
##   69   44
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
##  Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 59, df = 16, p-value = 7e-07
```

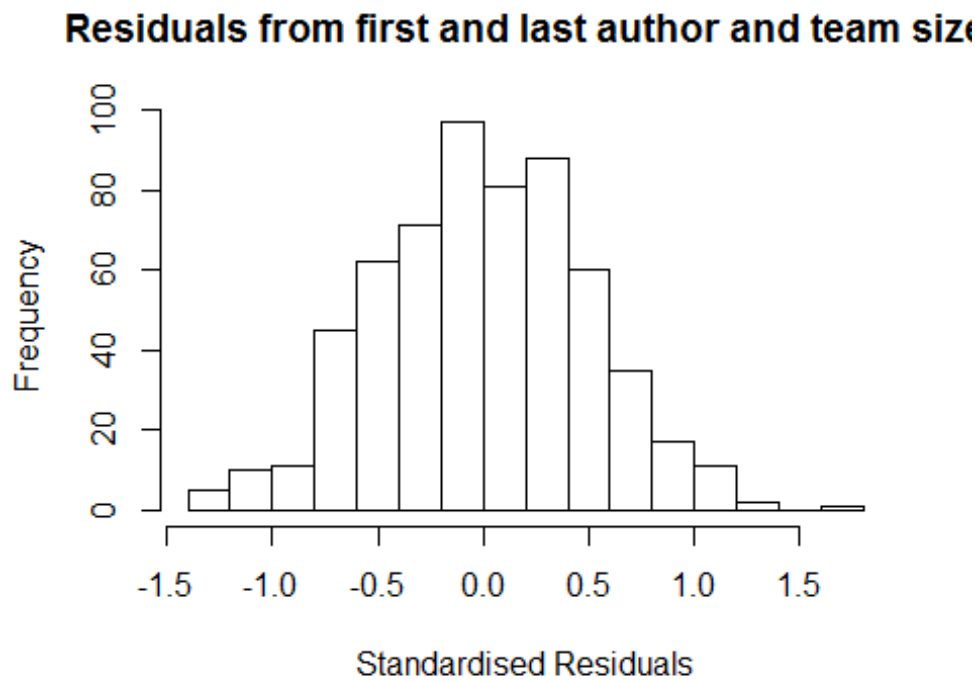



```
##
##  Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 0.026, df = 1, p-value = 0.9

## [1] "Female first author team size 2018 geometric mean: 5.29363964083957"
## [1] "Male first author team size 2018 geometric mean: 2.33321114339426"

## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties
```

```
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 170, p-value = 7e-04
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 3.30192724889463"
## [1] "Male last author team size 2018 geometric mean: 3.04303674900777"

## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties
```

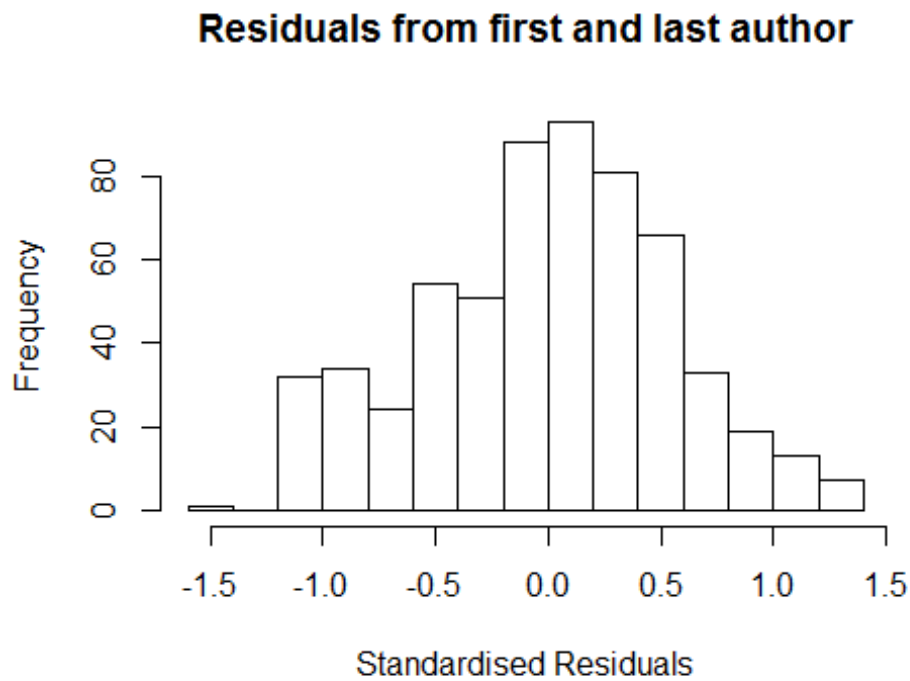



```
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 66, p-value = 0.9
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##                     GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.119  1           1.058
## LastAuthorFemale  1.208  1           1.099
## UniqueAuthors     1.974  4           1.089
## Year              2.241 16           1.026
```



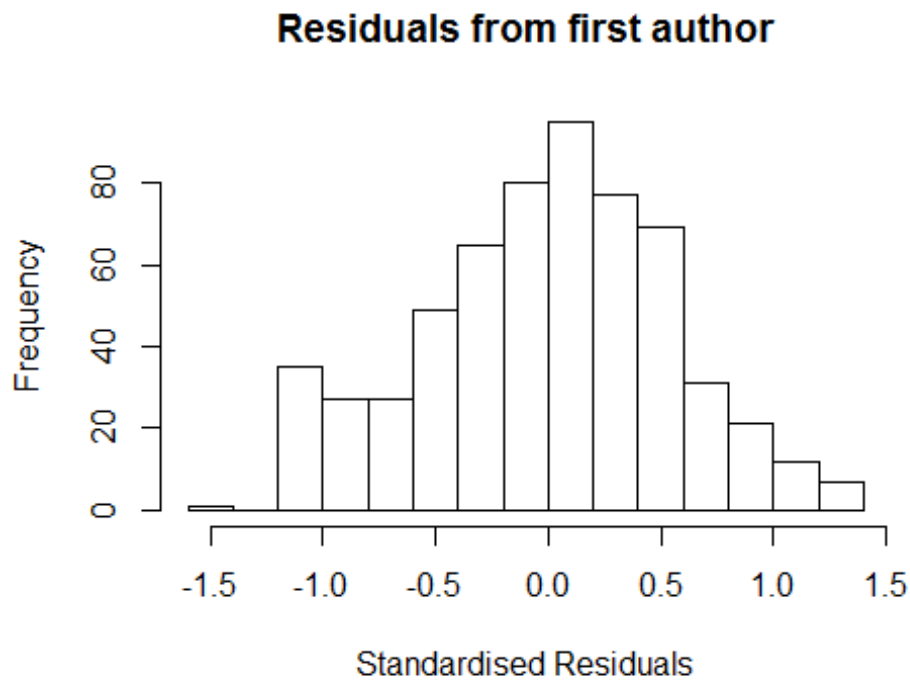

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
UniqueAuthors +
##     Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##     k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.3498 -0.3408 -0.0104  0.3393  1.6473
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         0.60580    0.09519    6.36  4.0e-10 ***
## FirstAuthorFemale1  0.04132    0.05168    0.80   0.4243
## LastAuthorFemale1  -0.09219    0.05372   -1.72   0.0867 .
## UniqueAuthors2      0.47933    0.06951    6.90  1.4e-11 ***
## UniqueAuthors3      0.49931    0.06690    7.46  3.2e-13 ***
## UniqueAuthors4      0.54588    0.06701    8.15  2.4e-15 ***
## UniqueAuthors5      0.71972    0.07727    9.31  < 2e-16 ***
## Year1997           -0.01176    0.21478   -0.05   0.9563
## Year1998           -0.01544    0.12882   -0.12   0.9046
## Year1999           -0.21450    0.10379   -2.07   0.0392 *
```

```
## Year2000            0.02431    0.10159    0.24   0.8110
## Year2001            0.05848    0.10546    0.55   0.5794
## Year2002           -0.10202    0.09763   -1.04   0.2965
## Year2003           -0.00869    0.09866   -0.09   0.9298
## Year2004           -0.10399    0.10941   -0.95   0.3423
## Year2005            0.00686    0.10217    0.07   0.9465
## Year2006            0.16898    0.11767    1.44   0.1515
## Year2007            0.27428    0.11061    2.48   0.0134 *
## Year2008            0.40784    0.12976    3.14   0.0018 **
## Year2009            0.12662    0.12362    1.02   0.3061
## Year2010           -0.06910    0.11786   -0.59   0.5579
## Year2011            0.02022    0.10344    0.20   0.8451
## Year2012            0.13951    0.12210    1.14   0.2537
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.508
## Multiple R-squared:  0.254,  Adjusted R-squared:  0.226
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
##  50 weights are ~= 1. The remaining 546 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.272   0.876   0.949   0.911   0.985   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.68e-04          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.321  1           1.149
## LastAuthorFemale  1.830  1           1.353
## Year              2.206 16           1.025
```



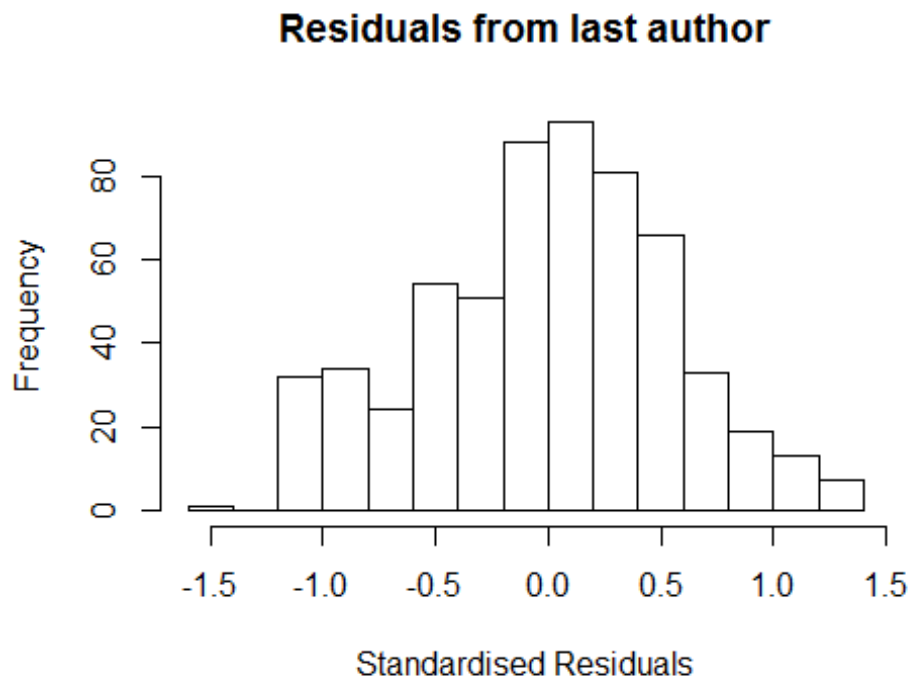

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.4518 -0.3690  0.0283  0.3817  1.3505
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)           1.2305     0.0302   40.73  < 2e-16 ***
## FirstAuthorFemale1    0.0958     0.0586    1.64  0.10234
## LastAuthorFemale1    -0.0759     0.0603   -1.26  0.20858
## Year1997             -0.2552     0.1502   -1.70  0.08992 .
## Year1998             -0.1729     0.0949   -1.82  0.06881 .
## Year1999             -0.3587     0.1016   -3.53  0.00045 ***
## Year2000             -0.1780     0.0789   -2.26  0.02444 *
## Year2001             -0.1973     0.0767   -2.57  0.01040 *
## Year2002             -0.3687     0.0844   -4.37  1.5e-05 ***
## Year2003             -0.1626     0.0646   -2.52  0.01211 *
## Year2004             -0.4199     0.0871   -4.82  1.8e-06 ***
## Year2005             -0.2185     0.0719   -3.04  0.00250 **
```

```
## Year2006            -0.0470     0.0798   -0.59  0.55623
## Year2007             0.1051     0.0869    1.21  0.22689
## Year2008             0.2213     0.1114    1.99  0.04735 *
## Year2009            -0.1189     0.1209   -0.98  0.32560
## Year2010            -0.2550     0.1022   -2.50  0.01282 *
## Year2011            -0.2590     0.0905   -2.86  0.00439 **
## Year2012            -0.0803     0.1217   -0.66  0.50961
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.533
## Multiple R-squared:  0.0722, Adjusted R-squared:  0.0433
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
##  46 weights are ~= 1. The remaining 550 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.438   0.868   0.949   0.901   0.986   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.68e-04          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.155  1           1.075
## Year              1.155 16           1.005
```




```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.4460 -0.3598  0.0238  0.3629  1.3644
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)           1.1925     0.0287   41.60  < 2e-16 ***
## FirstAuthorFemale1    0.0839     0.0603    1.39   0.1649
## Year1997             -0.2312     0.1555   -1.49   0.1377
## Year1998             -0.1392     0.0958   -1.45   0.1468
## Year1999             -0.3265     0.1045   -3.12   0.0019 **
## Year2000             -0.1494     0.0822   -1.82   0.0697 .
## Year2001             -0.1635     0.0780   -2.10   0.0365 *
## Year2002             -0.3440     0.0870   -3.95  8.7e-05 ***
## Year2003             -0.1302     0.0670   -1.94   0.0524 .
## Year2004             -0.3961     0.0929   -4.27  2.3e-05 ***
## Year2005             -0.1952     0.0761   -2.57   0.0105 *
## Year2006             -0.0148     0.0805   -0.18   0.8545
```

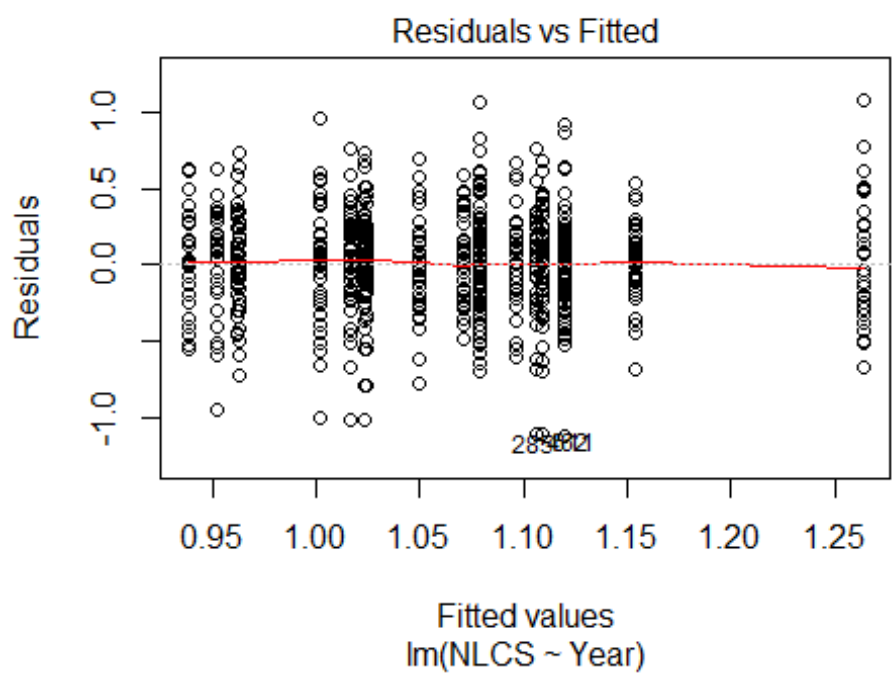
```
## Year2007             0.1359     0.0888    1.53   0.1264
## Year2008             0.2535     0.1134    2.24   0.0258 *
## Year2009            -0.0904     0.1228   -0.74   0.4620
## Year2010            -0.2250     0.1060   -2.12   0.0342 *
## Year2011            -0.2349     0.0938   -2.50   0.0126 *
## Year2012            -0.0563     0.1259   -0.45   0.6548
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.533
## Multiple R-squared:  0.0703, Adjusted R-squared:  0.0429
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
##  45 weights are ~= 1. The remaining 551 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.441   0.864   0.948   0.901   0.986   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.68e-04          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##                   GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.711  1           1.308
## Year             1.711 16           1.017
```




```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.4697 -0.3608  0.0155  0.3795  1.4076
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.2208     0.0314   38.86  < 2e-16 ***
## LastAuthorFemale1  -0.0567     0.0604   -0.94   0.3489
## Year1997           -0.2289     0.1461   -1.57   0.1177
## Year1998           -0.1514     0.0971   -1.56   0.1193
## Year1999           -0.3306     0.1026   -3.22   0.0013 **
## Year2000           -0.1516     0.0782   -1.94   0.0531 .
## Year2001           -0.1713     0.0773   -2.22   0.0271 *
## Year2002           -0.3464     0.0845   -4.10  4.7e-05 ***
## Year2003           -0.1336     0.0642   -2.08   0.0378 *
## Year2004           -0.3944     0.0845   -4.66  3.8e-06 ***
## Year2005           -0.1917     0.0707   -2.71   0.0069 **
## Year2006           -0.0236     0.0782   -0.30   0.7630
```

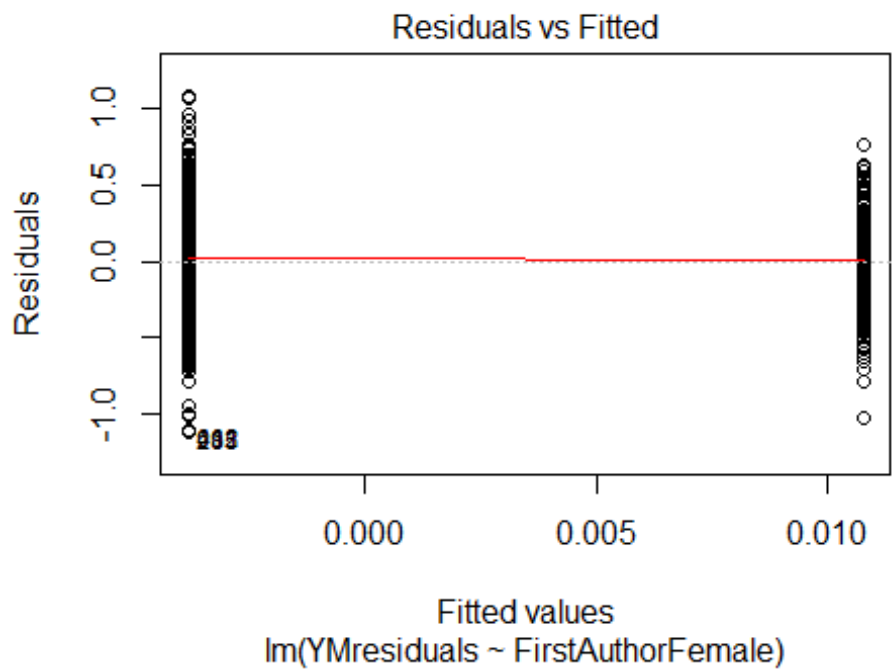
```
## Year2007            0.1410     0.0859    1.64   0.1015
## Year2008            0.2489     0.1094    2.28   0.0232 *
## Year2009           -0.0943     0.1206   -0.78   0.4345
## Year2010           -0.2181     0.1041   -2.10   0.0366 *
## Year2011           -0.2314     0.0895   -2.59   0.0099 **
## Year2012           -0.0648     0.1234   -0.53   0.5995
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.535
## Multiple R-squared:  0.0681, Adjusted R-squared:  0.0406
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
##  46 weights are ~= 1. The remaining 550 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.431   0.870   0.950   0.901   0.987   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.68e-04          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Sample size for the above analysis:  596"
## [1] ""
## [1] ""
## [1] "####################################"
## [1] "Analysis of AJSC 3506"
## [1] "####################################"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  111  109   88   55   68   57   62   47   55   57   41   63   51   78   53
## 2011 2012
##   60   51
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##   78   63   55   35   52   27   46   33   45   42   29   44   33   54   32
## 2011 2012
```

```
##   33   37
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##   74   57   50   27   37   22   38   29   35   31   24   33   28   42   28
## 2011 2012
##   27   26
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
##  Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 29, df = 16, p-value = 0.02
```

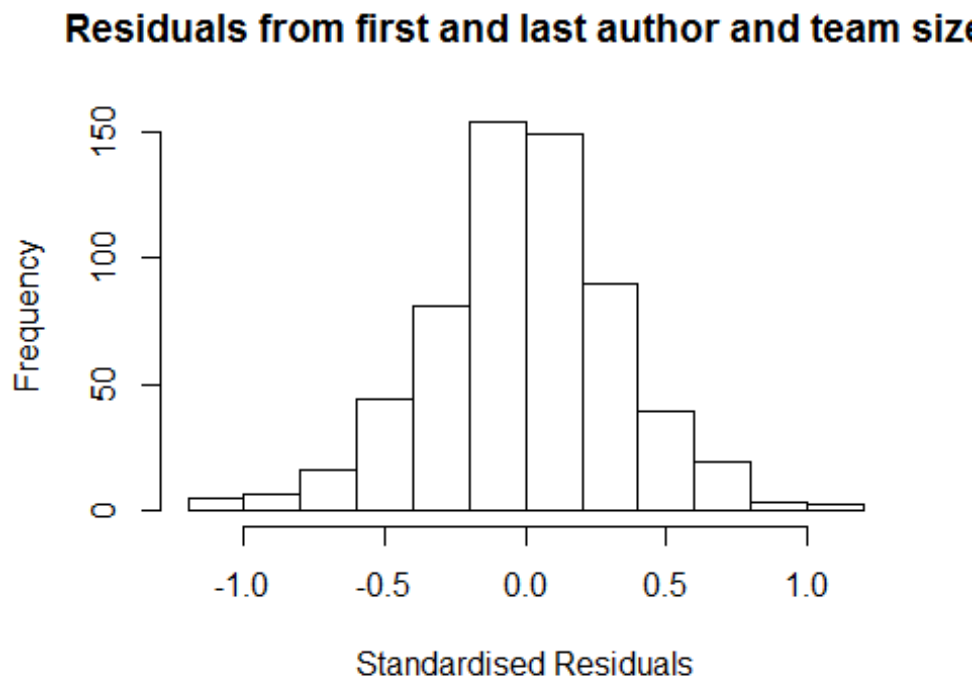



```
##
##  Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 7.5, df = 1, p-value = 0.006

## [1] "Female first author team size 2018 geometric mean: 4.88619991986741"
## [1] "Male first author team size 2018 geometric mean: 4.68700783368141"

## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties
```

```
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 100, p-value = 0.9
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 6.47983039827449"
## [1] "Male last author team size 2018 geometric mean: 4.47580587016569"

## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties
```

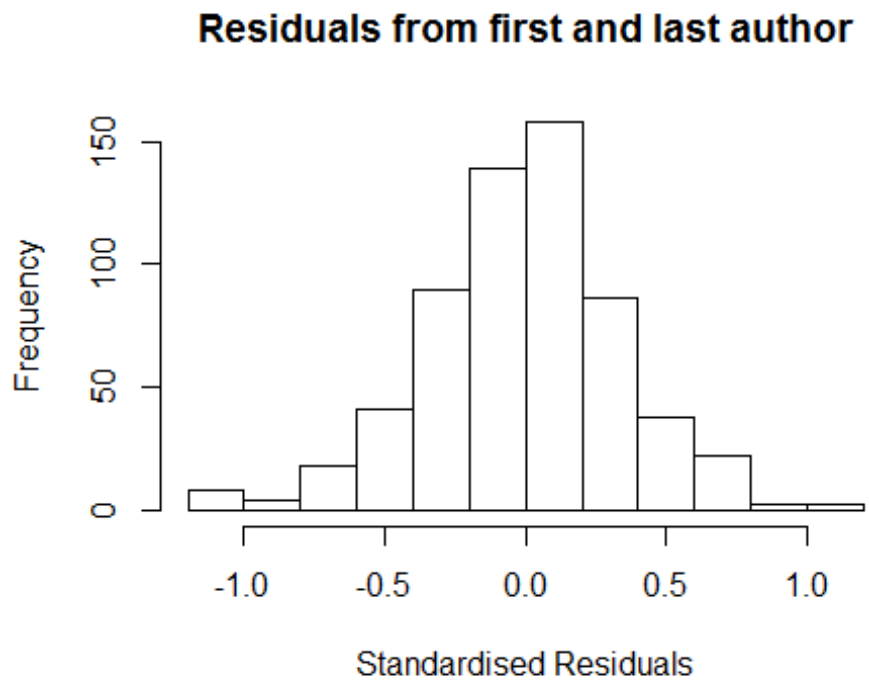



```
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 82, p-value = 0.4
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##                     GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.306  1           1.143
## LastAuthorFemale  1.262  1           1.123
## UniqueAuthors     2.239  4           1.106
## Year              3.283 16           1.038
```



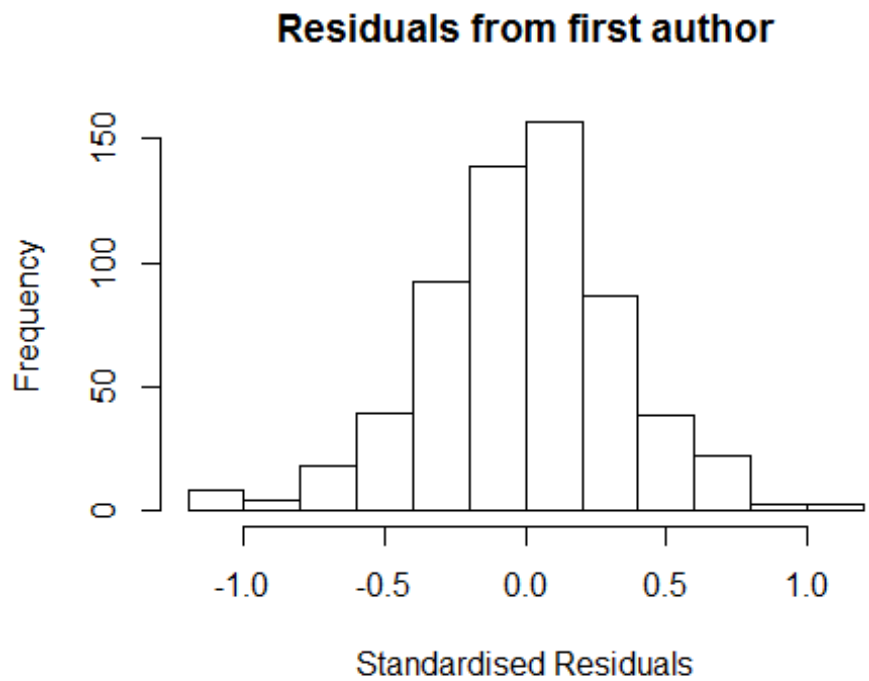

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
UniqueAuthors +
##     Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##     k.max = 1000))
##  \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.08862 -0.19642 -0.00208  0.20102  1.08205
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         0.95733    0.05732   16.70   <2e-16 ***
## FirstAuthorFemale1  0.00307    0.03182    0.10   0.9231
## LastAuthorFemale1  -0.00549    0.03566   -0.15   0.8778
## UniqueAuthors2      0.11468    0.06488    1.77   0.0777 .
## UniqueAuthors3      0.12327    0.06931    1.78   0.0758 .
## UniqueAuthors4      0.15858    0.06527    2.43   0.0154 *
## UniqueAuthors5      0.19090    0.06018    3.17   0.0016 **
## Year1997            0.04248    0.06050    0.70   0.4829
## Year1998           -0.02453    0.06773   -0.36   0.7173
## Year1999            0.01660    0.06385    0.26   0.7949
```

```
## Year2000            0.01330    0.07623    0.17   0.8616
## Year2001           -0.19133    0.09798   -1.95   0.0513 .
## Year2002           -0.10440    0.06625   -1.58   0.1156
## Year2003           -0.06016    0.07966   -0.76   0.4504
## Year2004           -0.17721    0.08871   -2.00   0.0462 *
## Year2005            0.01400    0.07161    0.20   0.8450
## Year2006           -0.22546    0.07656   -2.94   0.0034 **
## Year2007           -0.09272    0.07343   -1.26   0.2072
## Year2008           -0.13414    0.07849   -1.71   0.0880 .
## Year2009           -0.06690    0.07382   -0.91   0.3652
## Year2010           -0.17737    0.08473   -2.09   0.0367 *
## Year2011           -0.10444    0.09250   -1.13   0.2593
## Year2012            0.10872    0.10680    1.02   0.3091
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.302
## Multiple R-squared:  0.0767, Adjusted R-squared:  0.042
## Convergence in 15 IRWLS iterations
##
## Robustness weights:
##  42 weights are ~= 1. The remaining 566 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.167   0.858   0.954   0.891   0.987   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.64e-04          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.260  1           1.122
## LastAuthorFemale  1.206  1           1.098
## Year              1.487 16           1.012
```



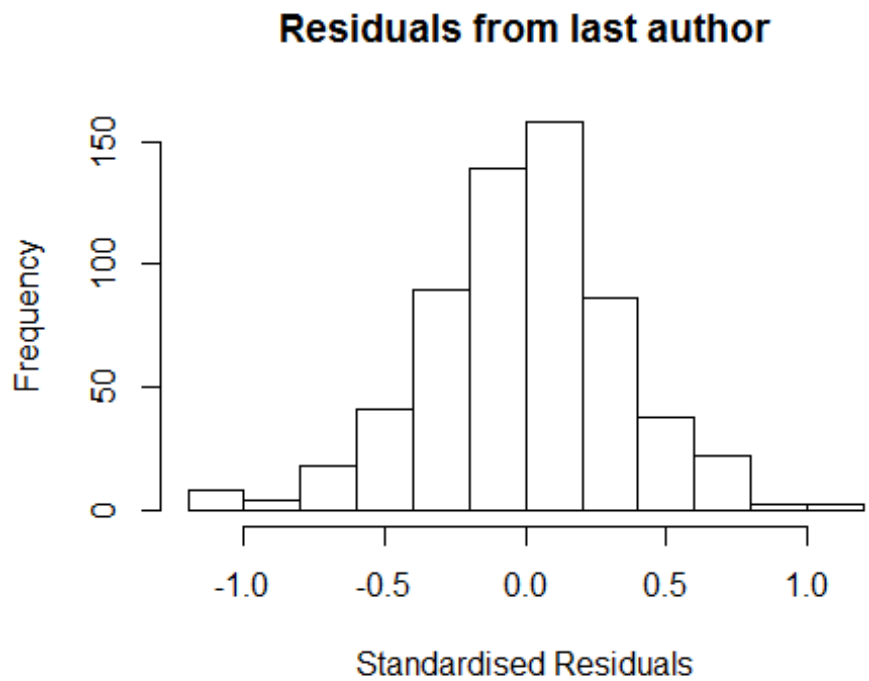

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.10287 -0.20746  0.00481  0.19828  1.12148
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.05100    0.04622   22.74   <2e-16 ***
## FirstAuthorFemale1  0.00821    0.03207    0.26    0.798
## LastAuthorFemale1  -0.00176    0.03531   -0.05    0.960
## Year1997            0.05187    0.06081    0.85    0.394
## Year1998            0.02188    0.06753    0.32    0.746
## Year1999            0.05174    0.06804    0.76    0.447
## Year2000            0.06702    0.07374    0.91    0.364
## Year2001           -0.14232    0.09816   -1.45    0.148
## Year2002           -0.03375    0.06098   -0.55    0.580
## Year2003            0.00344    0.07336    0.05    0.963
## Year2004           -0.10769    0.08299   -1.30    0.195
## Year2005            0.07942    0.06758    1.18    0.240
```

```
## Year2006           -0.15122    0.07246   -2.09    0.037 *
## Year2007           -0.02771    0.06770   -0.41    0.682
## Year2008           -0.07204    0.07982   -0.90    0.367
## Year2009           -0.00272    0.07090   -0.04    0.969
## Year2010           -0.10717    0.08066   -1.33    0.184
## Year2011           -0.03616    0.08541   -0.42    0.672
## Year2012            0.16652    0.10553    1.58    0.115
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.313
## Multiple R-squared:  0.0491, Adjusted R-squared:  0.02
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
##  44 weights are ~= 1. The remaining 564 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.172   0.859   0.954   0.894   0.986   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.64e-04          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.258  1           1.122
## Year              1.258 16           1.007
```




```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.1027 -0.2072  0.0051  0.1977  1.1221
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.05072    0.04621   22.74   <2e-16 ***
## FirstAuthorFemale1  0.00812    0.03208    0.25    0.800
## Year1997            0.05200    0.06084    0.85    0.393
## Year1998            0.02205    0.06754    0.33    0.744
## Year1999            0.05177    0.06781    0.76    0.445
## Year2000            0.06701    0.07356    0.91    0.363
## Year2001           -0.14213    0.09821   -1.45    0.148
## Year2002           -0.03367    0.06094   -0.55    0.581
## Year2003            0.00337    0.07339    0.05    0.963
## Year2004           -0.10755    0.08309   -1.29    0.196
## Year2005            0.07938    0.06721    1.18    0.238
## Year2006           -0.15153    0.07126   -2.13    0.034 *
```

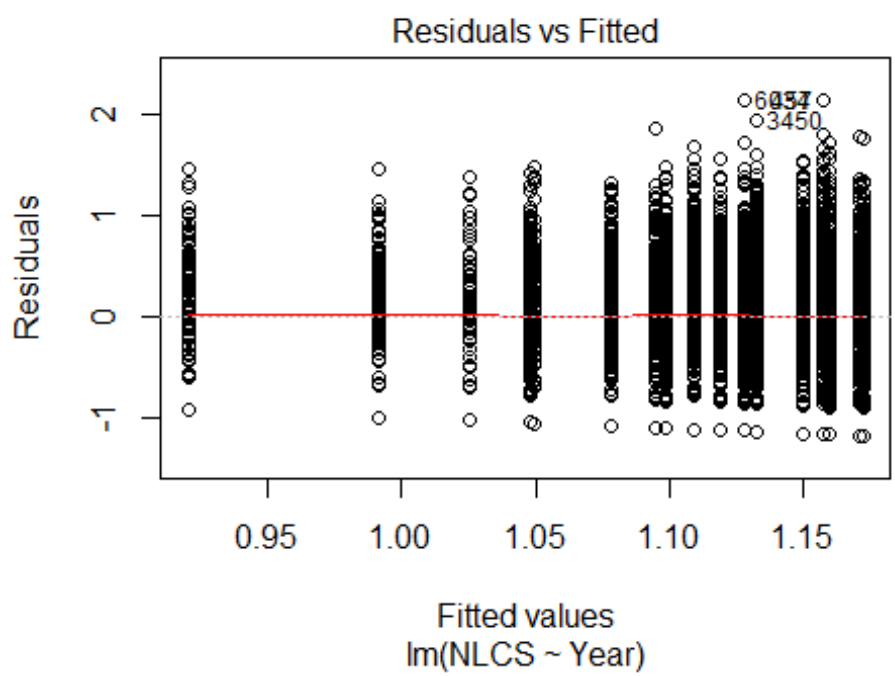
```
## Year2007           -0.02781    0.06755   -0.41    0.681
## Year2008           -0.07188    0.07966   -0.90    0.367
## Year2009           -0.00225    0.07081   -0.03    0.975
## Year2010           -0.10737    0.08053   -1.33    0.183
## Year2011           -0.03646    0.08441   -0.43    0.666
## Year2012            0.16622    0.10478    1.59    0.113
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.312
## Multiple R-squared:  0.0491, Adjusted R-squared:  0.0217
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
##  45 weights are ~= 1. The remaining 563 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.170   0.859   0.954   0.894   0.986   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.64e-04          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##                   GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.206  1           1.098
## Year             1.206 16           1.006
```




```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.10414 -0.20672  0.00865  0.19696  1.11683
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.05205    0.04587   22.93   <2e-16 ***
## LastAuthorFemale1 -0.00126    0.03536   -0.04    0.971
## Year1997            0.05209    0.06084    0.86    0.392
## Year1998            0.02342    0.06720    0.35    0.728
## Year1999            0.05178    0.06792    0.76    0.446
## Year2000            0.06729    0.07368    0.91    0.361
## Year2001           -0.14195    0.09856   -1.44    0.150
## Year2002           -0.03203    0.06050   -0.53    0.597
## Year2003            0.00331    0.07329    0.05    0.964
## Year2004           -0.10655    0.08327   -1.28    0.201
## Year2005            0.08154    0.06680    1.22    0.223
## Year2006           -0.14810    0.07139   -2.07    0.038 *
```

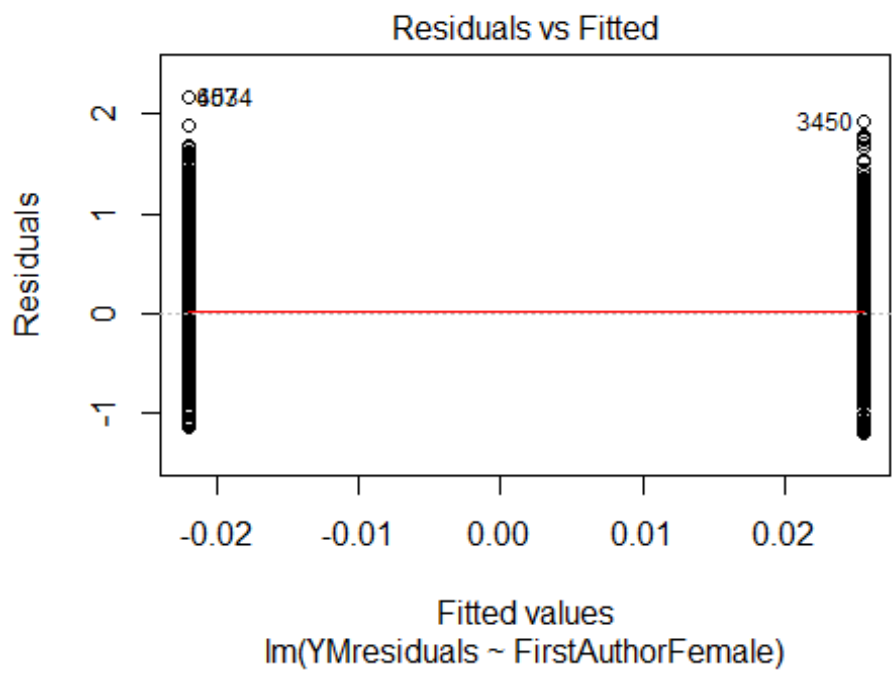
```
## Year2007          -0.02617    0.06772   -0.39    0.699
## Year2008          -0.07129    0.07976   -0.89    0.372
## Year2009          -0.00192    0.07075   -0.03    0.978
## Year2010          -0.10461    0.07985   -1.31    0.191
## Year2011          -0.03529    0.08528   -0.41    0.679
## Year2012           0.17013    0.10313    1.65    0.100 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.312
## Multiple R-squared:  0.0491, Adjusted R-squared:  0.0217
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
##  44 weights are ~= 1. The remaining 564 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.174   0.860   0.954   0.894   0.986   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.64e-04          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Sample size for the above analysis:  608"
## [1] ""
## [1] ""
## [1] "####################################"
## [1] "Analysis of AJSC 3600"
## [1] "####################################"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  481  489  433  500  475  589  638  440  529  583  557  600  516  163  153
## 2011 2012
##  120  135
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  337  324  291  370  268  317  546  393  437  505  486  524  435  144  140
## 2011 2012
```

```
##   97  108
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  306  298  259  336  244  288  511  357  406  454  449  471  390  123  114
## 2011 2012
##   86   97
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
##  Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 80, df = 16, p-value = 2e-10
```

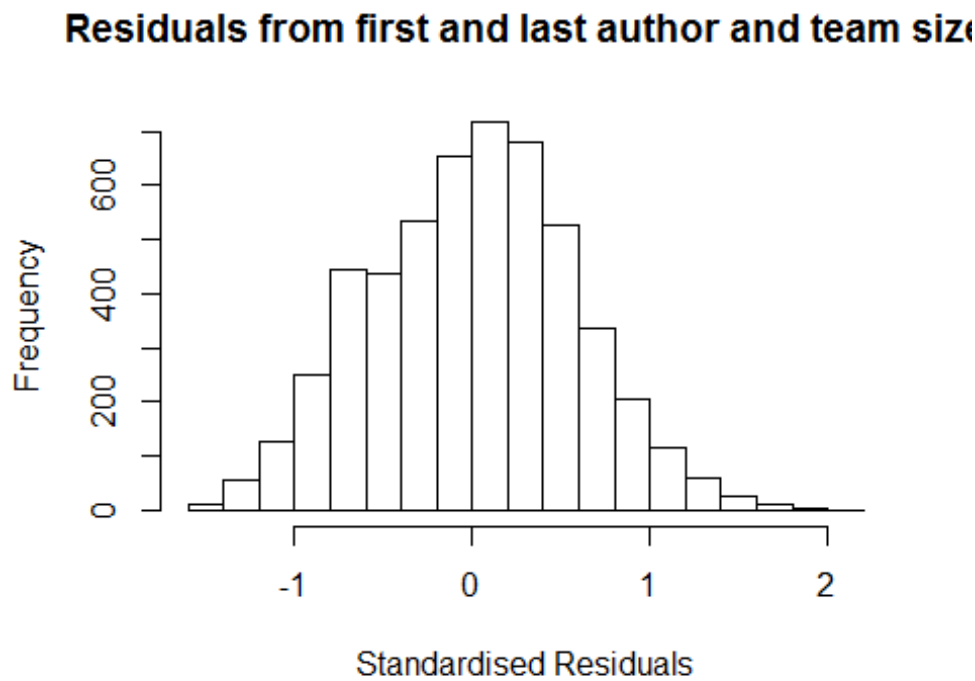



```
##
##  Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 3.1, df = 1, p-value = 0.08

## [1] "Female first author team size 2018 geometric mean: 2.54883016866429"
## [1] "Male first author team size 2018 geometric mean: 2.32194768848717"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 1000, p-value = 0.5
```

```
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 2.5242947636763"
## [1] "Male last author team size 2018 geometric mean: 2.38007194059049"

## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties
```

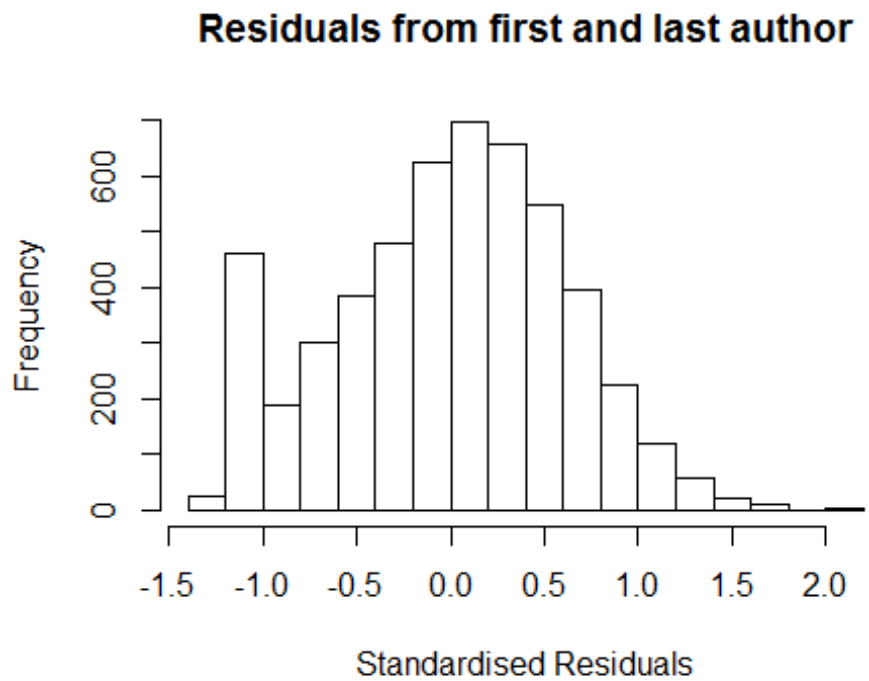



```
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 1000, p-value = 0.7
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##                     GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.130  1           1.063
## LastAuthorFemale  1.140  1           1.068
## UniqueAuthors     1.098  4           1.012
## Year              1.070 16           1.002
```



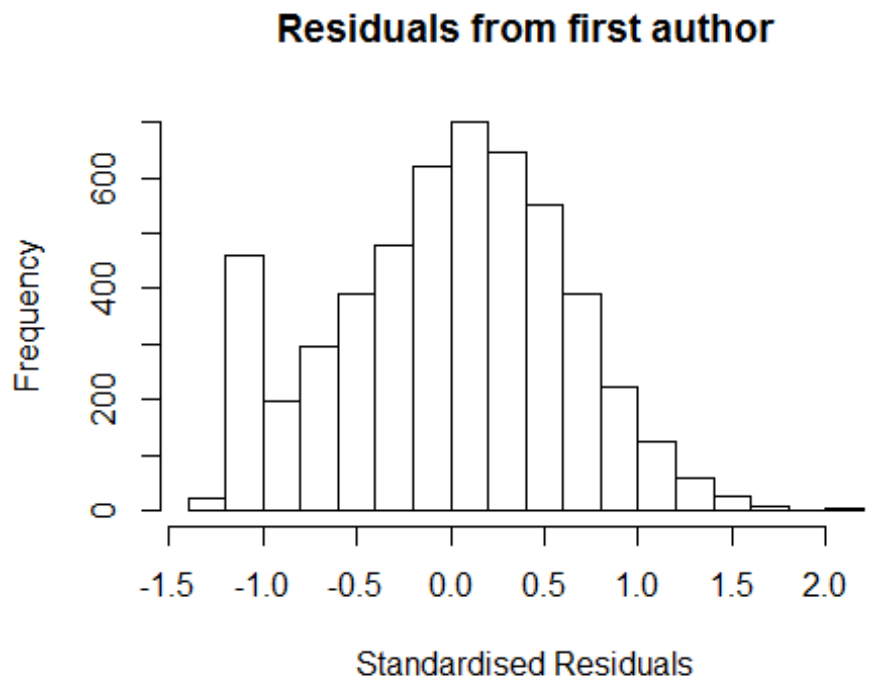

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
UniqueAuthors +
##     Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##     k.max = 1000))
##  \--> method = "MM"
## Residuals:
##    Min     1Q Median     3Q    Max
## -1.539 -0.412  0.023  0.391  2.101
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)          0.87195    0.04796   18.18  < 2e-16 ***
## FirstAuthorFemale1   0.03041    0.01759    1.73  0.08382 .
## LastAuthorFemale1    0.00189    0.01778    0.11  0.91530
## UniqueAuthors2       0.32386    0.02469   13.12  < 2e-16 ***
## UniqueAuthors3       0.42750    0.02519   16.97  < 2e-16 ***
## UniqueAuthors4       0.51707    0.02847   18.16  < 2e-16 ***
## UniqueAuthors5       0.63619    0.02552   24.93  < 2e-16 ***
## Year1997            -0.07314    0.06129   -1.19  0.23283
## Year1998             0.00796    0.06472    0.12  0.90212
## Year1999             0.01178    0.05728    0.21  0.83704
```

```
## Year2000           -0.12680    0.05840   -2.17  0.02997 *
## Year2001           -0.11329    0.05989   -1.89  0.05860 .
## Year2002           -0.08650    0.05172   -1.67  0.09449 .
## Year2003           -0.13232    0.05371   -2.46  0.01379 *
## Year2004           -0.09328    0.05439   -1.72  0.08638 .
## Year2005           -0.08358    0.05235   -1.60  0.11044
## Year2006           -0.12448    0.05238   -2.38  0.01751 *
## Year2007           -0.14310    0.05119   -2.80  0.00520 **
## Year2008           -0.19710    0.05295   -3.72  0.00020 ***
## Year2009           -0.26920    0.07289   -3.69  0.00022 ***
## Year2010           -0.22885    0.07054   -3.24  0.00118 **
## Year2011           -0.18554    0.07816   -2.37  0.01765 *
## Year2012           -0.19392    0.07737   -2.51  0.01222 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.587
## Multiple R-squared:  0.14,   Adjusted R-squared:  0.137
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
##  451 weights are ~= 1. The remaining 4738 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.173   0.864   0.949   0.910   0.985   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.93e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.084  1           1.041
## LastAuthorFemale  1.081  1           1.040
## Year              1.023 16           1.001
```



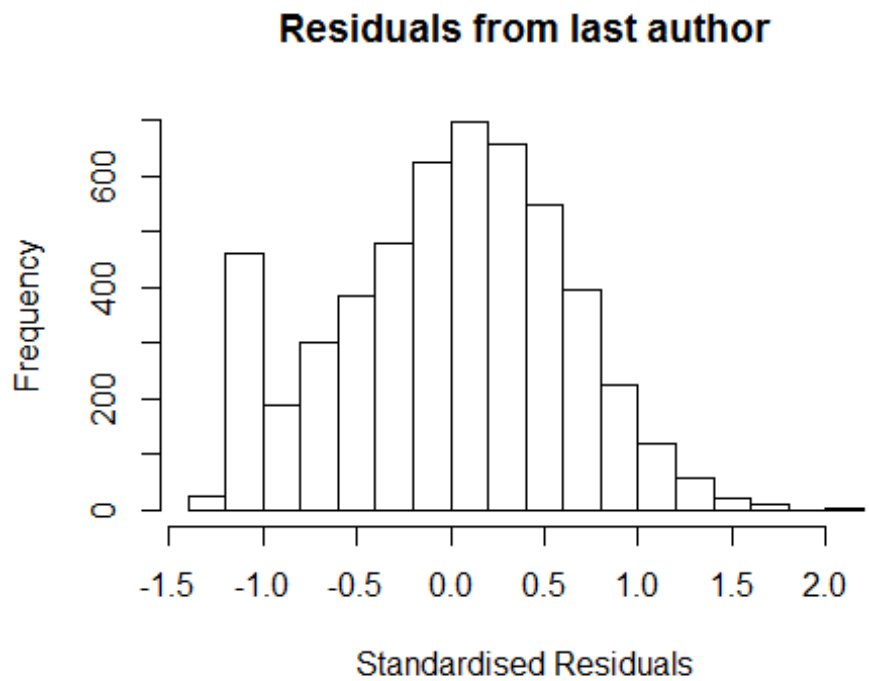

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.2221 -0.4274  0.0388  0.4285  2.1828
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)          1.15073    0.05144   22.37  < 2e-16 ***
## FirstAuthorFemale1   0.06159    0.01864    3.30  0.00096 ***
## LastAuthorFemale1   -0.01986    0.01872   -1.06  0.28879
## Year1997            -0.06447    0.06774   -0.95  0.34123
## Year1998             0.00912    0.06856    0.13  0.89420
## Year1999             0.00977    0.06205    0.16  0.87490
## Year2000            -0.09500    0.06323   -1.50  0.13304
## Year2001            -0.10871    0.06387   -1.70  0.08883 .
## Year2002            -0.03660    0.05811   -0.63  0.52889
## Year2003            -0.07673    0.05903   -1.30  0.19372
## Year2004            -0.02828    0.06026   -0.47  0.63891
## Year2005            -0.01836    0.05820   -0.32  0.75247
```

```
## Year2006           -0.04170    0.05822   -0.72  0.47392
## Year2007           -0.05724    0.05660   -1.01  0.31196
## Year2008           -0.12481    0.05878   -2.12  0.03377 *
## Year2009           -0.25065    0.08236   -3.04  0.00235 **
## Year2010           -0.24079    0.07967   -3.02  0.00252 **
## Year2011           -0.13277    0.08430   -1.57  0.11532
## Year2012           -0.14482    0.08669   -1.67  0.09485 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.617
## Multiple R-squared:  0.0102, Adjusted R-squared:  0.00673
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
##  419 weights are ~= 1. The remaining 4770 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.184   0.860   0.948   0.908   0.986   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.93e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.014  1           1.007
## Year              1.014 16           1.000
```




```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.2104 -0.4293  0.0397  0.4286  2.1681
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)          1.14519    0.05080   22.54   <2e-16 ***
## FirstAuthorFemale1  0.05452    0.01813    3.01   0.0027 **
## Year1997           -0.06375    0.06757   -0.94   0.3455
## Year1998            0.00891    0.06849    0.13   0.8965
## Year1999            0.01074    0.06192    0.17   0.8623
## Year2000           -0.09485    0.06309   -1.50   0.1328
## Year2001           -0.10981    0.06383   -1.72   0.0854 .
## Year2002           -0.03618    0.05800   -0.62   0.5328
## Year2003           -0.07688    0.05895   -1.30   0.1922
## Year2004           -0.02816    0.06017   -0.47   0.6398
## Year2005           -0.01824    0.05809   -0.31   0.7535
## Year2006           -0.04127    0.05810   -0.71   0.4775
```

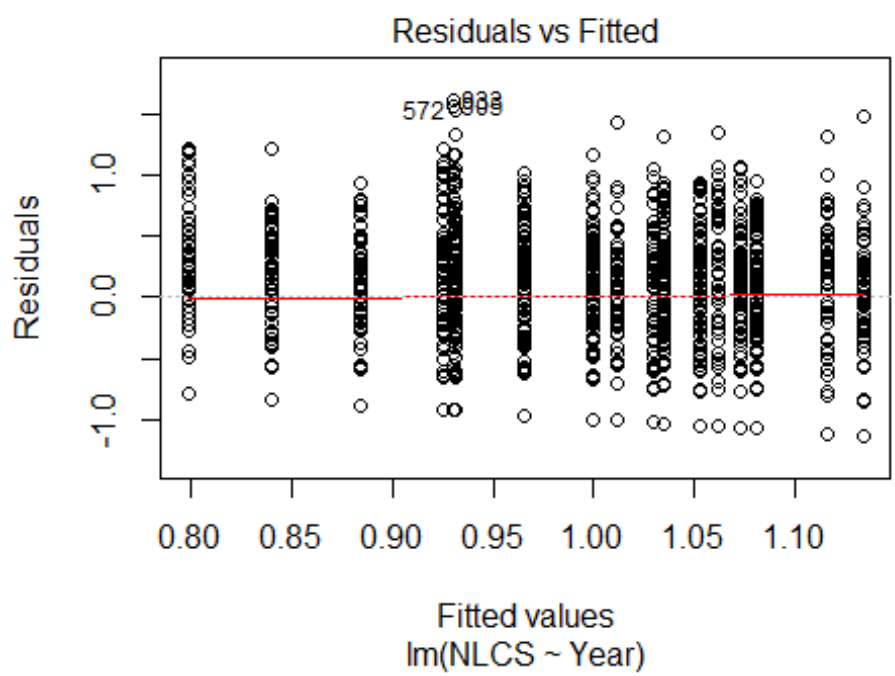
```
## Year2007           -0.05674    0.05650   -1.00   0.3153
## Year2008           -0.12405    0.05866   -2.11   0.0345 *
## Year2009           -0.25193    0.08213   -3.07   0.0022 **
## Year2010           -0.24047    0.07925   -3.03   0.0024 **
## Year2011           -0.13283    0.08406   -1.58   0.1141
## Year2012           -0.14718    0.08654   -1.70   0.0890 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.617
## Multiple R-squared:  0.00998,    Adjusted R-squared:  0.00673
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
##  432 weights are ~= 1. The remaining 4757 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.192   0.860   0.947   0.908   0.985   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.93e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##                   GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.011  1           1.006
## Year             1.011 16           1.000
```



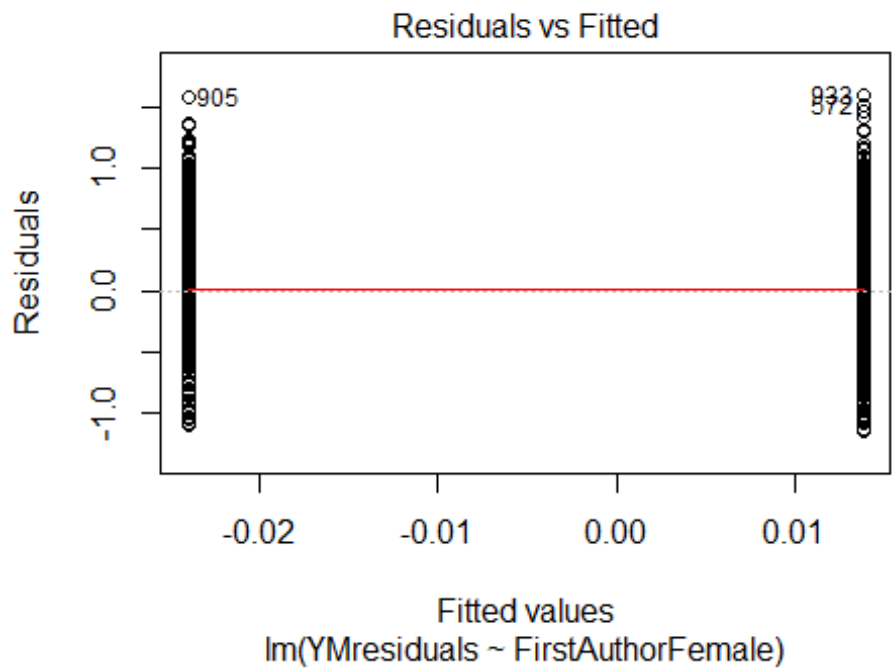

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.1805 -0.4274  0.0426  0.4249  2.1438
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.16436    0.05143   22.64   <2e-16 ***
## LastAuthorFemale1   0.00226    0.01820    0.12   0.9011
## Year1997           -0.05731    0.06788   -0.84   0.3986
## Year1998            0.01294    0.06868    0.19   0.8505
## Year1999            0.01385    0.06208    0.22   0.8235
## Year2000           -0.09012    0.06324   -1.43   0.1542
## Year2001           -0.10173    0.06387   -1.59   0.1113
## Year2002           -0.03299    0.05814   -0.57   0.5704
## Year2003           -0.06988    0.05905   -1.18   0.2367
## Year2004           -0.02185    0.06038   -0.36   0.7175
## Year2005           -0.01426    0.05827   -0.24   0.8067
## Year2006           -0.03844    0.05828   -0.66   0.5096
```

```
## Year2007          -0.05085    0.05667   -0.90   0.3696
## Year2008          -0.11895    0.05889   -2.02   0.0435 *
## Year2009          -0.24263    0.08252   -2.94   0.0033 **
## Year2010          -0.23062    0.07994   -2.89   0.0039 **
## Year2011          -0.12195    0.08447   -1.44   0.1489
## Year2012          -0.13519    0.08668   -1.56   0.1189
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.617
## Multiple R-squared:  0.00811,    Adjusted R-squared:  0.00485
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
##  435 weights are ~= 1. The remaining 4754 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.202   0.859   0.947   0.908   0.985   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.93e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Sample size for the above analysis:  5189"
## [1] ""
## [1] ""
## [1] "####################################"
## [1] "Analysis of AJSC 3601"
## [1] "####################################"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##   92   79   75   84   76  131   87   90   73   73   72  102  127  161  163
## 2011 2012
##  158  167
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##   83   68   65   78   65  105   73   78   62   70   66   91  109  143  145
## 2011 2012
```

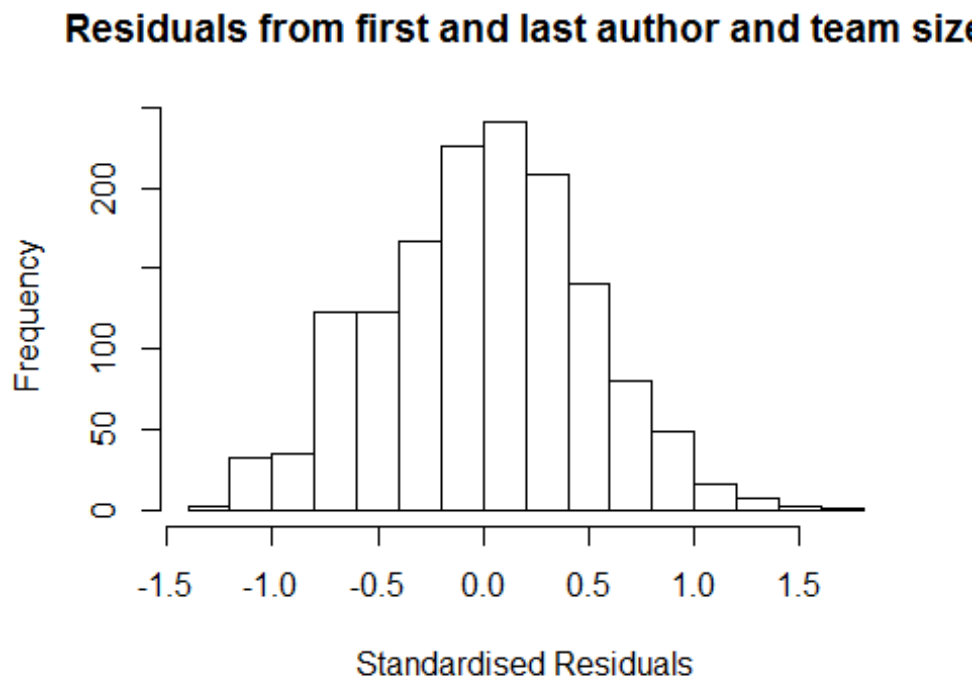
```
##  140  147
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##   80   66   61   70   58   98   63   74   59   66   57   83   98  130  132
## 2011 2012
##  129  132
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
##  Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 43, df = 16, p-value = 3e-04
```




```
##
##  Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 4, df = 1, p-value = 0.04
```

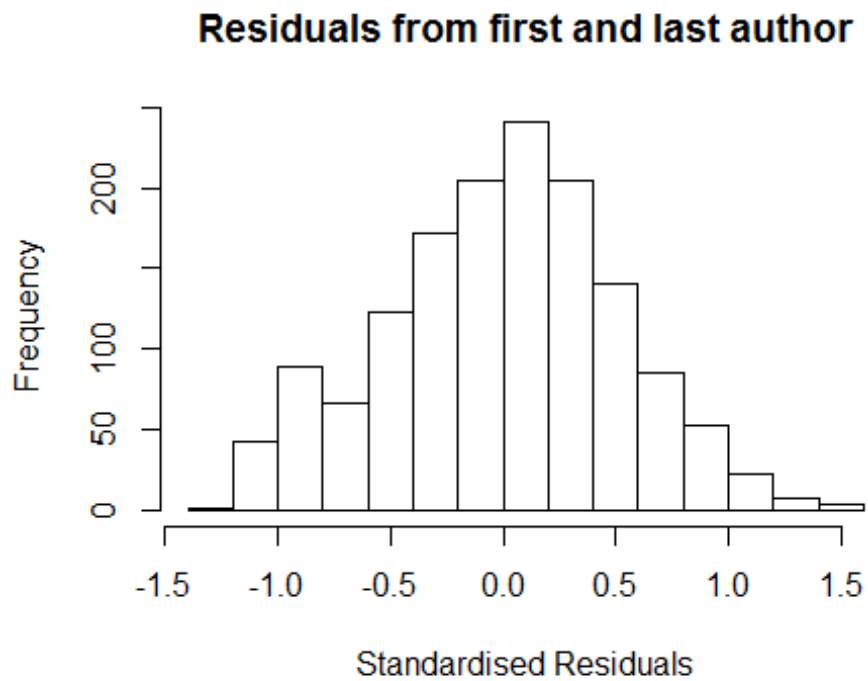



```
## [1] "Female first author team size 2018 geometric mean: 3.53045740291816"
## [1] "Male first author team size 2018 geometric mean: 2.99381610780763"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 1700, p-value = 0.3
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 3.21752006217463"
## [1] "Male last author team size 2018 geometric mean: 3.76304001877365"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 1600, p-value = 0.08
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.321  1           1.149
## LastAuthorFemale  1.384  1           1.176
## UniqueAuthors     1.332  4           1.037
## Year              1.410 16           1.011
```




```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
UniqueAuthors +
##     Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##     k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.3832 -0.3488  0.0156  0.3366  1.7838
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)           0.5851     0.0655    8.93  < 2e-16 ***
## FirstAuthorFemale1    0.0539     0.0333    1.62  0.10607
## LastAuthorFemale1    -0.0445     0.0317   -1.40  0.16144
## UniqueAuthors2        0.2497     0.0409    6.10  1.4e-09 ***
## UniqueAuthors3        0.3701     0.0425    8.71  < 2e-16 ***
## UniqueAuthors4        0.3723     0.0457    8.15  7.7e-16 ***
## UniqueAuthors5        0.4456     0.0416   10.71  < 2e-16 ***
## Year1997              0.0834     0.0864    0.97  0.33419
## Year1998              0.2985     0.0893    3.34  0.00086 ***
## Year1999              0.2513     0.0765    3.29  0.00104 **
```

```
## Year2000             0.2145     0.0938    2.29  0.02235 *
## Year2001             0.0916     0.0796    1.15  0.24972
## Year2002             0.0316     0.0851    0.37  0.71064
## Year2003             0.0453     0.0982    0.46  0.64453
## Year2004             0.1831     0.1050    1.74  0.08140 .
## Year2005             0.1283     0.1080    1.19  0.23511
## Year2006             0.2547     0.0922    2.76  0.00580 **
## Year2007             0.1808     0.0804    2.25  0.02471 *
## Year2008             0.2043     0.0745    2.74  0.00615 **
## Year2009             0.1623     0.0755    2.15  0.03161 *
## Year2010             0.1900     0.0735    2.59  0.00980 **
## Year2011             0.0987     0.0729    1.35  0.17623
## Year2012             0.1031     0.0740    1.39  0.16347
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.501
## Multiple R-squared:  0.136,  Adjusted R-squared:  0.123
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
##  144 weights are ~= 1. The remaining 1312 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.178   0.865   0.949   0.907   0.985   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          6.87e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.275  1           1.129
## LastAuthorFemale  1.301  1           1.141
## Year              1.087 16           1.003
```

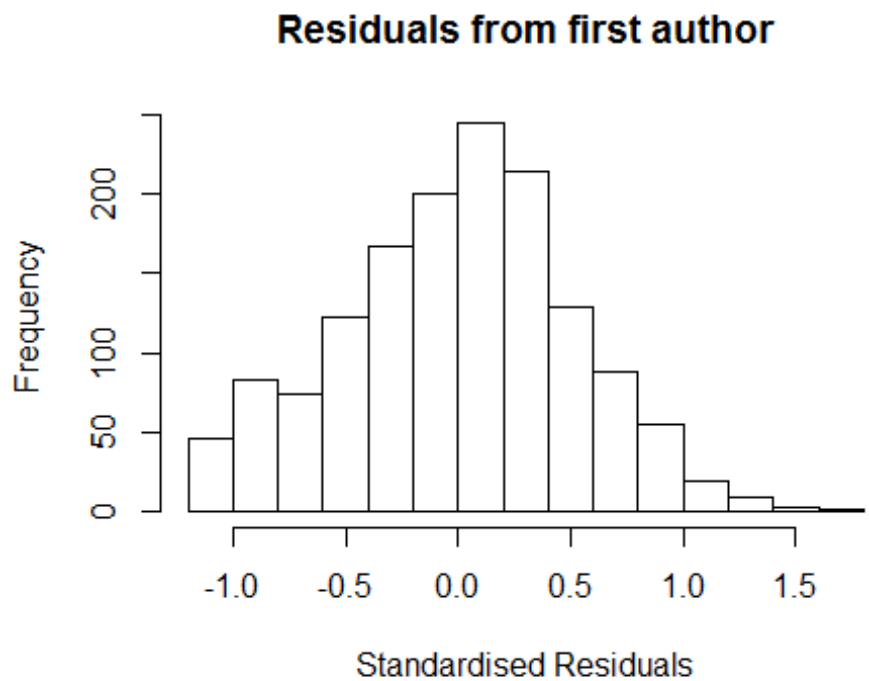



```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.2179 -0.3482  0.0267  0.3386  1.5824
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)           0.7703     0.0709   10.86  < 2e-16 ***
## FirstAuthorFemale1    0.0848     0.0347    2.45  0.01452 *
## LastAuthorFemale1    -0.0921     0.0326   -2.82  0.00480 **
## Year1997              0.1170     0.0931    1.26  0.20925
## Year1998              0.3628     0.1020    3.56  0.00039 ***
## Year1999              0.2673     0.0829    3.22  0.00129 **
## Year2000              0.2886     0.0976    2.96  0.00317 **
## Year2001              0.1246     0.0908    1.37  0.17010
## Year2002              0.0794     0.0979    0.81  0.41772
## Year2003              0.1226     0.1082    1.13  0.25716
## Year2004              0.2999     0.1141    2.63  0.00869 **
## Year2005              0.1475     0.1220    1.21  0.22678
```

```
## Year2006             0.3530     0.0924    3.82  0.00014 ***
## Year2007             0.2743     0.0891    3.08  0.00211 **
## Year2008             0.2927     0.0811    3.61  0.00032 ***
## Year2009             0.2335     0.0835    2.80  0.00522 **
## Year2010             0.3170     0.0801    3.96  7.9e-05 ***
## Year2011             0.2133     0.0804    2.65  0.00809 **
## Year2012             0.1931     0.0816    2.36  0.01817 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.521
## Multiple R-squared:  0.0379, Adjusted R-squared:  0.0259
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  141 weights are ~= 1. The remaining 1315 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.337   0.863   0.951   0.905   0.984   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          6.87e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.031  1           1.015
## Year              1.031 16           1.001

## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
```

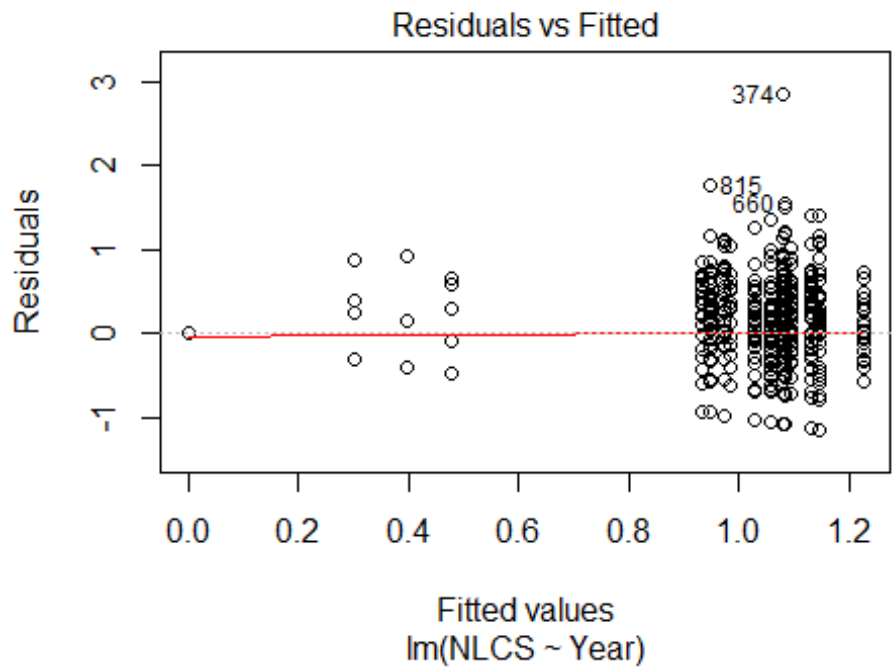
```
## -1.1541 -0.3475  0.0246  0.3424  1.6011
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)          0.7377     0.0712   10.36  < 2e-16 ***
## FirstAuthorFemale1   0.0440     0.0311    1.41  0.15733
## Year1997             0.1292     0.0938    1.38  0.16879
## Year1998             0.3724     0.1032    3.61  0.00032 ***
## Year1999             0.2685     0.0835    3.21  0.00134 **
## Year2000             0.2818     0.0984    2.86  0.00424 **
## Year2001             0.1297     0.0920    1.41  0.15904
## Year2002             0.0894     0.0990    0.90  0.36658
## Year2003             0.1256     0.1083    1.16  0.24651
## Year2004             0.3030     0.1151    2.63  0.00859 **
## Year2005             0.1632     0.1226    1.33  0.18334
## Year2006             0.3561     0.0932    3.82  0.00014 ***
## Year2007             0.2751     0.0900    3.06  0.00228 **
## Year2008             0.3029     0.0824    3.68  0.00025 ***
## Year2009             0.2437     0.0843    2.89  0.00391 **
## Year2010             0.3241     0.0807    4.01  6.3e-05 ***
## Year2011             0.2293     0.0812    2.83  0.00479 **
## Year2012             0.2046     0.0825    2.48  0.01330 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.522
## Multiple R-squared:  0.032,  Adjusted R-squared:  0.0205
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
##  119 weights are ~= 1. The remaining 1337 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.326   0.868   0.953   0.907   0.986   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          6.87e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
```

```
## Warning in lmrob.S(x, y, control = control, mf = mf): S refinements did
not
## converge (to refine.tol=1e-07) in 1000 (= k.max) steps
```



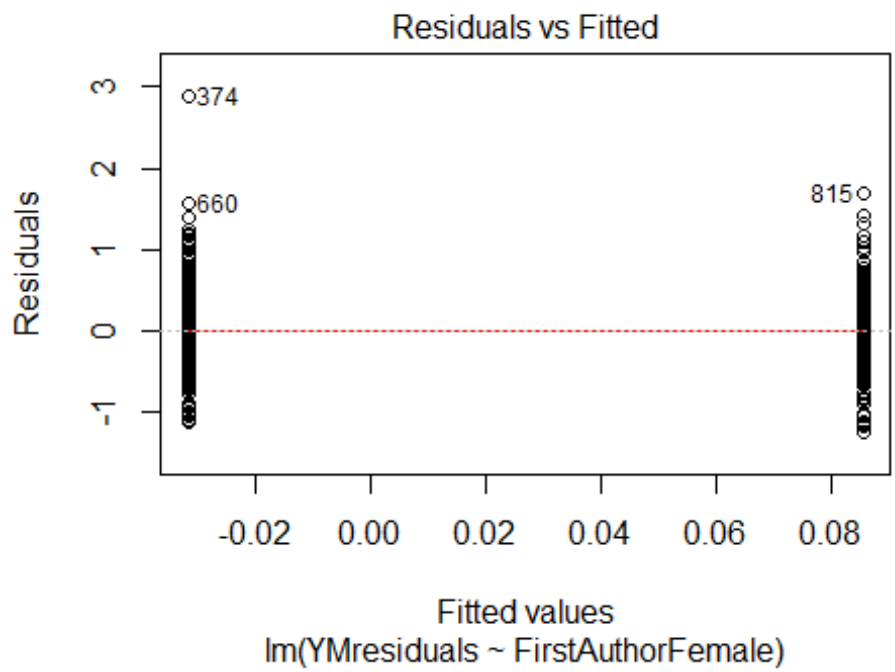

```
## [1] "Sample size for the above analysis:  1456"
## [1] ""
## [1] ""
## [1] "####################################"
## [1] "Analysis of AJSC 3602"
## [1] "####################################"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##   13   15   20   31   38   40   36   31   25   56   69   56   59   72   60
## 2011 2012
##   82   58
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    7    6    8    2   33   31   29   25   23   53   62   47   49   64   57
## 2011 2012
##   72   48
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    7    6    8    2   31   30   27   21   22   49   60   44   46   57   50
## 2011 2012
##   65   41
```

```
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
##  Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = Inf, df = 16, p-value <2e-16
```

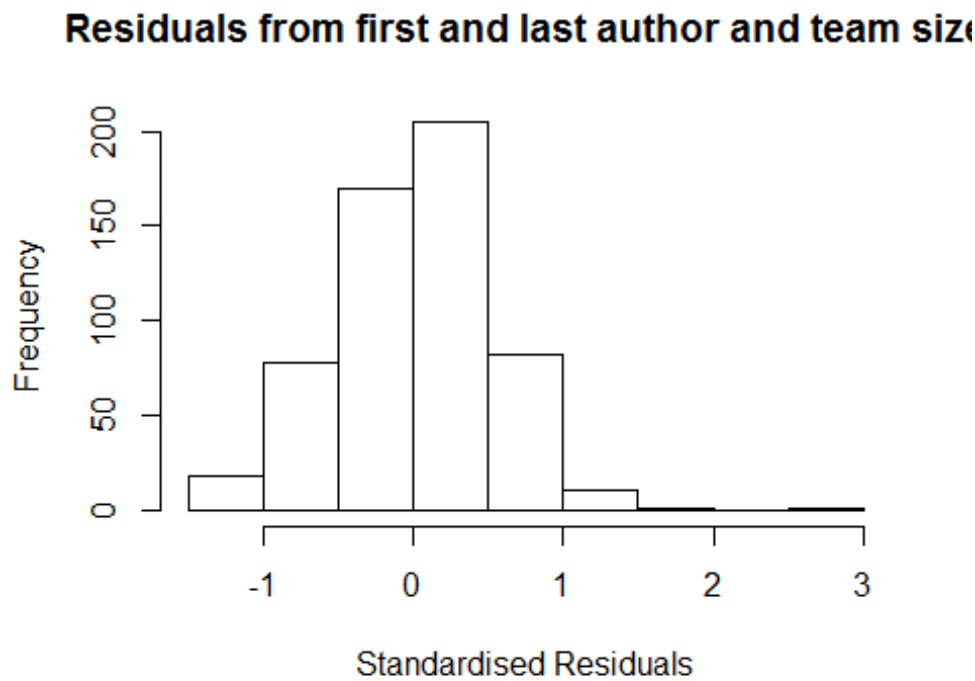



```
##
##  Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 3.3, df = 1, p-value = 0.07

## [1] "Female first author team size 2018 geometric mean: 3.52795024181346"
## [1] "Male first author team size 2018 geometric mean: 2.53456500823867"

## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties

##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 340, p-value = 0.1
## alternative hypothesis: true location shift is not equal to 0
##
```

```
## [1] "Female last author team size 2018 geometric mean: 2.98539388081357"
## [1] "Male last author team size 2018 geometric mean: 2.85254264826902"

## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties
```

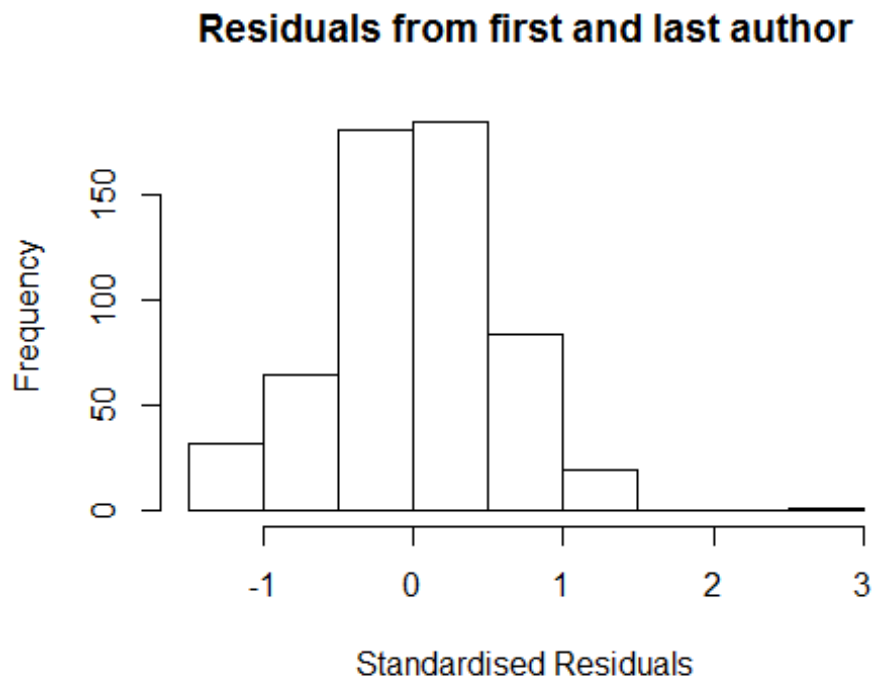



```
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 240, p-value = 0.8
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##                     GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.766  1           1.329
## LastAuthorFemale  1.663  1           1.290
## UniqueAuthors     5.240  4           1.230
## Year              8.303 16           1.068
```



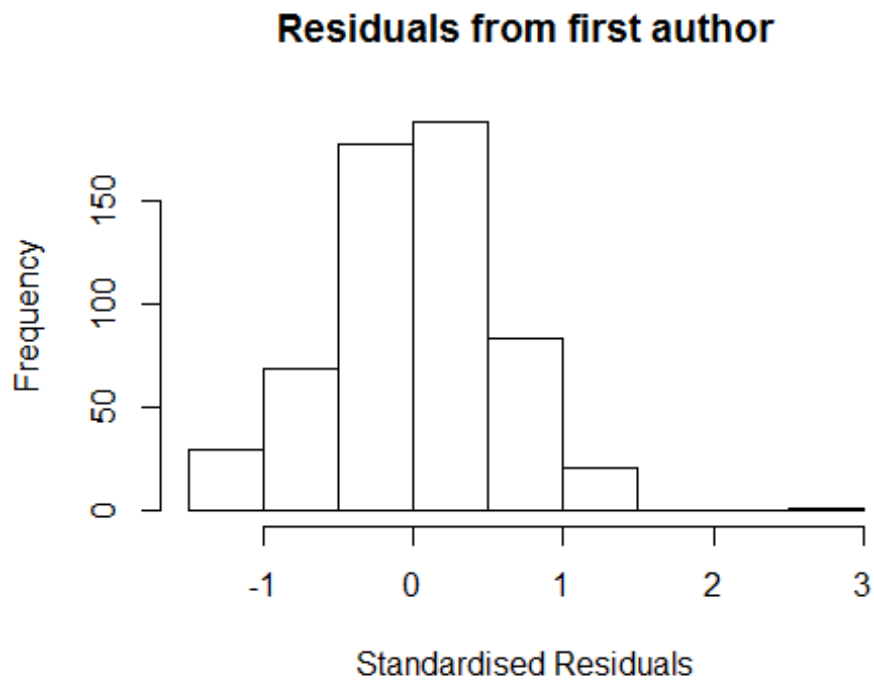

```
## [1] "List of  1 outliers with residuals above  2.5"
##        ScopusId  NLCS Year OneField Fields residuals
## 374 34848925195 3.932 2006     3602      1     2.822
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
UniqueAuthors +
##     Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##     k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.4053 -0.3448  0.0455  0.3436  2.8224
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)          0.4404     0.2226    1.98   0.0484 *
## FirstAuthorFemale1   0.0861     0.0584    1.47   0.1411
## LastAuthorFemale1   -0.0707     0.0568   -1.24   0.2139
## UniqueAuthors2       0.1907     0.0686    2.78   0.0056 **
## UniqueAuthors3       0.2666     0.0732    3.64   0.0003 ***
## UniqueAuthors4       0.3667     0.0916    4.00  7.1e-05 ***
## UniqueAuthors5       0.5623     0.0742    7.58  1.5e-13 ***
## Year1997            -0.0800     0.2980   -0.27   0.7884
## Year1998            -0.1663     0.2700   -0.62   0.5382
## Year1999            -0.4481     0.2256   -1.99   0.0475 *
```

```
## Year2000             0.4351     0.2336    1.86   0.0631 .
## Year2001             0.5527     0.2377    2.32   0.0204 *
## Year2002             0.2092     0.2452    0.85   0.3941
## Year2003             0.2882     0.2474    1.16   0.2446
## Year2004             0.4018     0.2537    1.58   0.1139
## Year2005             0.4634     0.2534    1.83   0.0680 .
## Year2006             0.4026     0.2524    1.60   0.1112
## Year2007             0.5004     0.2411    2.08   0.0384 *
## Year2008             0.4121     0.2382    1.73   0.0843 .
## Year2009             0.3391     0.2396    1.42   0.1576
## Year2010             0.3532     0.2328    1.52   0.1299
## Year2011             0.2290     0.2325    0.99   0.3250
## Year2012             0.2342     0.2460    0.95   0.3415
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.525
## Multiple R-squared:  0.187,  Adjusted R-squared:  0.154
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
##  observation 217 is an outlier with |weight| = 0 ( < 0.00018);
##  32 weights are ~= 1. The remaining 533 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.375   0.874   0.958   0.911   0.986   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.77e-04          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.301  1           1.141
## LastAuthorFemale  1.302  1           1.141
## Year              1.404 16           1.011
```



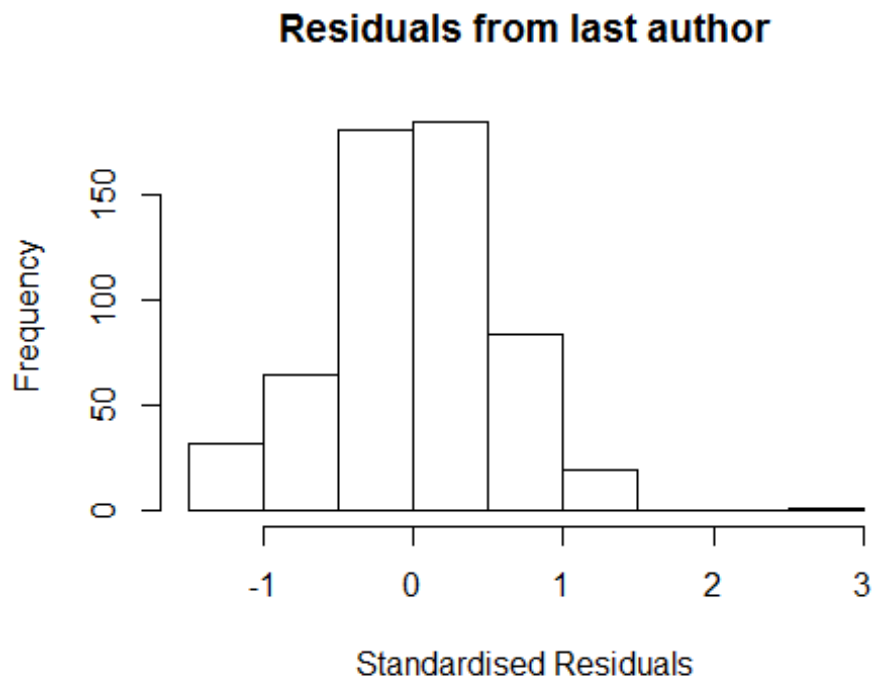

```
## [1] "List of  1 outliers with residuals above  2.5"
##        ScopusId  NLCS Year OneField Fields residuals
## 374 34848925195 3.932 2006     3602      1      2.91
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.2125 -0.3649  0.0151  0.3827  2.9104
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)           0.4542     0.2090    2.17  0.03017 *
## FirstAuthorFemale1    0.1325     0.0624    2.12  0.03404 *
## LastAuthorFemale1    -0.0276     0.0593   -0.47  0.64153
## Year1997             -0.0886     0.2878   -0.31  0.75830
## Year1998             -0.1890     0.2560   -0.74  0.46067
## Year1999             -0.5067     0.2145   -2.36  0.01851 *
## Year2000              0.5758     0.2190    2.63  0.00880 **
## Year2001              0.7348     0.2182    3.37  0.00081 ***
## Year2002              0.4142     0.2407    1.72  0.08585 .
## Year2003              0.4577     0.2366    1.93  0.05354 .
## Year2004              0.5295     0.2516    2.10  0.03582 *
## Year2005              0.6257     0.2438    2.57  0.01052 *
```

```
## Year2006             0.5673     0.2377    2.39  0.01734 *
## Year2007             0.6581     0.2290    2.87  0.00422 **
## Year2008             0.6033     0.2207    2.73  0.00646 **
## Year2009             0.5573     0.2228    2.50  0.01267 *
## Year2010             0.5458     0.2184    2.50  0.01273 *
## Year2011             0.4054     0.2185    1.85  0.06414 .
## Year2012             0.4547     0.2358    1.93  0.05437 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.561
## Multiple R-squared:  0.0897, Adjusted R-squared:  0.0598
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
##  observation 217 is an outlier with |weight| = 0 ( < 0.00018);
##  49 weights are ~= 1. The remaining 516 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.469   0.881   0.949   0.911   0.984   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.77e-04          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                  psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.187  1           1.089
## Year              1.187 16           1.005
```




```
## [1] "List of  1 outliers with residuals above  2.5"
##        ScopusId  NLCS Year OneField Fields residuals
## 374 34848925195 3.932 2006     3602      1      2.91
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.1976 -0.3655  0.0164  0.3847  2.9203
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)           0.4518     0.2107    2.14  0.03247 *
## FirstAuthorFemale1    0.1234     0.0608    2.03  0.04282 *
## Year1997             -0.0863     0.2891   -0.30  0.76542
## Year1998             -0.1893     0.2570   -0.74  0.46178
## Year1999             -0.5135     0.2167   -2.37  0.01815 *
## Year2000              0.5763     0.2208    2.61  0.00931 **
## Year2001              0.7326     0.2198    3.33  0.00092 ***
## Year2002              0.4140     0.2422    1.71  0.08799 .
## Year2003              0.4620     0.2381    1.94  0.05288 .
## Year2004              0.5289     0.2532    2.09  0.03716 *
## Year2005              0.6223     0.2449    2.54  0.01134 *
## Year2006              0.5599     0.2379    2.35  0.01897 *
```

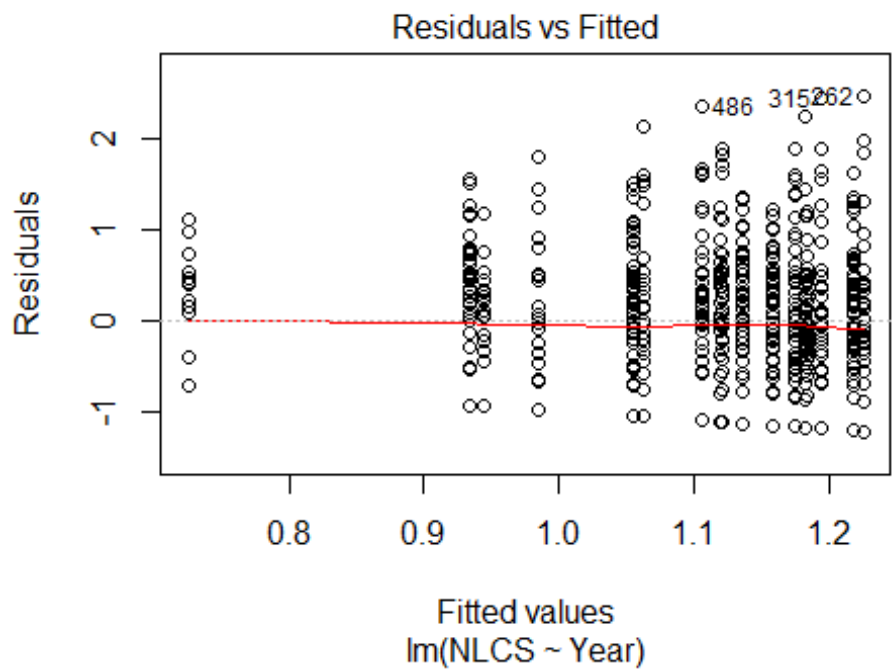
```
## Year2007             0.6565     0.2306    2.85  0.00458 **
## Year2008             0.5987     0.2221    2.70  0.00724 **
## Year2009             0.5542     0.2244    2.47  0.01384 *
## Year2010             0.5412     0.2197    2.46  0.01405 *
## Year2011             0.4024     0.2200    1.83  0.06793 .
## Year2012             0.4493     0.2368    1.90  0.05826 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.56
## Multiple R-squared:  0.0895, Adjusted R-squared:  0.0612
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
##  observation 217 is an outlier with |weight| = 0 ( < 0.00018);
##  44 weights are ~= 1. The remaining 521 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.475   0.882   0.950   0.911   0.986   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.77e-04          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##                   GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.478  1           1.216
## Year             1.478 16           1.012
```




```
## [1] "List of  1 outliers with residuals above  2.5"
##        ScopusId  NLCS Year OneField Fields residuals
## 374 34848925195 3.932 2006     3602      1      2.91
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.14340 -0.36612  0.00362  0.38091  2.89094
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)          0.4667     0.2011    2.32  0.02070 *
## LastAuthorFemale1    0.0151     0.0576    0.26  0.79278
## Year1997            -0.1005     0.2822   -0.36  0.72178
## Year1998            -0.1885     0.2520   -0.75  0.45479
## Year1999            -0.4742     0.2023   -2.34  0.01943 *
## Year2000             0.5842     0.2112    2.77  0.00586 **
## Year2001             0.7413     0.2113    3.51  0.00049 ***
## Year2002             0.4267     0.2324    1.84  0.06690 .
## Year2003             0.4726     0.2311    2.05  0.04132 *
## Year2004             0.5205     0.2448    2.13  0.03390 *
## Year2005             0.6428     0.2349    2.74  0.00642 **
## Year2006             0.5744     0.2309    2.49  0.01315 *
```

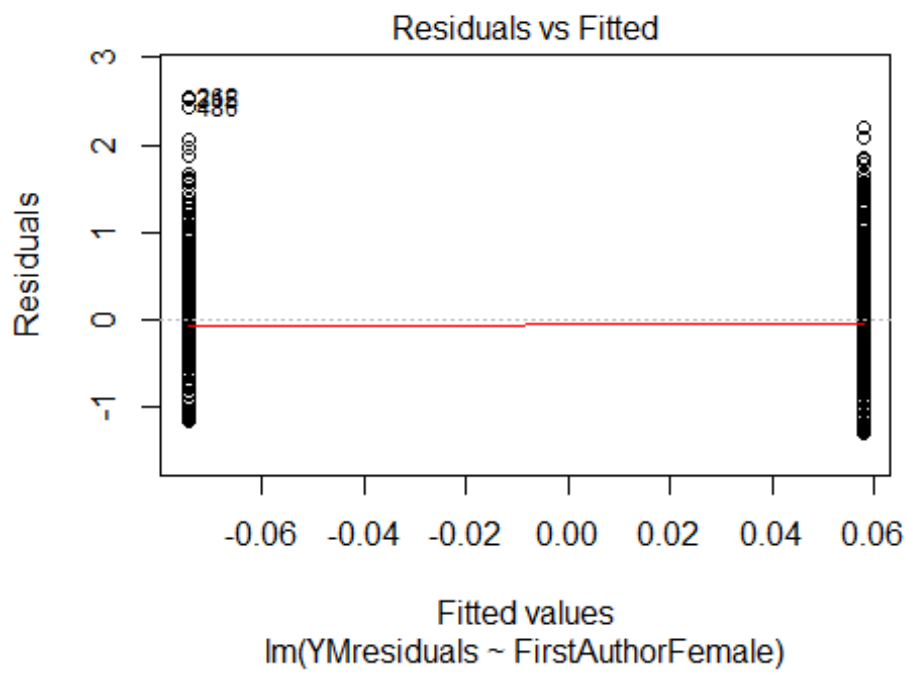
```
## Year2007            0.6616     0.2221    2.98  0.00303 **
## Year2008            0.6086     0.2134    2.85  0.00452 **
## Year2009            0.5579     0.2152    2.59  0.00977 **
## Year2010            0.5595     0.2104    2.66  0.00806 **
## Year2011            0.4198     0.2112    1.99  0.04734 *
## Year2012            0.4660     0.2301    2.02  0.04337 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.565
## Multiple R-squared:  0.0812, Adjusted R-squared:  0.0527
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
##  observation 217 is an outlier with |weight| = 0 ( < 0.00018);
##  46 weights are ~= 1. The remaining 519 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.430   0.875   0.951   0.912   0.984   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.77e-04          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                    psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Sample size for the above analysis:  566"
## [1] ""
## [1] ""
## [1] "####################################"
## [1] "Analysis of AJSC 3603"
## [1] "####################################"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##   43   33   24   26   42   44   59   54   39   43   59   60   66   92   76
## 2011 2012
##   80   72
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##   38   27   21   21   33   29   47   51   34   38   55   53   60   78   62
```

```
## 2011 2012
##   64   63
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##   36   26   21   21   33   29   40   47   31   36   52   52   59   74   59
## 2011 2012
##   62   62
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
##  Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 31, df = 16, p-value = 0.01
```

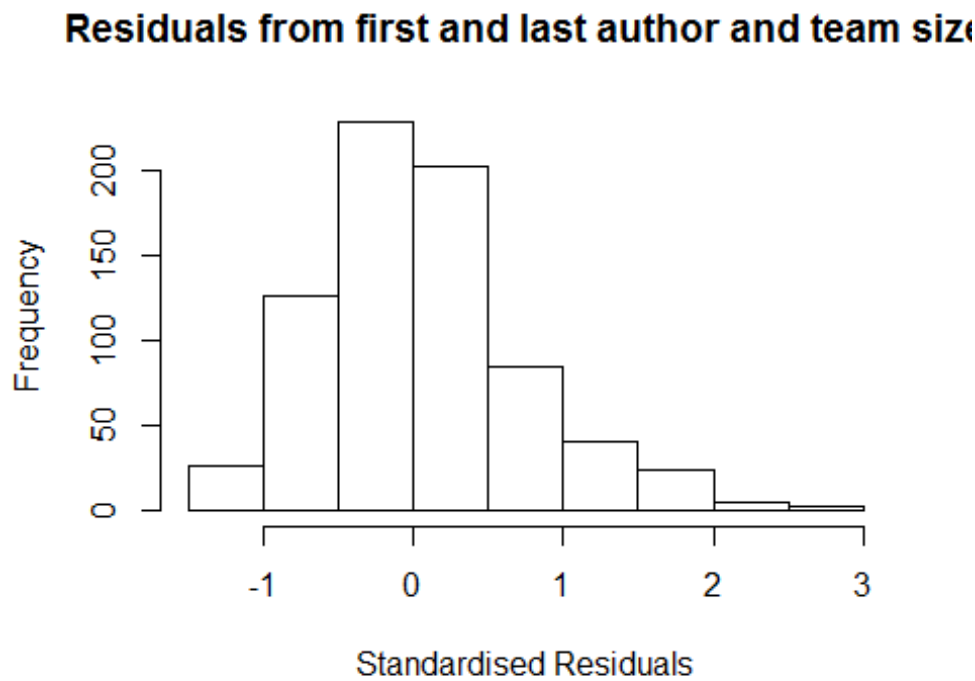



```
##
##  Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 0.084, df = 1, p-value = 0.8

## [1] "Female first author team size 2018 geometric mean: 2.9201363581422"
## [1] "Male first author team size 2018 geometric mean: 2.76996691275158"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
```

```
## W = 740, p-value = 0.8
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 2.8287292665956"
## [1] "Male last author team size 2018 geometric mean: 2.92060856494819"

## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties
```

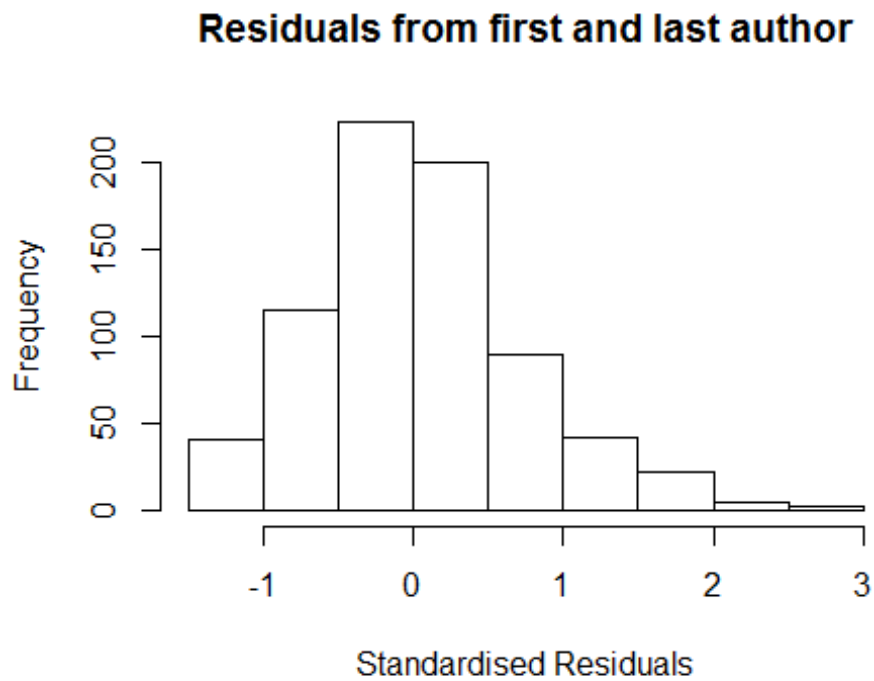



```
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 780, p-value = 0.9
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.459  1           1.208
## LastAuthorFemale  1.525  1           1.235
## UniqueAuthors     1.716  4           1.070
## Year              1.993 16           1.022
```



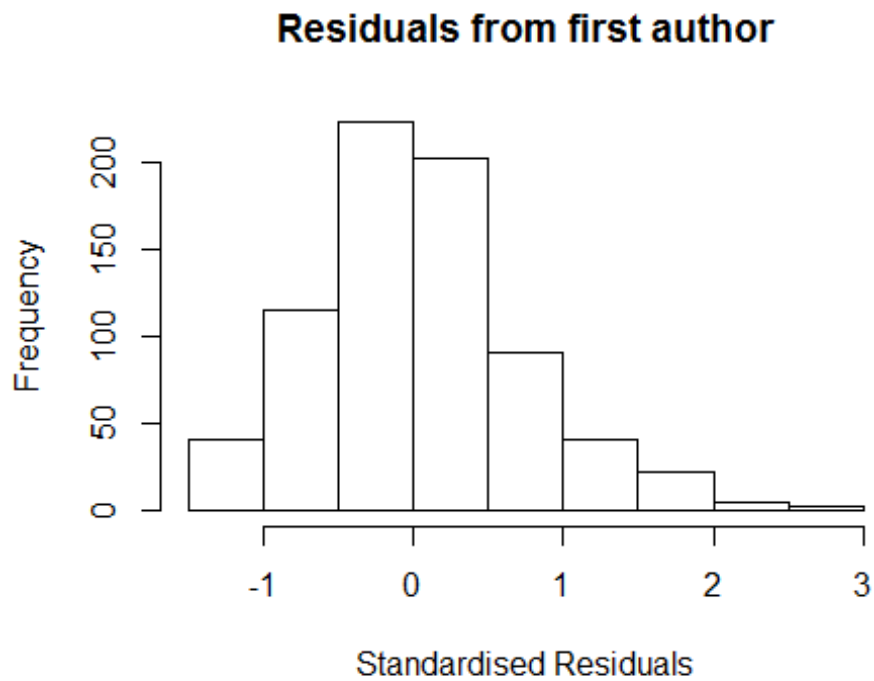

```
## [1] "List of  2 outliers with residuals above  2.5"
##       ScopusId  NLCS Year OneField Fields residuals
## 262 0036616956 3.693 2002     1210      2     2.783
## 315 0141815500 3.632 2003     1210      2     2.818
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
UniqueAuthors +
##     Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##     k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.3493 -0.3859 -0.0158  0.4174  2.8184
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         0.81001    0.14872    5.45  7.1e-08 ***
## FirstAuthorFemale1  0.15332    0.05627    2.72  0.00660 **
## LastAuthorFemale1   0.01288    0.05658    0.23  0.82004
## UniqueAuthors2      0.15572    0.06143    2.53  0.01146 *
## UniqueAuthors3      0.24727    0.07471    3.31  0.00098 ***
## UniqueAuthors4      0.25519    0.08378    3.05  0.00240 **
## UniqueAuthors5      0.59544    0.09171    6.49  1.6e-10 ***
## Year1997           -0.07253    0.20845   -0.35  0.72798
## Year1998           -0.08263    0.18337   -0.45  0.65239
```

```
## Year1999           -0.19049    0.19902   -0.96  0.33881
## Year2000           -0.05536    0.19368   -0.29  0.77507
## Year2001            0.05543    0.19204    0.29  0.77296
## Year2002            0.09964    0.18013    0.55  0.58035
## Year2003            0.00361    0.17961    0.02  0.98399
## Year2004            0.04088    0.18264    0.22  0.82295
## Year2005            0.05977    0.16440    0.36  0.71628
## Year2006           -0.07632    0.17118   -0.45  0.65585
## Year2007            0.06583    0.17618    0.37  0.70877
## Year2008            0.18876    0.16875    1.12  0.26370
## Year2009           -0.01493    0.16160   -0.09  0.92640
## Year2010            0.05721    0.16982    0.34  0.73631
## Year2011            0.11801    0.16362    0.72  0.47099
## Year2012           -0.20949    0.16434   -1.27  0.20281
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.624
## Multiple R-squared:  0.101,  Adjusted R-squared:  0.0737
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
##  74 weights are ~= 1. The remaining 666 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0051  0.8610  0.9540  0.8920  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.35e-04          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                    psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##                     GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.405  1           1.185
## LastAuthorFemale  1.448  1           1.204
## Year              1.192 16           1.006
```




```
## [1] "List of  2 outliers with residuals above  2.5"
##       ScopusId  NLCS Year OneField Fields residuals
## 262 0036616956 3.693 2002     1210      2     2.673
## 315 0141815500 3.632 2003     1210      2     2.646
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.2958 -0.4047 -0.0178  0.4278  2.6734
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)          0.94412    0.14515    6.50  1.5e-10 ***
## FirstAuthorFemale1   0.14687    0.05862    2.51    0.012 *
## LastAuthorFemale1    0.00497    0.05862    0.08    0.932
## Year1997            -0.11169    0.21363   -0.52    0.601
## Year1998            -0.11562    0.18789   -0.62    0.539
## Year1999            -0.30154    0.19892   -1.52    0.130
## Year2000            -0.08132    0.19984   -0.41    0.684
## Year2001             0.06487    0.20132    0.32    0.747
## Year2002             0.07548    0.18386    0.41    0.682
## Year2003             0.04180    0.18120    0.23    0.818
## Year2004             0.01978    0.18337    0.11    0.914
```

```
## Year2005            0.11125    0.16734    0.66    0.506
## Year2006           -0.03097    0.17952   -0.17    0.863
## Year2007            0.07511    0.18165    0.41    0.679
## Year2008            0.19986    0.17376    1.15    0.250
## Year2009            0.00184    0.16655    0.01    0.991
## Year2010            0.11486    0.17177    0.67    0.504
## Year2011            0.13477    0.16422    0.82    0.412
## Year2012           -0.12283    0.17083   -0.72    0.472
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.658
## Multiple R-squared:  0.0384, Adjusted R-squared:  0.0143
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
##  67 weights are ~= 1. The remaining 673 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0613  0.8580  0.9550  0.8980  0.9860  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.35e-04          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.079  1           1.039
## Year              1.079 16           1.002
```



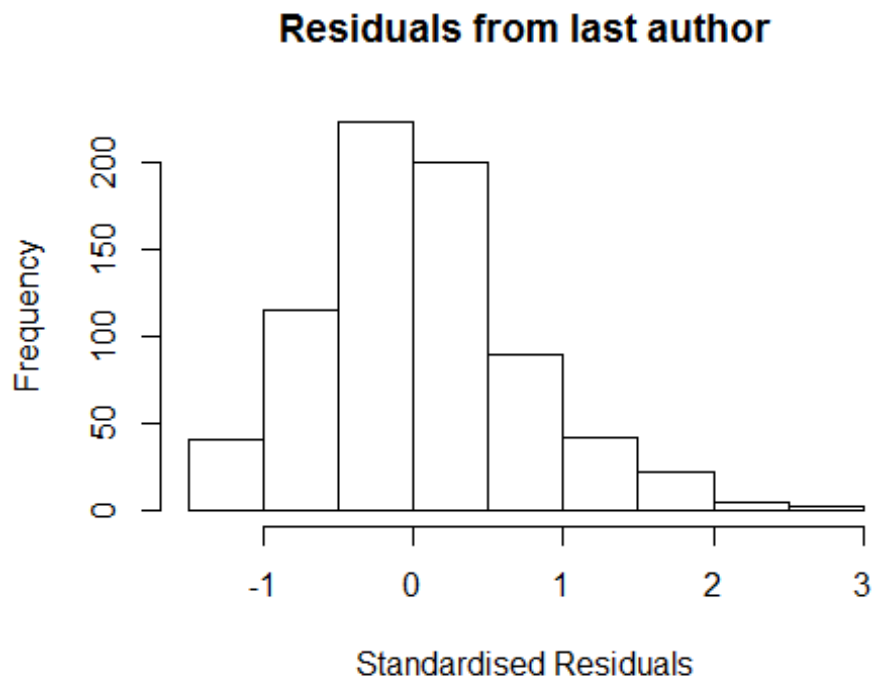

```
## [1] "List of  2 outliers with residuals above  2.5"
##       ScopusId  NLCS Year OneField Fields residuals
## 262 0036616956 3.693 2002     1210      2     2.673
## 315 0141815500 3.632 2003     1210      2     2.646
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.2945 -0.4052 -0.0165  0.4268  2.6727
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)           0.9445     0.1450    6.51  1.4e-10 ***
## FirstAuthorFemale1    0.1497     0.0516    2.90   0.0038 **
## Year1997             -0.1113     0.2135   -0.52   0.6022
## Year1998             -0.1147     0.1877   -0.61   0.5414
## Year1999             -0.3012     0.1988   -1.52   0.1301
## Year2000             -0.0808     0.1992   -0.41   0.6852
## Year2001              0.0656     0.2011    0.33   0.7444
## Year2002              0.0757     0.1836    0.41   0.6801
## Year2003              0.0423     0.1807    0.23   0.8148
## Year2004              0.0200     0.1832    0.11   0.9132
## Year2005              0.1117     0.1670    0.67   0.5038
```

```
## Year2006            -0.0301     0.1785   -0.17   0.8659
## Year2007             0.0750     0.1815    0.41   0.6794
## Year2008             0.2004     0.1735    1.15   0.2487
## Year2009             0.0026     0.1661    0.02   0.9875
## Year2010             0.1158     0.1711    0.68   0.4987
## Year2011             0.1358     0.1632    0.83   0.4054
## Year2012            -0.1227     0.1709   -0.72   0.4732
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.657
## Multiple R-squared:  0.0384, Adjusted R-squared:  0.0157
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
##  67 weights are ~= 1. The remaining 673 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0601  0.8570  0.9550  0.8980  0.9860  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.35e-04          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##                   GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.101  1           1.049
## Year             1.101 16           1.003

## [1] "List of  2 outliers with residuals above  2.5"
##       ScopusId  NLCS Year OneField Fields residuals
## 262 0036616956 3.693 2002     1210      2     2.673
## 315 0141815500 3.632 2003     1210      2     2.646
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
```

```
##     Min      1Q  Median      3Q     Max
## -1.2636 -0.4211 -0.0191  0.4384  2.6329
##
## Coefficients:
##                   Estimate Std. Error t value Pr(>|t|)
## (Intercept)        0.98004    0.14629    6.70  4.2e-11 ***
## LastAuthorFemale1  0.08693    0.05176    1.68    0.094 .
## Year1997          -0.10404    0.21469   -0.48    0.628
## Year1998          -0.10008    0.18334   -0.55    0.585
## Year1999          -0.30116    0.20002   -1.51    0.133
## Year2000          -0.09417    0.19974   -0.47    0.637
## Year2001           0.08652    0.19966    0.43    0.665
## Year2002           0.08010    0.18603    0.43    0.667
## Year2003           0.04469    0.18139    0.25    0.805
## Year2004           0.02904    0.18441    0.16    0.875
## Year2005           0.13941    0.16696    0.83    0.404
## Year2006          -0.03455    0.17990   -0.19    0.848
## Year2007           0.08104    0.18395    0.44    0.660
## Year2008           0.19661    0.17361    1.13    0.258
## Year2009          -0.00161    0.16750   -0.01    0.992
## Year2010           0.10772    0.17337    0.62    0.535
## Year2011           0.13087    0.16446    0.80    0.426
## Year2012          -0.12435    0.17265   -0.72    0.472
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.659
## Multiple R-squared:  0.0307, Adjusted R-squared:  0.00791
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
##  64 weights are ~= 1. The remaining 676 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0747  0.8590  0.9520  0.8980  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.35e-04          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                    psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
```

```
## [1] "Sample size for the above analysis:  740"
## [1] ""
## [1] ""
## [1] "####################################"
## [1] "Analysis of AJSC 3604"
## [1] "####################################"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011 2012
##    2   30   25   33   17   25   21   35   13   21   26    5    2   16   23
##
## 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011 2012
##    2   26   23   33   13   17   19   30    9   14   19    4    1    0    0
##
## 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011 2012
##    2   24   23   33   11   16   18   29    9   12   19    4    1    0    0
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
## [1] "Female first author team size 2018 geometric mean: 4.58257569495584"
## [1] "Male first author team size 2018 geometric mean: 2.75892417638112"

## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties

##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 4, p-value = 0.8
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 4.58257569495584"
## [1] "Male last author team size 2018 geometric mean: 1"

## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties

##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 4, p-value = 0.3
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"

## Warning in lmrob.S(x, y, control = control, mf = mf): S-estimated scale ==
## 0: Probably exact fit; check your data
```

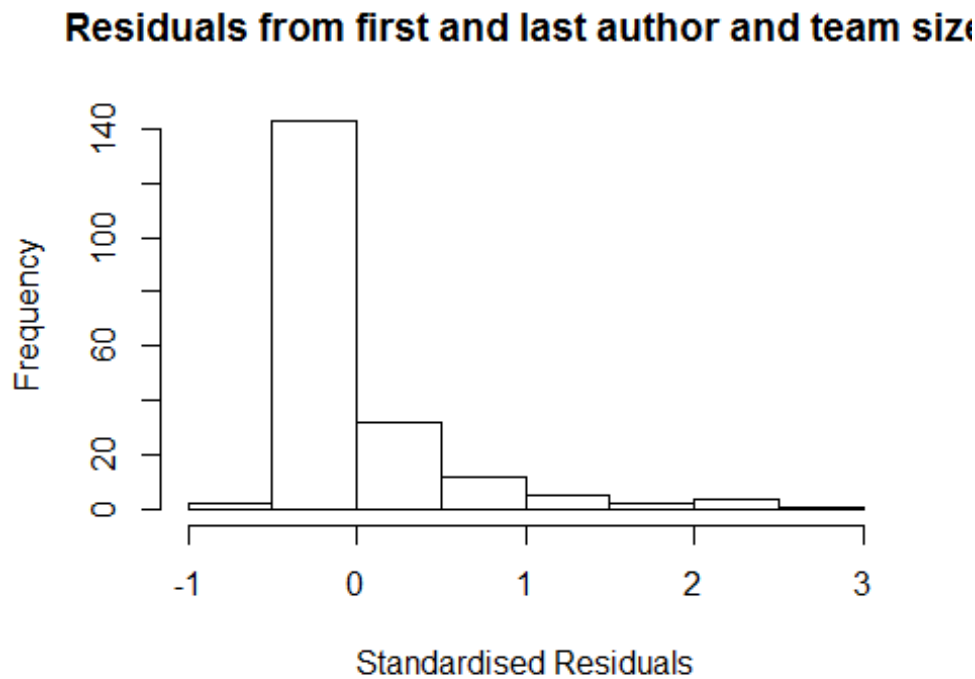
```
## Warning in lmrob.S(x, y, control = control, mf = mf): diag(.) had 0 or NA
## entries; non-finite result is doubtful
```




```
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale  NaN  1             NaN
## LastAuthorFemale   NaN  1             NaN
## UniqueAuthors      NaN  4             NaN
## Year               NaN 12             NaN

## [1] "List of  1 outliers with residuals above  2.5"
##       ScopusId  NLCS Year OneField Fields residuals
## 55 78951485576 2.085 2000     3311      3     2.577
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
UniqueAuthors +
##     Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##     k.max = 1000))
##  \--> method = "S"
## Residuals:
##    Min     1Q Median     3Q    Max
## -0.984  0.000  0.000  0.093  2.577
##
## Exact fit detected
##
## Coefficients:
```

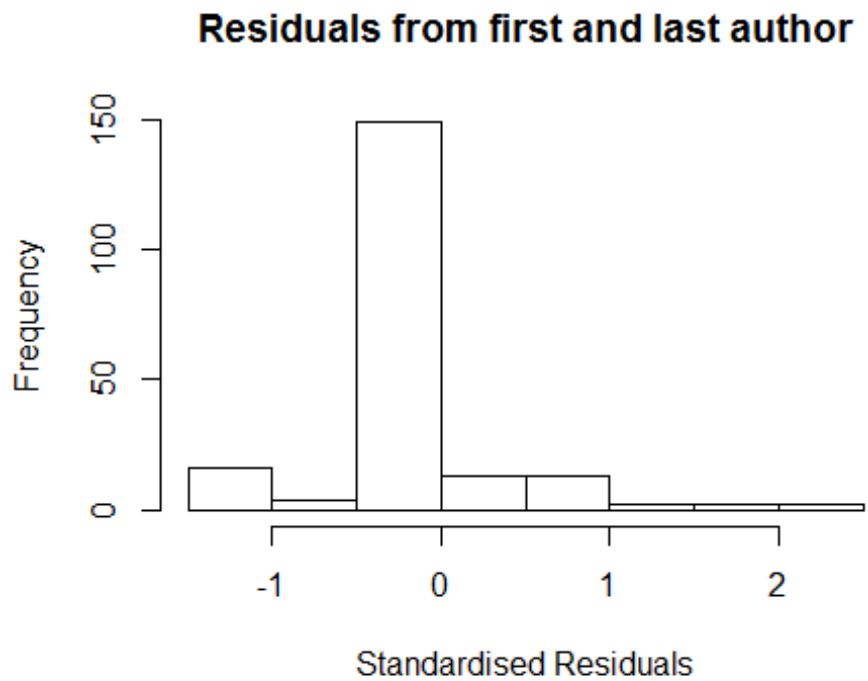
```
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)            0.000      0.000      NA       NA
## FirstAuthorFemale1    -0.492      0.000      NA       NA
## LastAuthorFemale1      0.492      0.000      NA       NA
## UniqueAuthors2         0.000      0.000      NA       NA
## UniqueAuthors3         0.000      0.000      NA       NA
## UniqueAuthors4         0.492      0.000      NA       NA
## UniqueAuthors5         0.000      0.000      NA       NA
## Year1999               0.000      0.000      NA       NA
## Year2000               0.000      0.000      NA       NA
## Year2001               0.000      0.000      NA       NA
## Year2002               0.492      0.000      NA       NA
## Year2003               0.000      0.000      NA       NA
## Year2004               0.000      0.000      NA       NA
## Year2005               0.000      0.000      NA       NA
## Year2006               0.000      0.000      NA       NA
## Year2007              -0.093      0.000      NA       NA
## Year2008               0.000      0.000      NA       NA
## Year2009               0.429      0.000      NA       NA
## Year2010               1.034      0.000      NA       NA
##
## Robustness weights:
##  76 observations
c(1,3,5,6,13,14,16,17,21,23,24,25,26,27,30,33,35,36,40,45,46,47,48,50,52,53,5
5,56,61,64,65,72,74,77,81,82,83,85,86,87,88,89,91,92,93,108,122,125,132,155,1
56,159,162,163,166,167,169,170,171,172,173,174,175,176,177,178,179,184,187,18
8,190,193,196,197,198,200)
##   are outliers with |weight| = 0 ( < 0.0005);
##  125 weights are ~= 1.
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          4.98e-04          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"

## Warning in lmrob.S(x, y, control = control, mf = mf): S-estimated scale ==
## 0: Probably exact fit; check your data
```

```
## Warning in lmrob.S(x, y, control = control, mf = mf): diag(.) had 0 or NA
## entries; non-finite result is doubtful
```



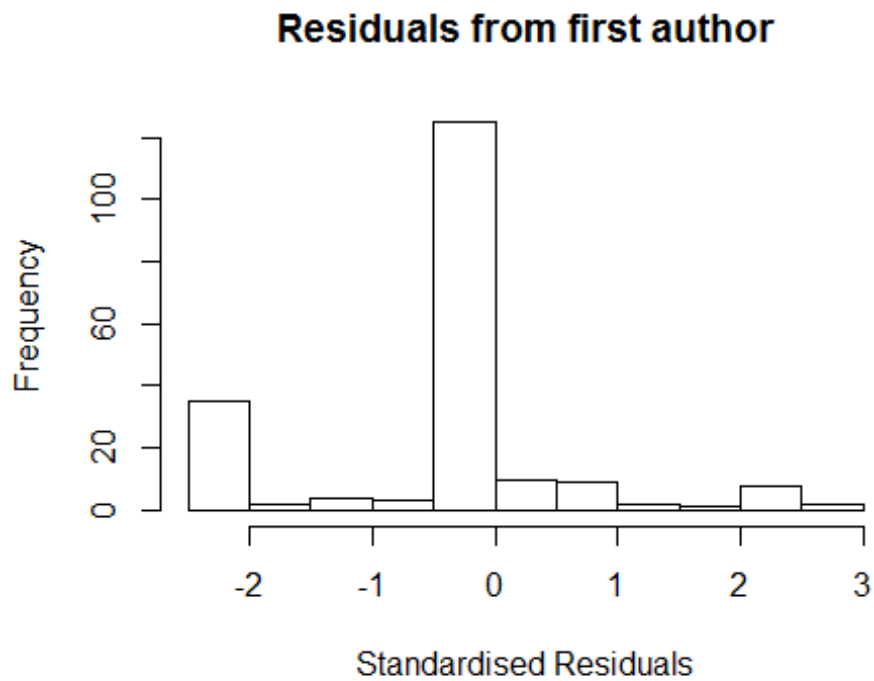

```
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale   NaN  1             NaN
## LastAuthorFemale    NaN  1             NaN
## Year                NaN 12             NaN

## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "S"
## Residuals:
##    Min     1Q Median     3Q    Max
##  -1.18   0.00   0.00   0.00   2.12
##
## Exact fit detected
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)             0.00       0.00      NA       NA
```

```
## FirstAuthorFemale1     0.00       0.00      NA       NA
## LastAuthorFemale1      0.00       0.00      NA       NA
## Year1999               0.00       0.00      NA       NA
## Year2000               1.10       0.00      NA       NA
## Year2001               0.00       0.00      NA       NA
## Year2002               0.00       0.00      NA       NA
## Year2003               0.00       0.00      NA       NA
## Year2004               0.00       0.00      NA       NA
## Year2005               0.00       0.00      NA       NA
## Year2006               0.00       0.00      NA       NA
## Year2007               0.00       0.00      NA       NA
## Year2008               0.00       0.00      NA       NA
## Year2009               1.18       0.00      NA       NA
## Year2010               1.03       0.00      NA       NA
##
## Robustness weights:
##  54 observations
c(1,3,5,6,14,23,24,25,26,27,28,29,30,31,32,33,34,35,36,37,38,39,40,41,42,43,4
4,45,47,48,49,50,61,72,74,77,81,82,83,85,145,155,156,159,162,168,174,178,187,
188,196,198,199,200)
##   are outliers with |weight| = 0 ( < 0.0005);
##  147 weights are ~= 1.
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          4.98e-04          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"

## Warning in lmrob.S(x, y, control = control, mf = mf): S-estimated scale ==
## 0: Probably exact fit; check your data

## Warning in lmrob.S(x, y, control = control, mf = mf): diag(.) had 0 or NA
## entries; non-finite result is doubtful
```

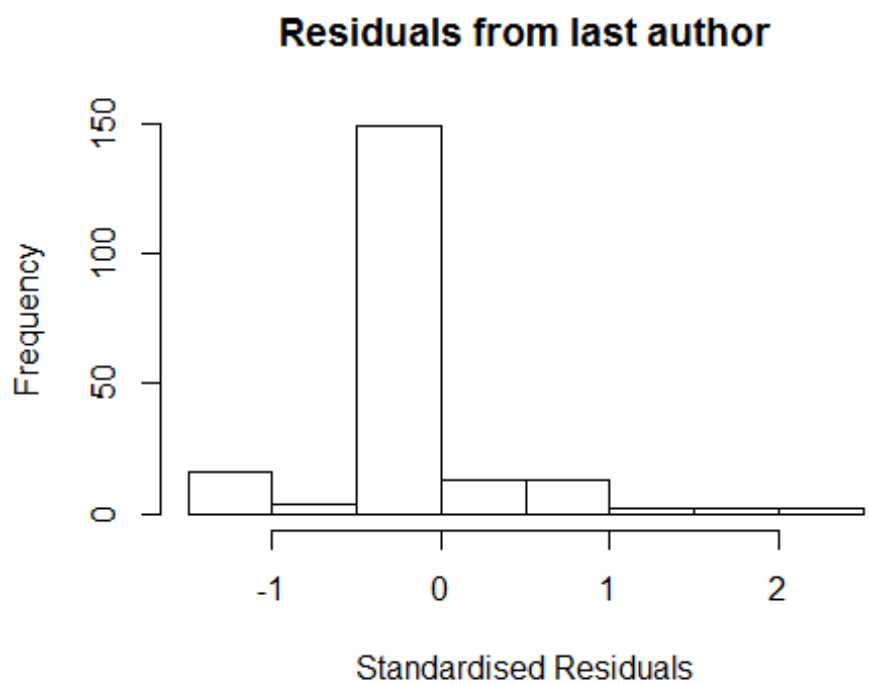



```
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale   NaN  1             NaN
## Year                NaN 12             NaN

## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "S"
## Residuals:
##    Min     1Q Median     3Q    Max
## -2.484 -0.399  0.000  0.000  2.952
##
## Exact fit detected
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)             2.12       0.00      NA       NA
## FirstAuthorFemale1      2.08       0.00      NA       NA
## Year1999               -2.12       0.00      NA       NA
## Year2000               -2.12       0.00      NA       NA
## Year2001               -2.12       0.00      NA       NA
## Year2002               -4.21       0.00      NA       NA
```

```
## Year2003              -2.12       0.00      NA       NA
## Year2004              -2.12       0.00      NA       NA
## Year2005              -2.12       0.00      NA       NA
## Year2006              -2.12       0.00      NA       NA
## Year2007              -1.73       0.00      NA       NA
## Year2008              -2.12       0.00      NA       NA
## Year2009              -2.84       0.00      NA       NA
## Year2010              -3.18       0.00      NA       NA
##
## Robustness weights:
##  86 observations
c(2,3,4,5,6,9,11,14,16,17,18,21,23,24,25,26,27,30,33,34,35,36,39,40,41,44,45,
46,48,50,57,59,60,61,63,65,66,72,74,77,78,79,81,82,83,85,86,87,88,91,92,93,10
8,132,137,140,145,147,155,156,158,159,160,162,166,167,169,170,171,172,173,175
,176,177,178,180,184,185,187,188,190,193,196,197,199,200)
##   are outliers with |weight| = 0 ( < 0.0005);
##  115 weights are ~= 1.
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          4.98e-04          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"

## Warning in lmrob.S(x, y, control = control, mf = mf): S-estimated scale ==
## 0: Probably exact fit; check your data

## Warning in lmrob.S(x, y, control = control, mf = mf): diag(.) had 0 or NA
## entries; non-finite result is doubtful
```

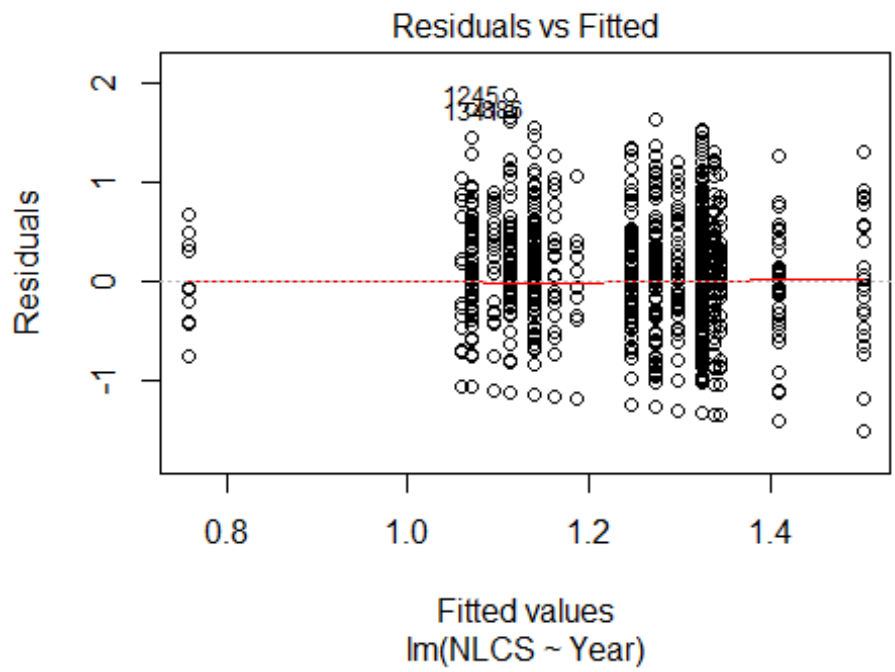



```
##                   GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale   NaN  1             NaN
## Year               NaN 12             NaN
```

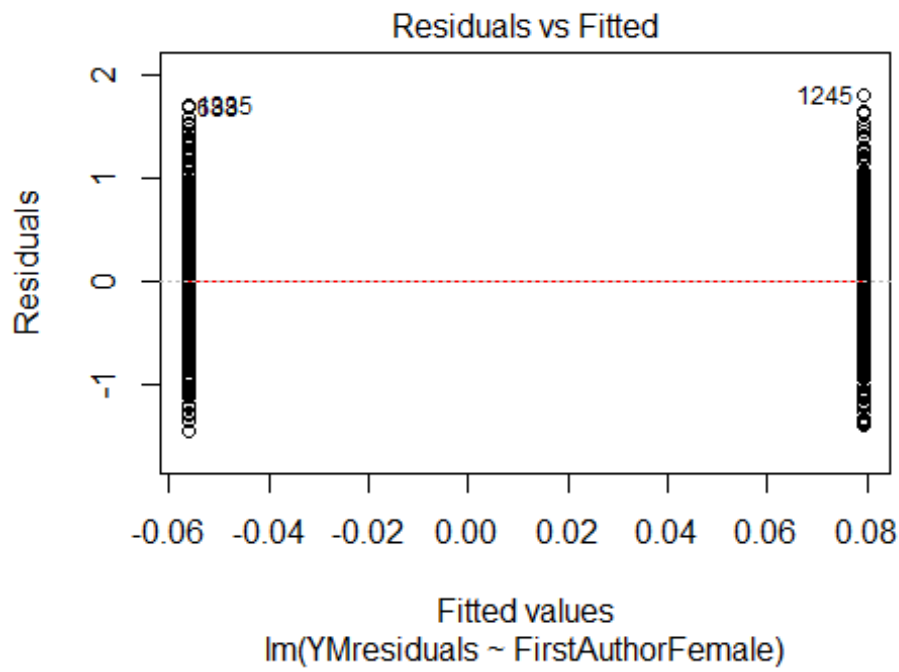


```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "S"
## Residuals:
##    Min     1Q Median     3Q    Max
## -1.651  0.000  0.000  0.486  3.222
##
## Exact fit detected
##
## Coefficients:
##                   Estimate Std. Error t value Pr(>|t|)
## (Intercept)           0.00       0.00      NA       NA
## LastAuthorFemale1    -1.65       0.00      NA       NA
## Year1999              0.00       0.00      NA       NA
## Year2000              0.00       0.00      NA       NA
## Year2001              0.00       0.00      NA       NA
## Year2002              1.65       0.00      NA       NA
## Year2003              0.00       0.00      NA       NA
## Year2004              0.00       0.00      NA       NA
## Year2005              0.00       0.00      NA       NA
## Year2006              0.00       0.00      NA       NA
## Year2007              0.00       0.00      NA       NA
## Year2008              1.65       0.00      NA       NA
## Year2009              1.18       0.00      NA       NA
## Year2010              2.68       0.00      NA       NA
##
## Robustness weights:
##  84 observations
c(1,3,4,5,6,11,14,18,23,24,25,26,27,30,34,35,36,39,41,44,45,46,47,48,50,52,53
,55,56,57,59,60,61,63,64,66,72,74,77,78,79,81,82,83,84,85,86,87,88,90,92,93,1
25,137,140,145,155,156,158,159,160,162,163,168,172,174,178,181,182,183,184,18
6,187,188,189,190,191,192,193,194,195,198,199,200)
##   are outliers with |weight| = 0 ( < 0.0005);
##  117 weights are ~= 1.
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          4.98e-04          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
```

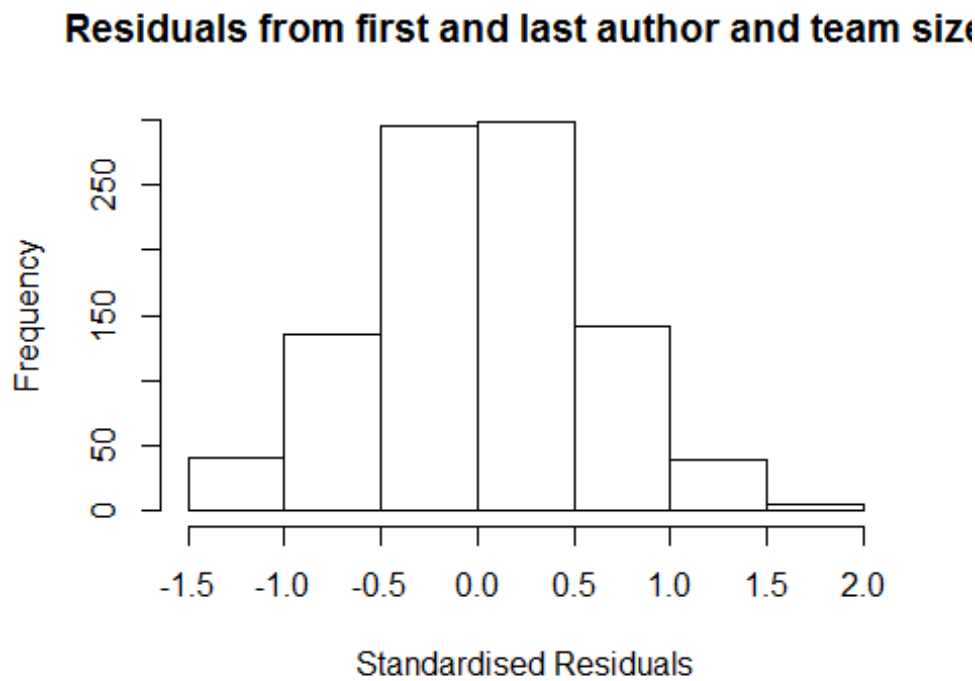
```
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Sample size for the above analysis:  201"
## [1] ""
## [1] ""
## [1] "##################################"
## [1] "Analysis of AJSC 3605"
## [1] "##################################"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##   53   44   33   35   43   44   56   56  114   84  141  147  151   87  131
## 2011 2012
##  220  247
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##   28   23   14   25   10   13   45   48   91   53  100   99  120   66   92
## 2011 2012
##  162  198
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##   24   20   10   21   10   11   40   39   64   33   79   92  101   58   69
## 2011 2012
##  126  159
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
##  Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 22, df = 16, p-value = 0.1
```




```
##
##  Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 0.0036, df = 1, p-value = 1
```

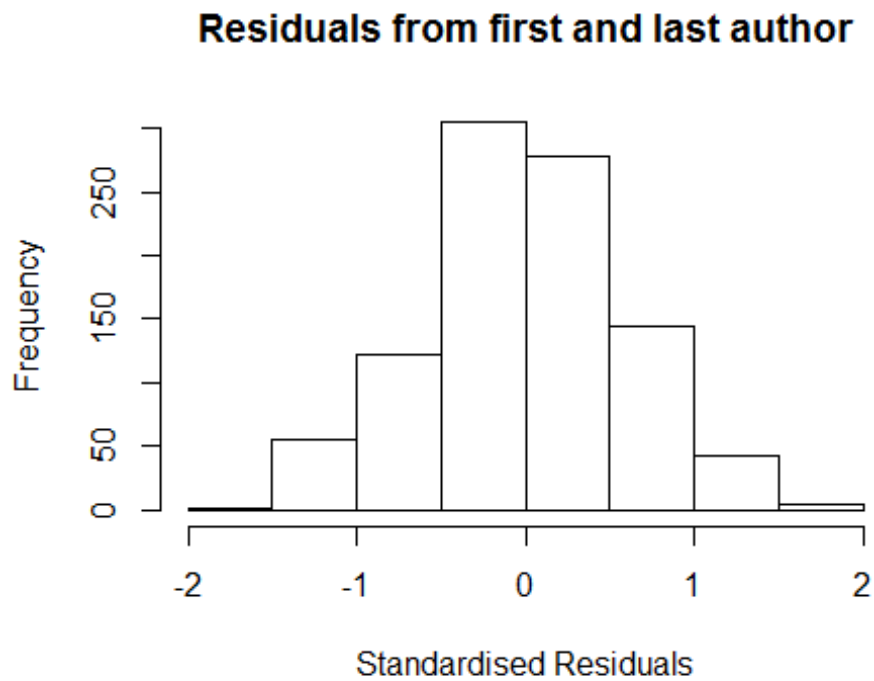



```
## [1] "Female first author team size 2018 geometric mean: 5.65938493572641"
## [1] "Male first author team size 2018 geometric mean: 4.63130642893535"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 10000, p-value = 0.03
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 5.21582150267972"
## [1] "Male last author team size 2018 geometric mean: 5.04256632893638"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 8700, p-value = 0.8
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.195  1           1.093
## LastAuthorFemale  1.105  1           1.051
## UniqueAuthors     1.667  4           1.066
## Year              1.807 16           1.019
```



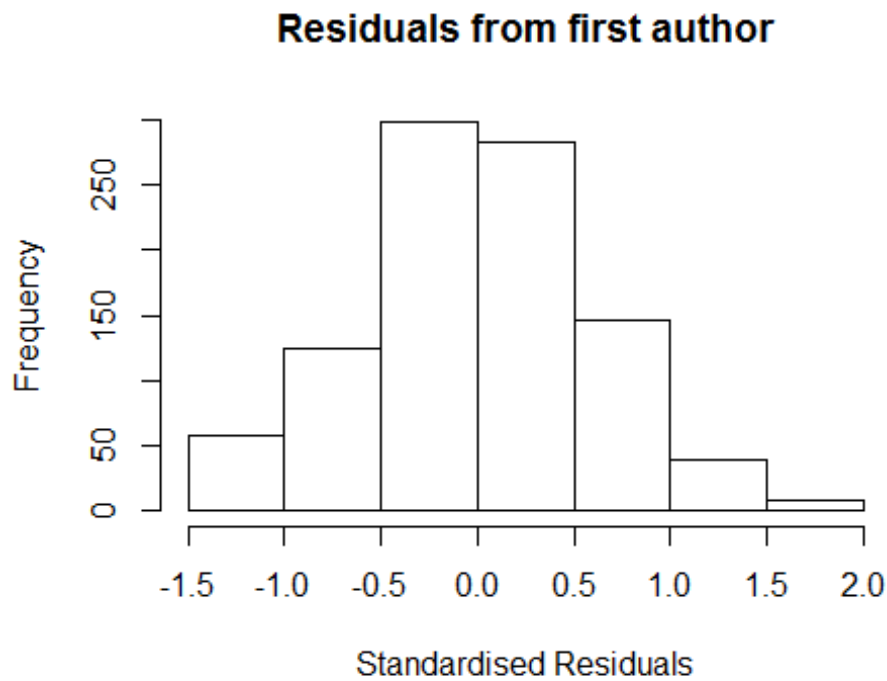

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
UniqueAuthors +
##     Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##     k.max = 1000))
##  \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.49888 -0.37723  0.00726  0.36333  1.90295
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         0.93315    0.14942    6.25  6.4e-10 ***
## FirstAuthorFemale1  0.07588    0.04109    1.85  0.06508 .
## LastAuthorFemale1   0.08755    0.04160    2.10  0.03562 *
## UniqueAuthors2      0.25981    0.07706    3.37  0.00078 ***
## UniqueAuthors3      0.28761    0.07760    3.71  0.00022 ***
## UniqueAuthors4      0.34004    0.07648    4.45  9.8e-06 ***
## UniqueAuthors5      0.52124    0.06918    7.53  1.2e-13 ***
## Year1997            0.26725    0.22239    1.20  0.22979
## Year1998           -0.25392    0.21301   -1.19  0.23355
## Year1999           -0.19927    0.23836   -0.84  0.40336
```

```
## Year2000           -0.02697    0.23884   -0.11  0.91013
## Year2001           -0.20846    0.26173   -0.80  0.42596
## Year2002            0.00956    0.17287    0.06  0.95591
## Year2003            0.07863    0.16216    0.48  0.62786
## Year2004           -0.16982    0.15709   -1.08  0.27997
## Year2005           -0.03138    0.19341   -0.16  0.87113
## Year2006            0.04040    0.16565    0.24  0.80735
## Year2007           -0.21198    0.15529   -1.37  0.17256
## Year2008           -0.18240    0.15422   -1.18  0.23720
## Year2009           -0.07398    0.17360   -0.43  0.67009
## Year2010           -0.18479    0.16533   -1.12  0.26398
## Year2011           -0.08815    0.15219   -0.58  0.56261
## Year2012           -0.05010    0.15340   -0.33  0.74403
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.559
## Multiple R-squared:  0.135,  Adjusted R-squared:  0.115
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
##  111 weights are ~= 1. The remaining 845 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.222   0.853   0.945   0.898   0.985   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.05e-04          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.131  1           1.063
## LastAuthorFemale  1.090  1           1.044
## Year              1.141 16           1.004
```



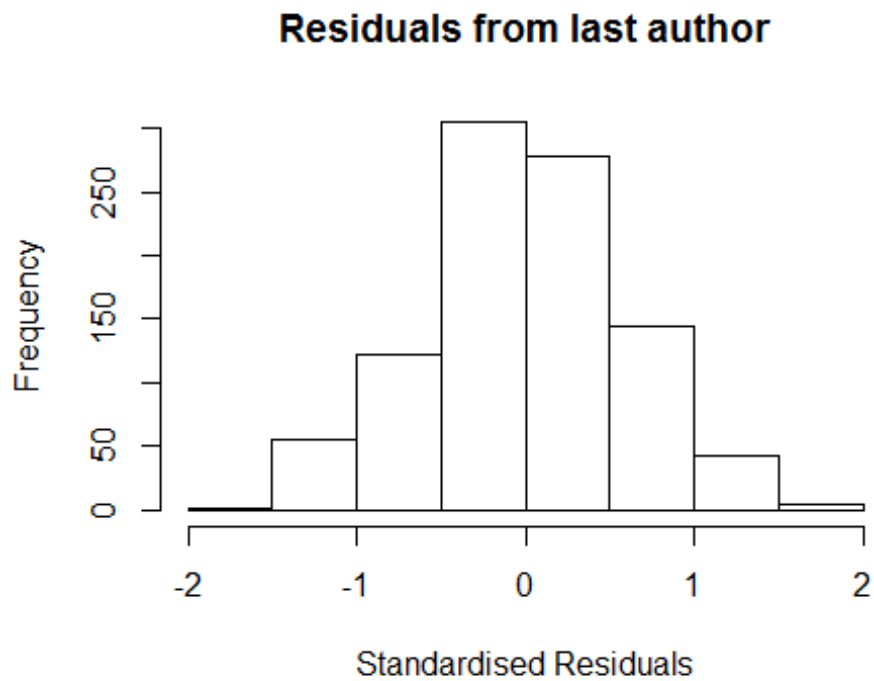

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.52730 -0.37811 -0.00751  0.38237  1.81536
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.11557    0.17565    6.35  3.3e-10 ***
## FirstAuthorFemale1  0.12885    0.04213    3.06   0.0023 **
## LastAuthorFemale1   0.07786    0.04355    1.79   0.0741 .
## Year1997            0.33074    0.25401    1.30   0.1932
## Year1998           -0.35375    0.23915   -1.48   0.1394
## Year1999           -0.08096    0.24514   -0.33   0.7413
## Year2000            0.04471    0.24447    0.18   0.8549
## Year2001           -0.08852    0.28941   -0.31   0.7598
## Year2002            0.25278    0.19347    1.31   0.1917
## Year2003            0.23561    0.18862    1.25   0.2119
## Year2004            0.00832    0.18447    0.05   0.9640
## Year2005            0.07515    0.21743    0.35   0.7297
```

```
## Year2006            0.20501    0.19198    1.07   0.2859
## Year2007           -0.13345    0.18504   -0.72   0.4710
## Year2008           -0.06295    0.18369   -0.34   0.7319
## Year2009            0.09604    0.19592    0.49   0.6241
## Year2010           -0.07279    0.18806   -0.39   0.6988
## Year2011            0.07948    0.18024    0.44   0.6593
## Year2012            0.17269    0.17893    0.97   0.3347
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.572
## Multiple R-squared:  0.0664, Adjusted R-squared:  0.0485
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
##  75 weights are ~= 1. The remaining 881 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.292   0.863   0.951   0.900   0.986   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.05e-04          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.102  1           1.050
## Year              1.102 16           1.003
```




```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.48836 -0.37186 -0.00424  0.38538  1.77691
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.13478    0.17297    6.56  8.8e-11 ***
## FirstAuthorFemale1  0.14936    0.04159    3.59  0.00035 ***
## Year1997            0.32504    0.25434    1.28  0.20159
## Year1998           -0.36064    0.23944   -1.51  0.13236
## Year1999           -0.09716    0.24427   -0.40  0.69092
## Year2000            0.02931    0.23976    0.12  0.90272
## Year2001           -0.07822    0.28886   -0.27  0.78661
## Year2002            0.25311    0.19191    1.32  0.18752
## Year2003            0.22704    0.18691    1.21  0.22478
## Year2004            0.00603    0.18271    0.03  0.97369
## Year2005            0.07651    0.21555    0.35  0.72270
## Year2006            0.20422    0.19056    1.07  0.28413
```

```
## Year2007           -0.13596    0.18333   -0.74  0.45849
## Year2008           -0.06675    0.18180   -0.37  0.71358
## Year2009            0.09516    0.19402    0.49  0.62393
## Year2010           -0.07405    0.18669   -0.40  0.69171
## Year2011            0.08428    0.17866    0.47  0.63725
## Year2012            0.17552    0.17736    0.99  0.32260
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.572
## Multiple R-squared:  0.0634, Adjusted R-squared:  0.0464
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
##  80 weights are ~= 1. The remaining 876 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.314   0.864   0.953   0.899   0.987   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.05e-04          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##                   GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.062  1           1.031
## Year             1.062 16           1.002
```

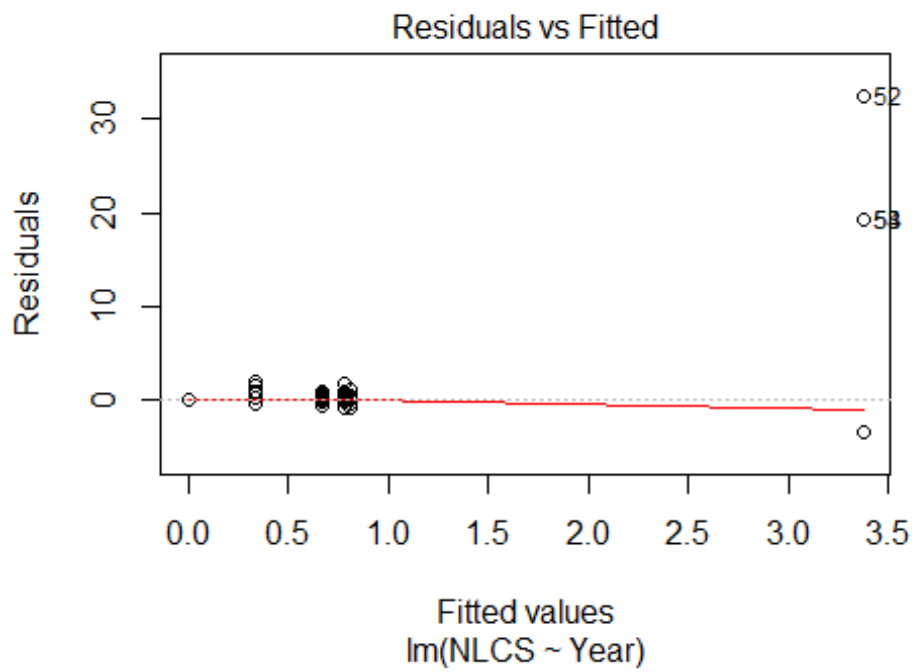



```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.46216 -0.39280 -0.00149  0.38824  1.90494
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)          1.1443     0.1799    6.36  3.1e-10 ***
## LastAuthorFemale1    0.1139     0.0433    2.63   0.0087 **
## Year1997             0.3164     0.2546    1.24   0.2142
## Year1998            -0.3615     0.2366   -1.53   0.1269
## Year1999            -0.0862     0.2490   -0.35   0.7294
## Year2000             0.0436     0.2536    0.17   0.8634
## Year2001            -0.1022     0.3043   -0.34   0.7371
## Year2002             0.2426     0.1978    1.23   0.2203
## Year2003             0.2564     0.1943    1.32   0.1872
## Year2004             0.0132     0.1892    0.07   0.9444
## Year2005             0.0865     0.2192    0.39   0.6934
## Year2006             0.2040     0.1967    1.04   0.3000
```

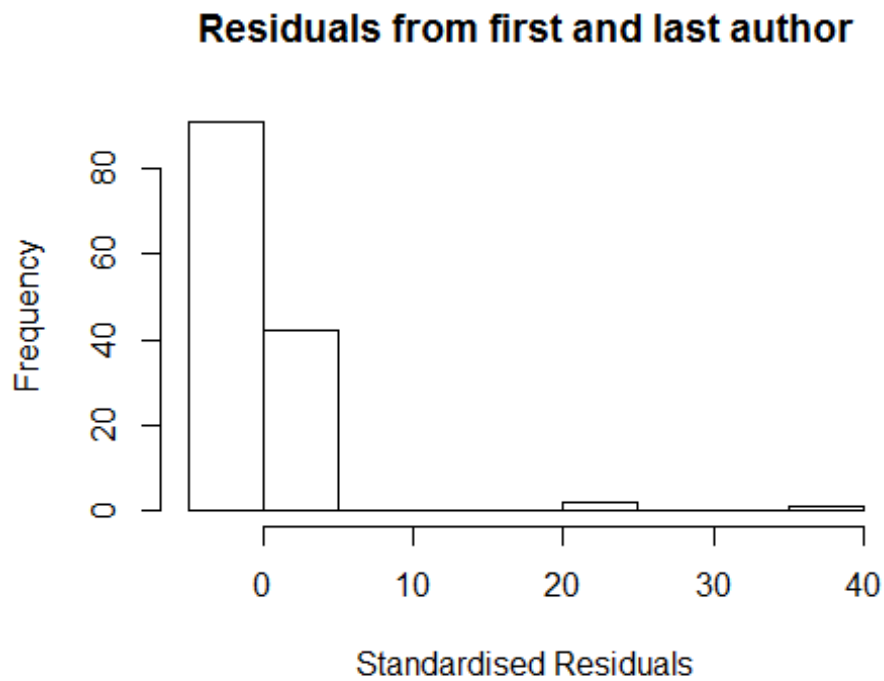
```
## Year2007           -0.1227     0.1901   -0.65   0.5186
## Year2008           -0.0496     0.1883   -0.26   0.7923
## Year2009            0.1195     0.2012    0.59   0.5526
## Year2010           -0.0622     0.1934   -0.32   0.7477
## Year2011            0.1064     0.1854    0.57   0.5660
## Year2012            0.1992     0.1838    1.08   0.2787
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.578
## Multiple R-squared:  0.0566, Adjusted R-squared:  0.0395
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
##  89 weights are ~= 1. The remaining 867 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.254   0.849   0.949   0.899   0.986   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.05e-04          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Sample size for the above analysis:  956"
## [1] ""
## [1] ""
## [1] "####################################"
## [1] "Analysis of AJSC 3606"
## [1] "####################################"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2004 2005 2006 2007 2008 2009 2010 2011 2012
##    2   20   20    9   25   13    3   10   19   36   27   26   23
##
## 1996 1997 1998 1999 2004 2005 2006 2007 2008 2009 2010 2011 2012
##    0   20   17    7   24   11    3   10   19   34   25   24   21
##
## 1996 1997 1998 1999 2004 2005 2006 2007 2008 2009 2010 2011 2012
##    0   20   17    7   24   11    3   10   19   32   23   23   19
```

```
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
##  Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = Inf, df = 6, p-value <2e-16

## [1] "Female first author team size 2018 geometric mean: 2.91295063024394"
## [1] "Male first author team size 2018 geometric mean: 2.71080601082953"

## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties

##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 8, p-value = 1
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 3.22370979547063"
## [1] "Male last author team size 2018 geometric mean: 2.44948974278318"

## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties

##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 10, p-value = 0.5
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"

## Warning in lmrob.S(x, y, control = control, mf = mf): find_scale() did not
## converge in 'maxit.scale' (= 200) iterations

## Warning in lmrob.S(x, y, control = control, mf = mf): find_scale() did not
## converge in 'maxit.scale' (= 200) iterations

## Warning in lmrob.S(x, y, control = control, mf = mf): find_scale() did not
## converge in 'maxit.scale' (= 200) iterations

## [1] "Regression 2: First author gender, Last author gender, Year as
factors"

## Warning in lmrob.S(x, y, control = control, mf = mf): find_scale() did not
## converge in 'maxit.scale' (= 200) iterations
```

```
## Warning in lmrob.S(x, y, control = control, mf = mf): S-estimated scale ==
## 0: Probably exact fit; check your data

## Warning in cov2cor(v): diag(.) had 0 or NA entries; non-finite result is
## doubtful
```

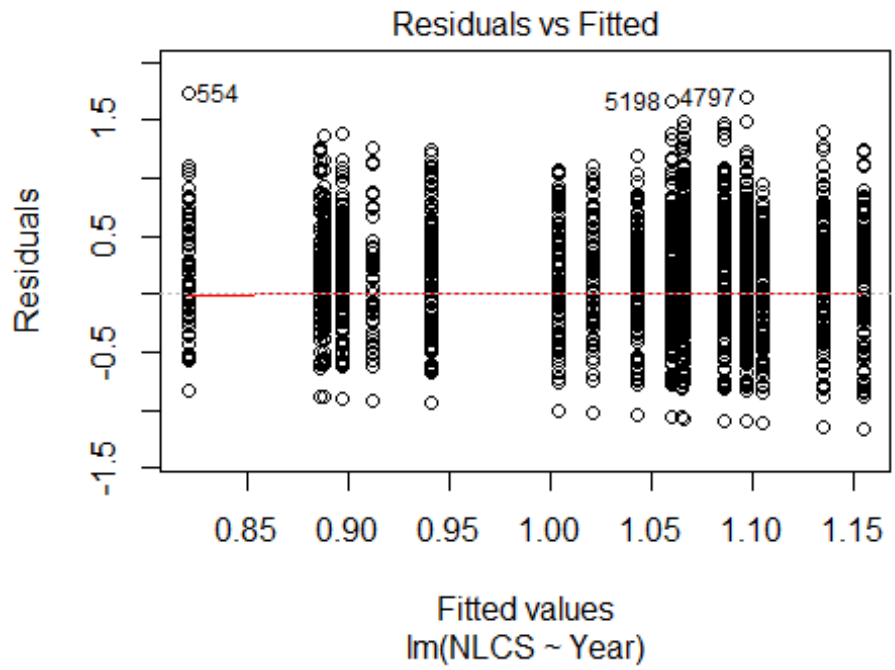



```
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale   NaN  1             NaN
## LastAuthorFemale    NaN  1             NaN
## Year                NaN  6             NaN
```

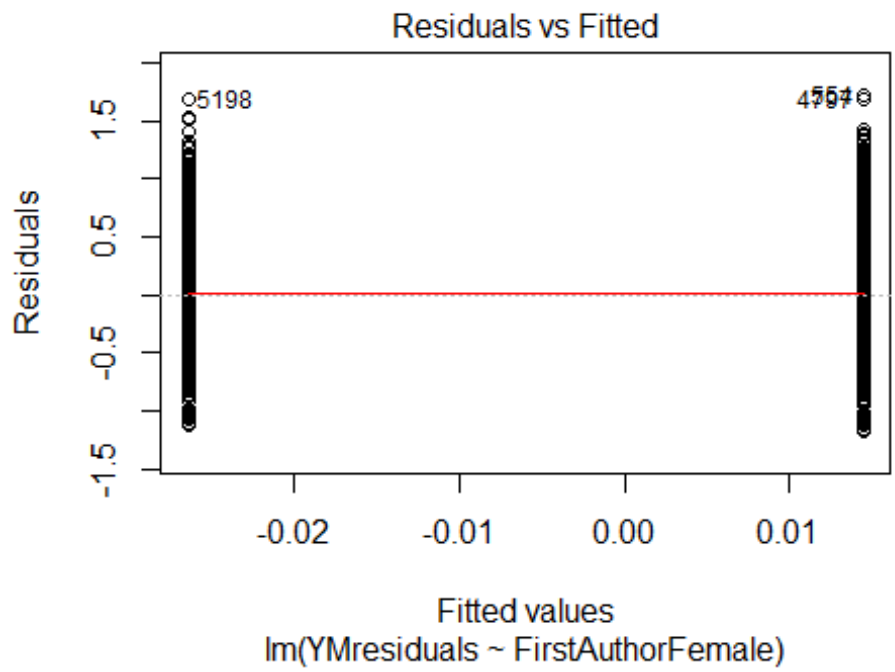



```
## [1] "Regression 3: First author gender, Year as factors"
## [1] "Regression 4: Last author gender, Year as factors"
## [1] "Sample size for the above analysis:  208"
## [1] ""
## [1] ""
## [1] "####################################"
## [1] "Analysis of AJSC 3607"
## [1] "####################################"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  231  253  230  260  423  353  350  316  357  291  272  235  227  317  407
## 2011 2012
##  354  326
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  171  184   97  105  107   59  129  229  261  219  211  186  173  246  313
## 2011 2012
##  255  241
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  149  163   83   94   90   50  118  197  217  175  186  164  146  206  263
## 2011 2012
##  225  201
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
```

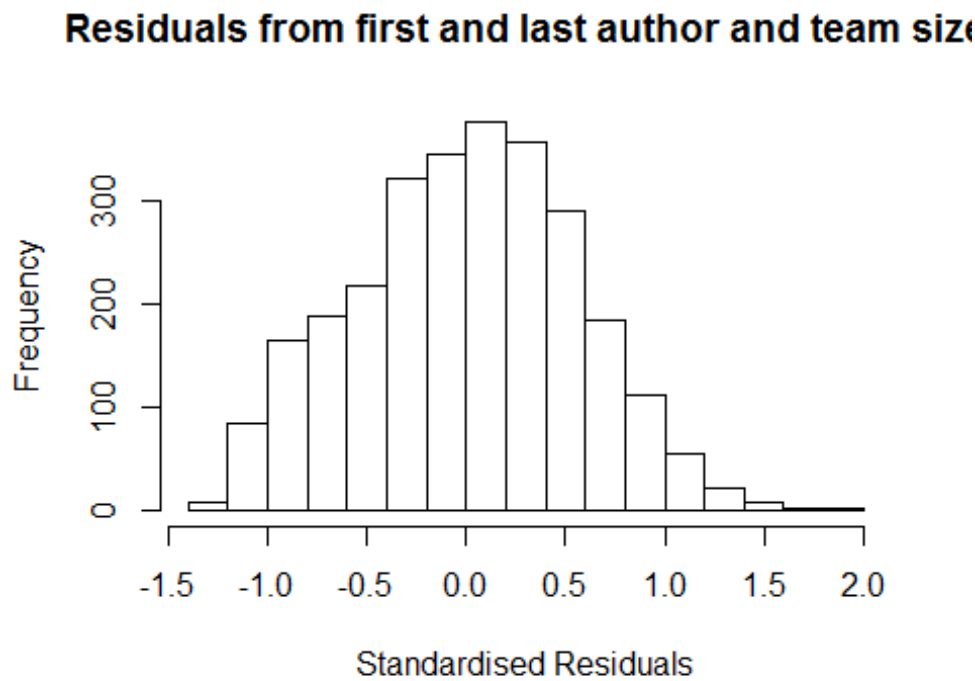
```
##
##  Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 33, df = 16, p-value = 0.006
```




```
##
##  Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 1.1, df = 1, p-value = 0.3
```

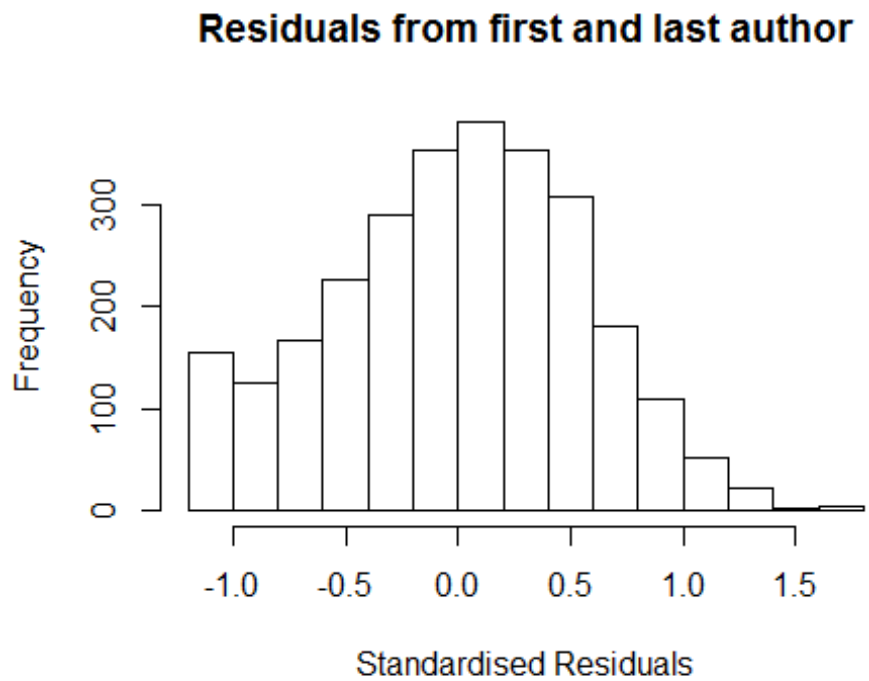



```
## [1] "Female first author team size 2018 geometric mean: 4.08982577232959"
## [1] "Male first author team size 2018 geometric mean: 3.5533622167172"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 6300, p-value = 0.2
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 4.1560559438621"
## [1] "Male last author team size 2018 geometric mean: 3.60190587094445"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 5800, p-value = 0.2
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.045  1           1.022
## LastAuthorFemale  1.026  1           1.013
## UniqueAuthors     1.182  4           1.021
## Year              1.213 16           1.006
```



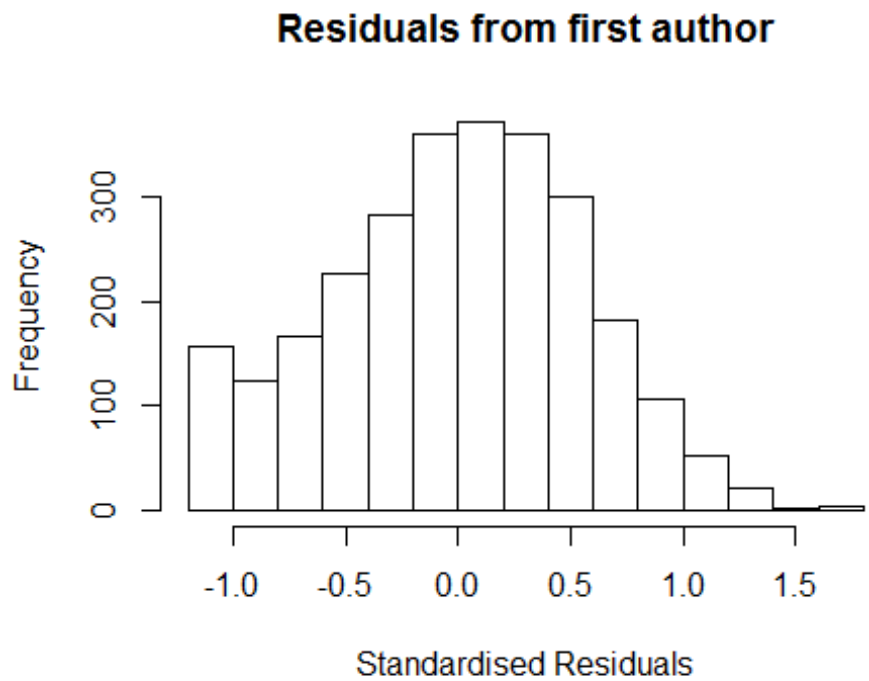

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
UniqueAuthors +
##     Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##     k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.2900 -0.3828  0.0241  0.3926  1.8624
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)          1.0151     0.0573   17.70  < 2e-16 ***
## FirstAuthorFemale1  -0.0510     0.0235   -2.17  0.03011 *
## LastAuthorFemale1   -0.0305     0.0260   -1.18  0.23934
## UniqueAuthors2       0.1310     0.0498    2.63  0.00852 **
## UniqueAuthors3       0.0968     0.0467    2.07  0.03830 *
## UniqueAuthors4       0.1642     0.0469    3.50  0.00047 ***
## UniqueAuthors5       0.3048     0.0446    6.83  1.1e-11 ***
## Year1997            -0.0910     0.0616   -1.48  0.13988
## Year1998            -0.3487     0.0821   -4.25  2.2e-05 ***
## Year1999            -0.2181     0.0867   -2.51  0.01199 *
```

```
## Year2000            -0.0639     0.0802   -0.80  0.42578
## Year2001            -0.2319     0.0905   -2.56  0.01044 *
## Year2002            -0.1735     0.0720   -2.41  0.01609 *
## Year2003            -0.1944     0.0596   -3.26  0.00111 **
## Year2004            -0.2711     0.0558   -4.86  1.3e-06 ***
## Year2005            -0.2983     0.0614   -4.86  1.2e-06 ***
## Year2006            -0.0900     0.0583   -1.54  0.12293
## Year2007            -0.0919     0.0600   -1.53  0.12545
## Year2008            -0.0298     0.0669   -0.45  0.65558
## Year2009            -0.0629     0.0631   -1.00  0.31873
## Year2010            -0.0894     0.0587   -1.52  0.12796
## Year2011            -0.1047     0.0606   -1.73  0.08435 .
## Year2012            -0.1173     0.0641   -1.83  0.06716 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.569
## Multiple R-squared:  0.0601, Adjusted R-squared:  0.0524
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
##  212 weights are ~= 1. The remaining 2515 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.262   0.868   0.950   0.912   0.985   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          3.67e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.030  1           1.015
## LastAuthorFemale  1.025  1           1.013
## Year              1.053 16           1.002
```



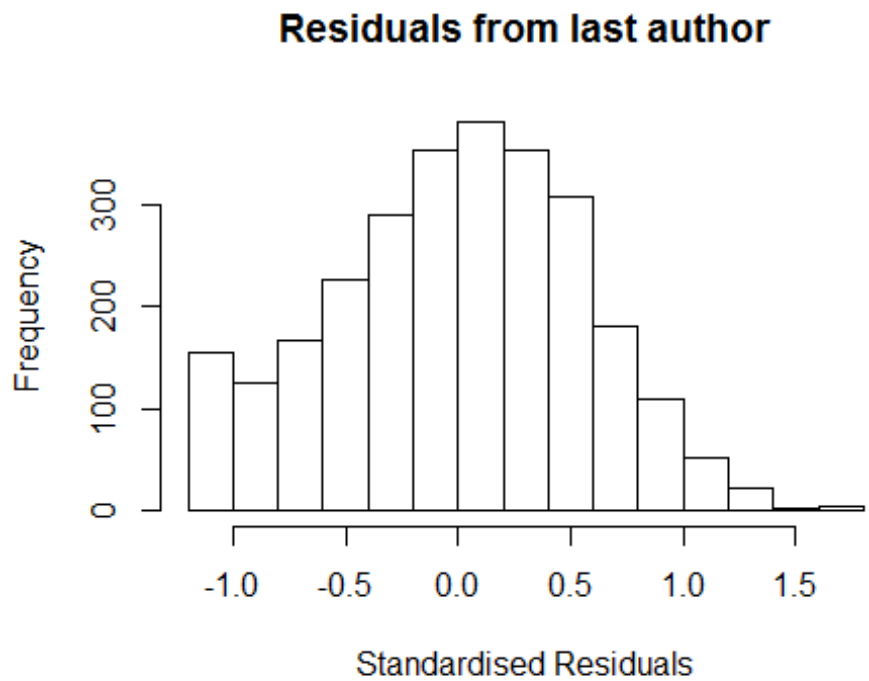

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.1698 -0.3899  0.0195  0.3919  1.7341
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)           1.1538     0.0436   26.44  < 2e-16 ***
## FirstAuthorFemale1   -0.0443     0.0238   -1.86   0.0624 .
## LastAuthorFemale1    -0.0268     0.0264   -1.01   0.3105
## Year1997             -0.0863     0.0613   -1.41   0.1591
## Year1998             -0.3368     0.0827   -4.07  4.8e-05 ***
## Year1999             -0.2471     0.0850   -2.91   0.0037 **
## Year2000             -0.0917     0.0830   -1.11   0.2691
## Year2001             -0.2232     0.0891   -2.51   0.0122 *
## Year2002             -0.1453     0.0714   -2.04   0.0419 *
## Year2003             -0.1567     0.0600   -2.61   0.0091 **
## Year2004             -0.2479     0.0568   -4.36  1.3e-05 ***
## Year2005             -0.2847     0.0629   -4.52  6.3e-06 ***
```

```
## Year2006            -0.0572     0.0580   -0.99   0.3238
## Year2007            -0.0424     0.0598   -0.71   0.4780
## Year2008             0.0160     0.0672    0.24   0.8121
## Year2009            -0.0198     0.0627   -0.32   0.7522
## Year2010            -0.0600     0.0578   -1.04   0.2993
## Year2011            -0.0662     0.0593   -1.12   0.2643
## Year2012            -0.0656     0.0634   -1.03   0.3009
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.581
## Multiple R-squared:  0.0287, Adjusted R-squared:  0.0222
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  216 weights are ~= 1. The remaining 2511 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.353   0.870   0.949   0.913   0.985   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          3.67e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.029  1           1.014
## Year              1.029 16           1.001
```




```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.1638 -0.3880  0.0214  0.3943  1.7411
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)           1.1511     0.0437   26.35  < 2e-16 ***
## FirstAuthorFemale1   -0.0473     0.0239   -1.98   0.0475 *
## Year1997             -0.0878     0.0615   -1.43   0.1531
## Year1998             -0.3412     0.0823   -4.14  3.5e-05 ***
## Year1999             -0.2499     0.0847   -2.95   0.0032 **
## Year2000             -0.0934     0.0830   -1.13   0.2602
## Year2001             -0.2250     0.0894   -2.52   0.0119 *
## Year2002             -0.1487     0.0715   -2.08   0.0377 *
## Year2003             -0.1600     0.0600   -2.67   0.0077 **
## Year2004             -0.2512     0.0567   -4.43  9.8e-06 ***
## Year2005             -0.2875     0.0630   -4.56  5.3e-06 ***
## Year2006             -0.0602     0.0581   -1.04   0.2997
```

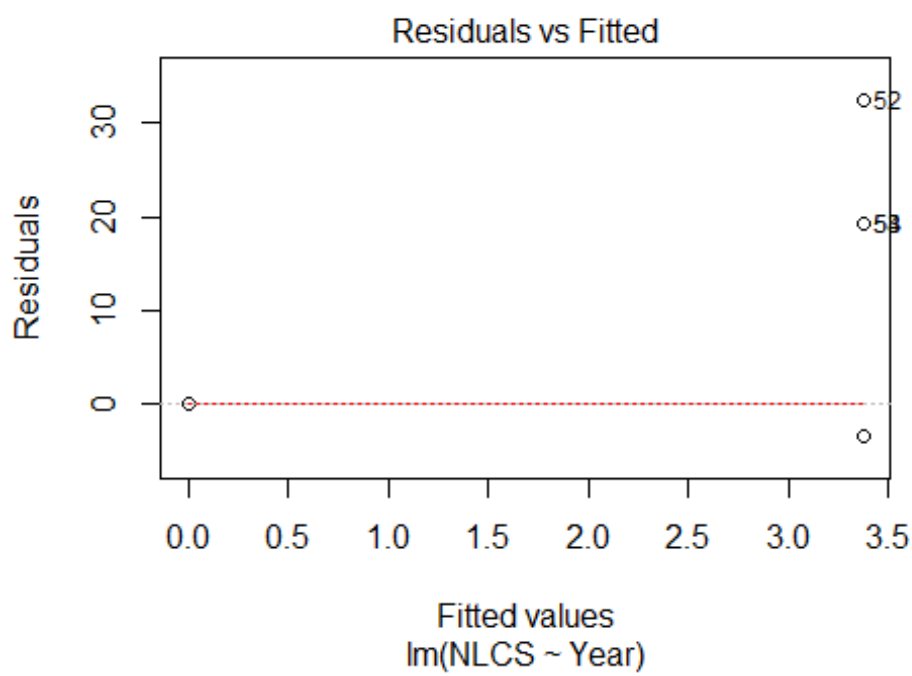
```
## Year2007            -0.0468     0.0596   -0.79   0.4320
## Year2008             0.0127     0.0672    0.19   0.8498
## Year2009            -0.0234     0.0627   -0.37   0.7095
## Year2010            -0.0626     0.0579   -1.08   0.2796
## Year2011            -0.0700     0.0592   -1.18   0.2371
## Year2012            -0.0687     0.0633   -1.09   0.2777
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.582
## Multiple R-squared:  0.0283, Adjusted R-squared:  0.0222
## Convergence in 10 IRWLS iterations
##
## Robustness weights:
##  224 weights are ~= 1. The remaining 2503 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.351   0.869   0.949   0.913   0.985   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          3.67e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##                   GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.024  1           1.012
## Year             1.024 16           1.001
```




```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.1545 -0.3923  0.0237  0.3991  1.7435
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.1433     0.0435   26.28  < 2e-16 ***
## LastAuthorFemale1  -0.0328     0.0265   -1.24   0.2165
## Year1997           -0.0881     0.0613   -1.44   0.1510
## Year1998           -0.3358     0.0829   -4.05  5.3e-05 ***
## Year1999           -0.2475     0.0851   -2.91   0.0037 **
## Year2000           -0.0916     0.0827   -1.11   0.2683
## Year2001           -0.2299     0.0890   -2.58   0.0098 **
## Year2002           -0.1475     0.0715   -2.06   0.0391 *
## Year2003           -0.1570     0.0602   -2.61   0.0092 **
## Year2004           -0.2509     0.0568   -4.42  1.0e-05 ***
## Year2005           -0.2865     0.0631   -4.54  5.9e-06 ***
## Year2006           -0.0635     0.0578   -1.10   0.2717
```

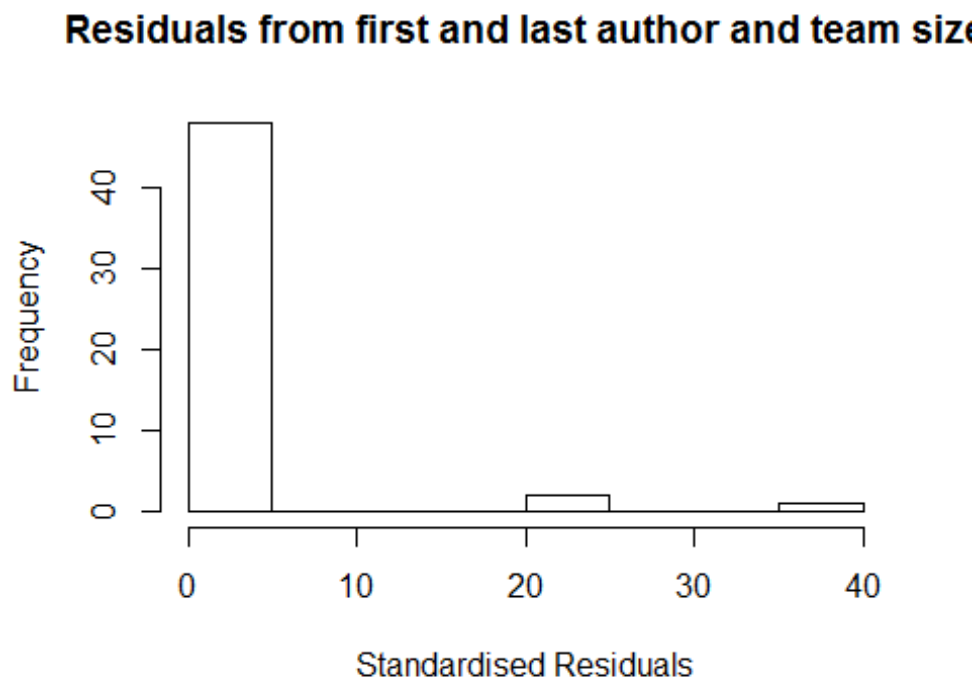
```
## Year2007           -0.0455     0.0598   -0.76   0.4466
## Year2008            0.0112     0.0672    0.17   0.8677
## Year2009           -0.0263     0.0628   -0.42   0.6755
## Year2010           -0.0640     0.0578   -1.11   0.2682
## Year2011           -0.0718     0.0591   -1.22   0.2245
## Year2012           -0.0713     0.0634   -1.12   0.2611
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.582
## Multiple R-squared:  0.0274, Adjusted R-squared:  0.0213
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  241 weights are ~= 1. The remaining 2486 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.349   0.867   0.948   0.913   0.984   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          3.67e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Sample size for the above analysis:  2727"
## [1] ""
## [1] ""
## [1] "####################################"
## [1] "Analysis of AJSC 3608"
## [1] "####################################"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2004 2005 2006 2007 2008 2009
##    2   20   20    9   25   13    3   10   19   12
##
## 1996 1997 1998 1999 2004 2005 2006 2007 2008 2009
##    0   20   17    7   24   11    3   10   19   12
##
## 1996 1997 1998 1999 2004 2005 2006 2007 2008 2009
##    0   20   17    7   24   11    3   10   19   12
```

```
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
##  Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = Inf, df = 3, p-value <2e-16

## [1] "Female first author team size 2018 geometric mean: NaN"
## [1] "Male first author team size 2018 geometric mean: NaN"
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"

## Warning in lmrob.S(x, y, control = control, mf = mf): S-estimated scale ==
## 0: Probably exact fit; check your data

## Warning in lmrob.S(x, y, control = control, mf = mf): diag(.) had 0 or NA
## entries; non-finite result is doubtful
```

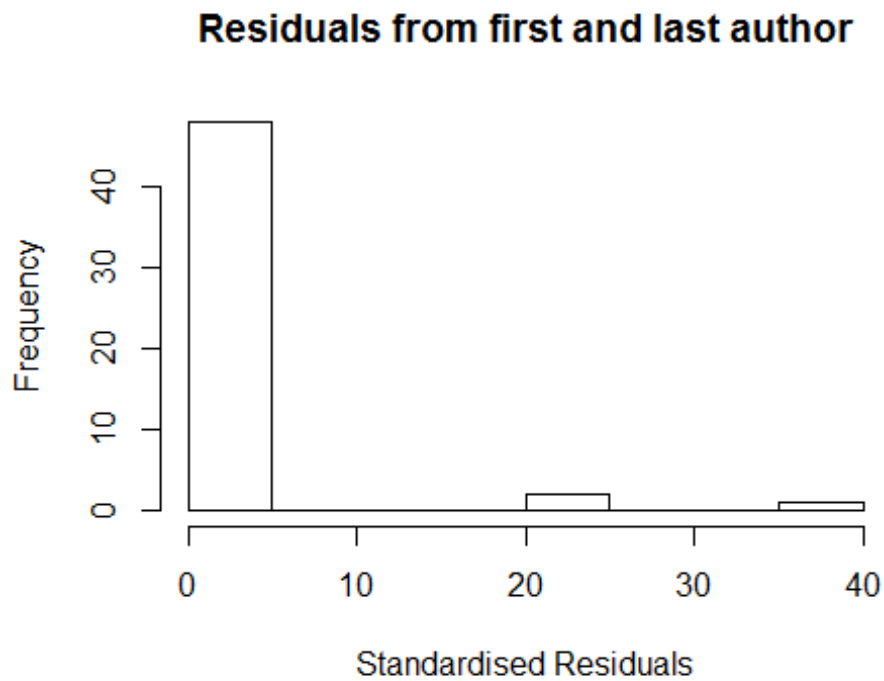



```
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale   NaN  1             NaN
## LastAuthorFemale    NaN  1             NaN
## UniqueAuthors       NaN  2             NaN
## Year                NaN  3             NaN

## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
```

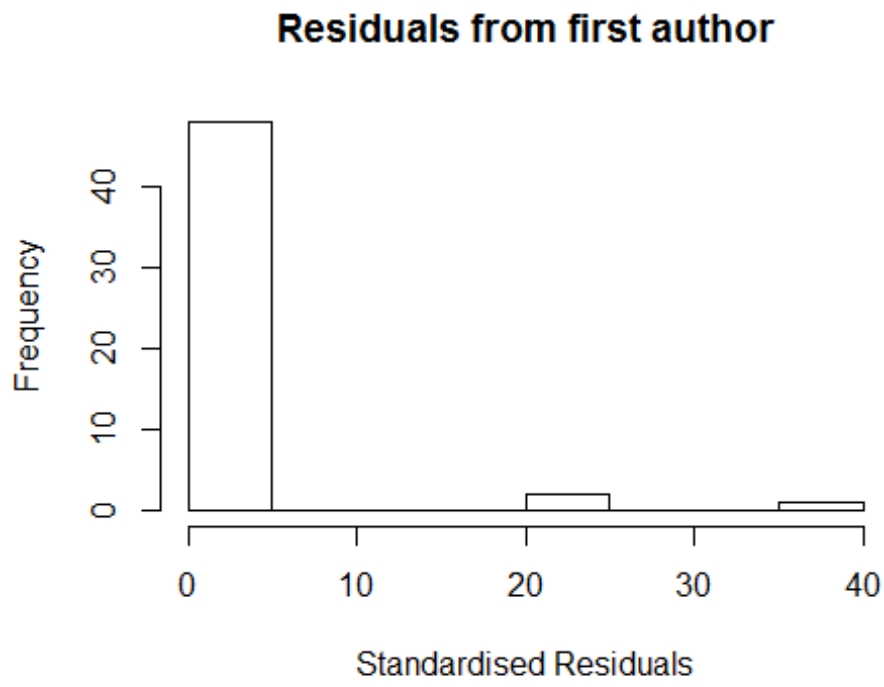
```
## Warning in lmrob.S(x, y, control = control, mf = mf): S-estimated scale ==
## 0: Probably exact fit; check your data

## Warning in lmrob.S(x, y, control = control, mf = mf): diag(.) had 0 or NA
## entries; non-finite result is doubtful
```



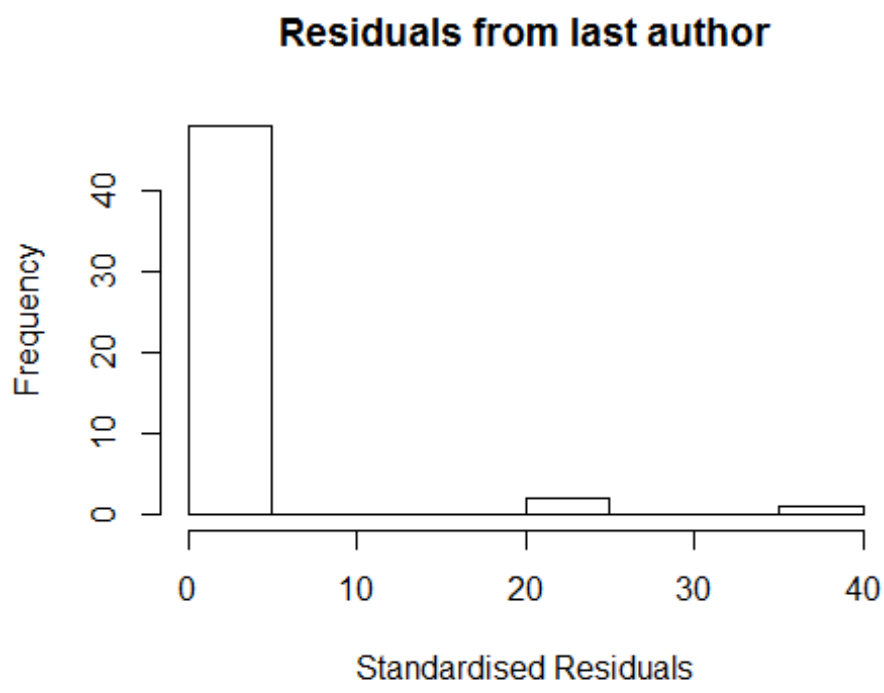

```
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale  NaN  1             NaN
## LastAuthorFemale   NaN  1             NaN
## Year               NaN  3             NaN

## [1] "Regression 3: First author gender, Year as factors"

## Warning in lmrob.S(x, y, control = control, mf = mf): S-estimated scale ==
## 0: Probably exact fit; check your data

## Warning in lmrob.S(x, y, control = control, mf = mf): diag(.) had 0 or NA
## entries; non-finite result is doubtful
```



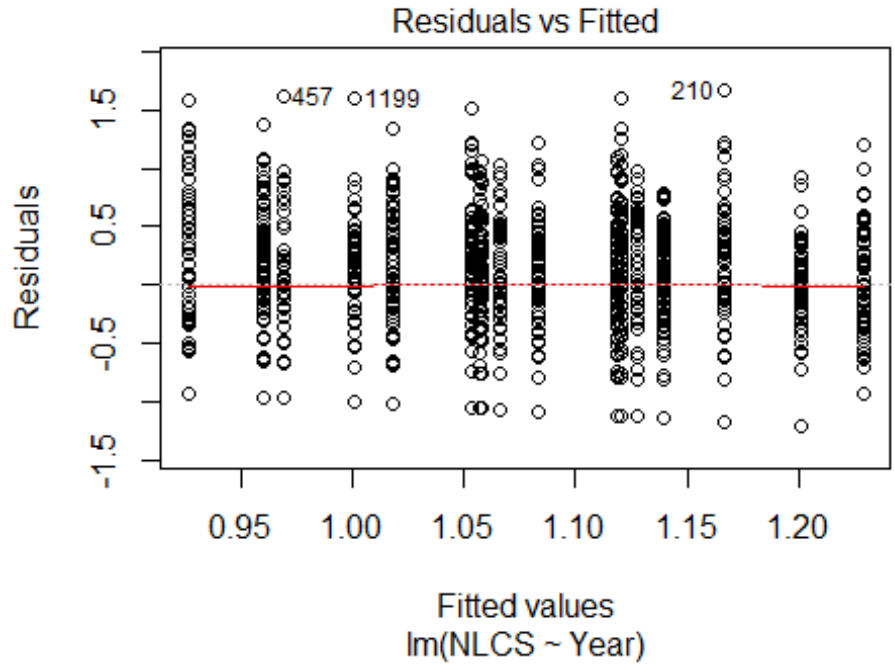

```
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale  NaN  1             NaN
## Year               NaN  3             NaN

## [1] "Regression 4: Last author gender, Year as factors"

## Warning in lmrob.S(x, y, control = control, mf = mf): S-estimated scale ==
## 0: Probably exact fit; check your data

## Warning in lmrob.S(x, y, control = control, mf = mf): diag(.) had 0 or NA
## entries; non-finite result is doubtful
```



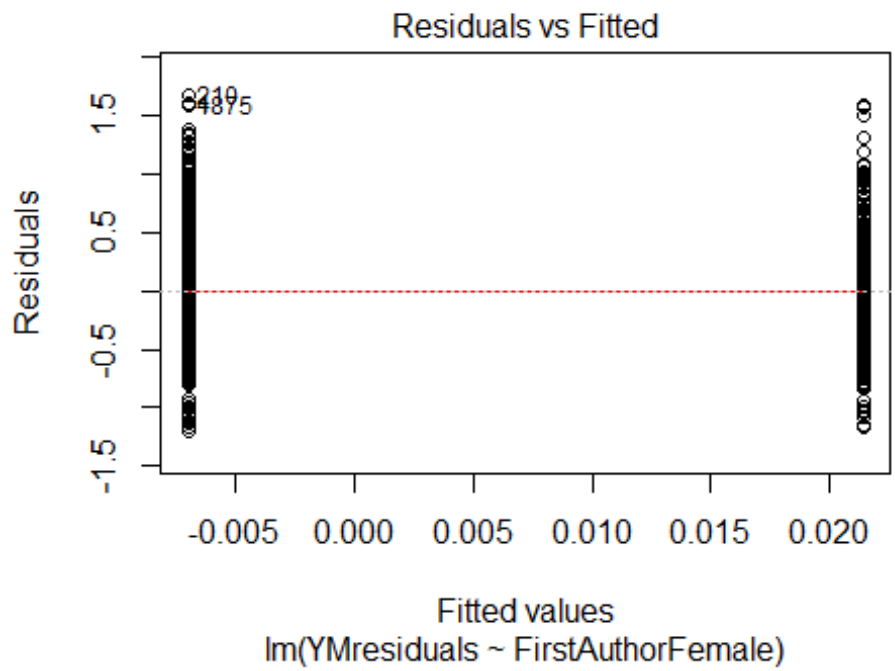

```
##                   GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale   NaN  1             NaN
## Year               NaN  3             NaN
```



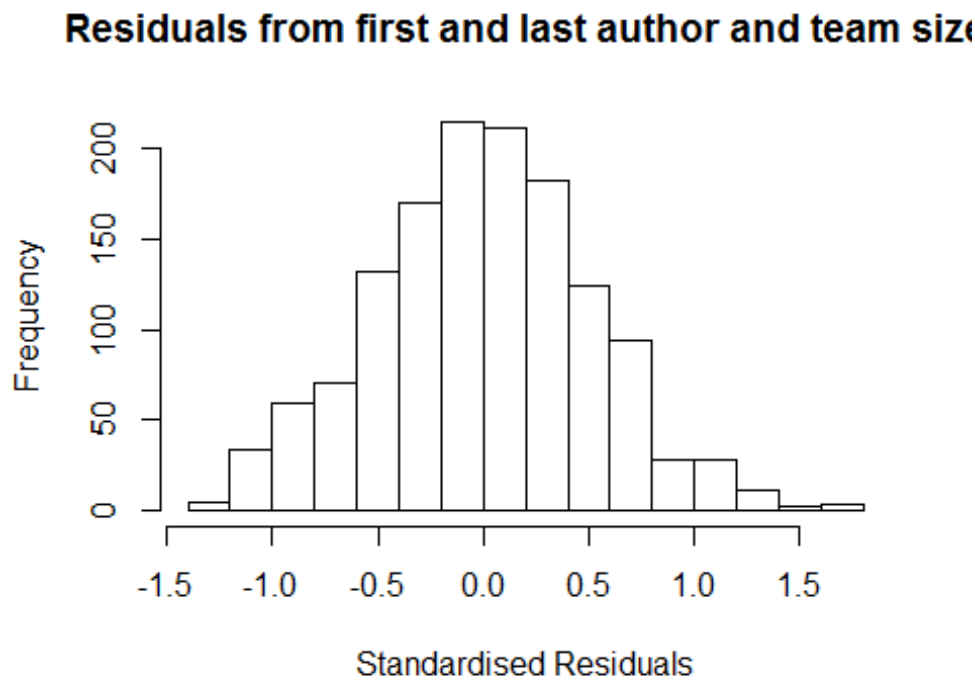
```
## [1] "Sample size for the above analysis:  123"
## [1] ""
## [1] ""
## [1] "####################################"
## [1] "Analysis of AJSC 3609"
## [1] "####################################"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  134  113  122   74   95  122   83   74   65   79   80   91  123  125  158
## 2011 2012
##  147  129
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##   98   82   97   56   61   88   74   68   57   73   71   80  112  106  128
## 2011 2012
##  127  120
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##   94   78   93   52   54   86   63   62   51   68   61   64  102   96  117
## 2011 2012
##  117  110
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
##  Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 48, df = 16, p-value = 4e-05
```




```
##
##  Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 0.42, df = 1, p-value = 0.5
```

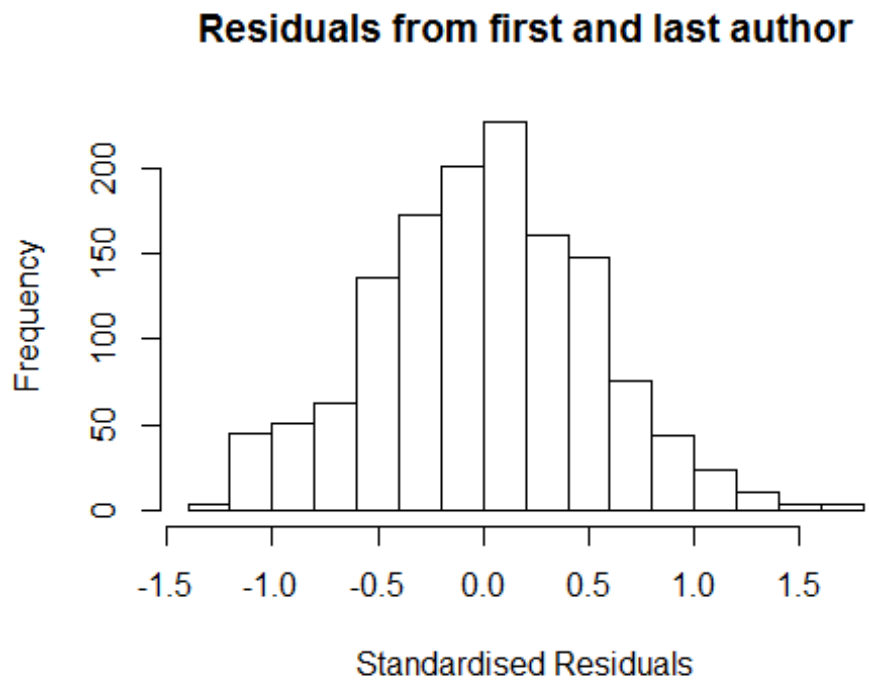



```
## [1] "Female first author team size 2018 geometric mean: 2.86716551938664"
## [1] "Male first author team size 2018 geometric mean: 3.59287880529774"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 570, p-value = 0.2
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 2.63776777429753"
## [1] "Male last author team size 2018 geometric mean: 3.91402691429484"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 680, p-value = 0.005
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##                     GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.197  1           1.094
## LastAuthorFemale  1.268  1           1.126
## UniqueAuthors     1.436  4           1.046
## Year              1.535 16           1.013
```



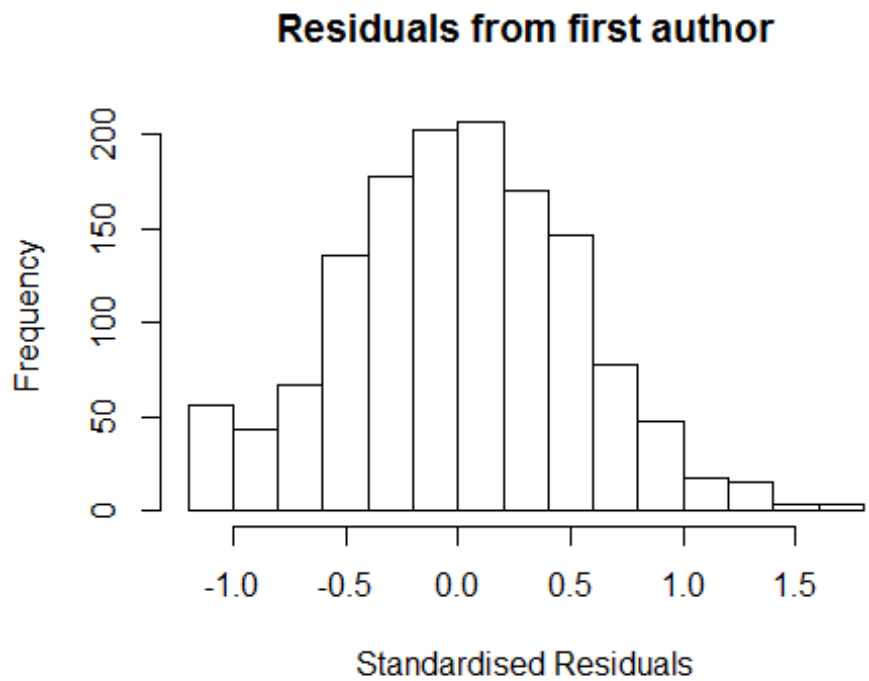

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
UniqueAuthors +
##     Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##     k.max = 1000))
##  \--> method = "MM"
## Residuals:
##       Min        1Q    Median        3Q       Max
## -1.323672 -0.333569  0.000413  0.346128  1.701903
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.14393    0.07853   14.57  < 2e-16 ***
## FirstAuthorFemale1  0.00504    0.03449    0.15  0.88386
## LastAuthorFemale1  -0.09492    0.03272   -2.90  0.00378 **
## UniqueAuthors2      0.09352    0.04111    2.28  0.02305 *
## UniqueAuthors3      0.14579    0.04328    3.37  0.00078 ***
## UniqueAuthors4      0.24183    0.04945    4.89  1.1e-06 ***
## UniqueAuthors5      0.20205    0.05071    3.98  7.1e-05 ***
## Year1997           -0.01648    0.09189   -0.18  0.85767
## Year1998           -0.09951    0.09541   -1.04  0.29714
## Year1999           -0.20719    0.09807   -2.11  0.03481 *
```

```
## Year2000           -0.12521    0.09902   -1.26  0.20624
## Year2001           -0.12386    0.09239   -1.34  0.18025
## Year2002           -0.12696    0.09890   -1.28  0.19946
## Year2003           -0.02735    0.08450   -0.32  0.74621
## Year2004           -0.04813    0.11774   -0.41  0.68278
## Year2005           -0.00944    0.08929   -0.11  0.91578
## Year2006           -0.12449    0.09703   -1.28  0.19971
## Year2007           -0.21410    0.09191   -2.33  0.01998 *
## Year2008           -0.24587    0.09272   -2.65  0.00810 **
## Year2009           -0.10255    0.09064   -1.13  0.25806
## Year2010           -0.09277    0.08226   -1.13  0.25964
## Year2011           -0.18374    0.09126   -2.01  0.04427 *
## Year2012           -0.30799    0.09638   -3.20  0.00143 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.51
## Multiple R-squared:  0.0577, Adjusted R-squared:  0.0423
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
##  115 weights are ~= 1. The remaining 1253 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.243   0.869   0.952   0.906   0.986   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          7.31e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.203  1           1.097
## LastAuthorFemale  1.238  1           1.113
## Year              1.119 16           1.004
```



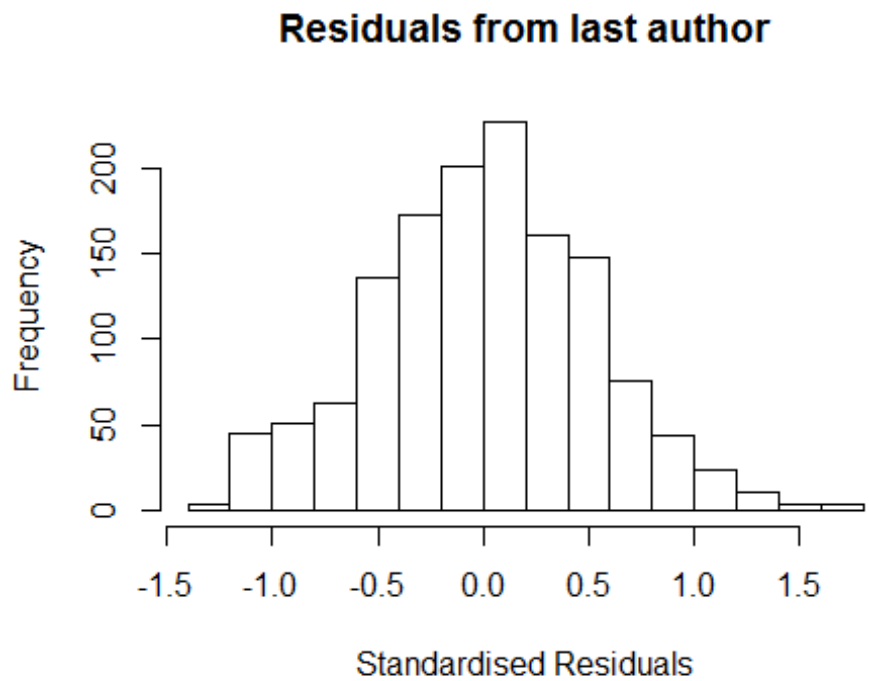

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.25801 -0.33801  0.00518  0.35414  1.72079
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.19774    0.07619   15.72  < 2e-16 ***
## FirstAuthorFemale1  0.00588    0.03491    0.17  0.86630
## LastAuthorFemale1  -0.11078    0.03292   -3.37  0.00079 ***
## Year1997            0.01938    0.09111    0.21  0.83161
## Year1998           -0.07570    0.09401   -0.81  0.42082
## Year1999           -0.16191    0.09565   -1.69  0.09073 .
## Year2000           -0.07226    0.09326   -0.77  0.43854
## Year2001           -0.05169    0.08764   -0.59  0.55542
## Year2002           -0.05431    0.09431   -0.58  0.56480
## Year2003            0.05439    0.08156    0.67  0.50495
## Year2004            0.03339    0.11114    0.30  0.76392
## Year2005            0.07720    0.08905    0.87  0.38614
```

```
## Year2006           -0.05079    0.09437   -0.54  0.59058
## Year2007           -0.11896    0.08902   -1.34  0.18165
## Year2008           -0.16856    0.09061   -1.86  0.06306 .
## Year2009           -0.02686    0.08739   -0.31  0.75862
## Year2010           -0.00909    0.07875   -0.12  0.90816
## Year2011           -0.10843    0.08808   -1.23  0.21851
## Year2012           -0.23583    0.09320   -2.53  0.01150 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.511
## Multiple R-squared:  0.0347, Adjusted R-squared:  0.0218
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
##  125 weights are ~= 1. The remaining 1243 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.234   0.873   0.947   0.904   0.985   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          7.31e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.041  1           1.020
## Year              1.041 16           1.001
```




```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##       Min        1Q    Median        3Q       Max
## -1.194801 -0.332243  0.000225  0.360239  1.698687
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)          1.14329    0.07447   15.35   <2e-16 ***
## FirstAuthorFemale1 -0.03975    0.03253   -1.22    0.222
## Year1997             0.03078    0.09164    0.34    0.737
## Year1998            -0.06447    0.09356   -0.69    0.491
## Year1999            -0.15288    0.09745   -1.57    0.117
## Year2000            -0.06278    0.09375   -0.67    0.503
## Year2001            -0.03832    0.08744   -0.44    0.661
## Year2002            -0.03597    0.09417   -0.38    0.703
## Year2003             0.09127    0.08145    1.12    0.263
## Year2004             0.03932    0.11163    0.35    0.725
## Year2005             0.08749    0.08880    0.99    0.325
## Year2006            -0.04801    0.09468   -0.51    0.612
```

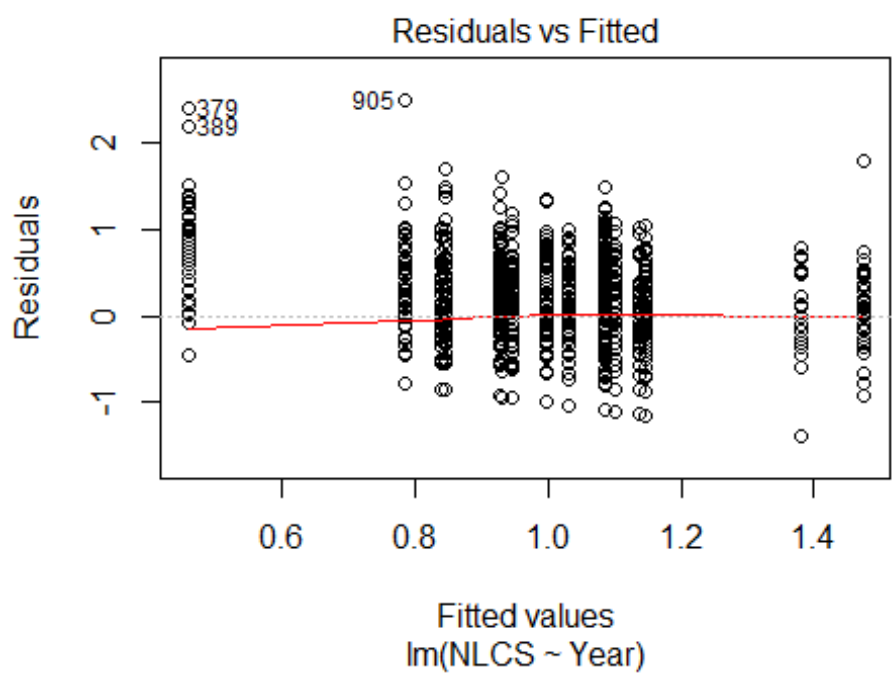
```
## Year2007           -0.11225    0.08897   -1.26    0.207
## Year2008           -0.15125    0.09166   -1.65    0.099 .
## Year2009           -0.00545    0.08718   -0.06    0.950
## Year2010            0.01210    0.07904    0.15    0.878
## Year2011           -0.10191    0.08830   -1.15    0.249
## Year2012           -0.22240    0.09310   -2.39    0.017 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.514
## Multiple R-squared:  0.0269, Adjusted R-squared:  0.0146
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
##  121 weights are ~= 1. The remaining 1247 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.252   0.871   0.947   0.904   0.984   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          7.31e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##                  GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.07  1           1.035
## Year             1.07 16           1.002
```




```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.25575 -0.33971  0.00367  0.35413  1.72150
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.20084    0.07343   16.35   <2e-16 ***
## LastAuthorFemale1 -0.10868    0.03061   -3.55   0.0004 ***
## Year1997            0.01935    0.09110    0.21   0.8318
## Year1998           -0.07527    0.09401   -0.80   0.4235
## Year1999           -0.16215    0.09563   -1.70   0.0902 .
## Year2000           -0.07257    0.09315   -0.78   0.4361
## Year2001           -0.05194    0.08763   -0.59   0.5534
## Year2002           -0.05444    0.09432   -0.58   0.5639
## Year2003            0.05491    0.08154    0.67   0.5008
## Year2004            0.03322    0.11116    0.30   0.7651
## Year2005            0.07674    0.08902    0.86   0.3888
## Year2006           -0.05154    0.09417   -0.55   0.5843
```

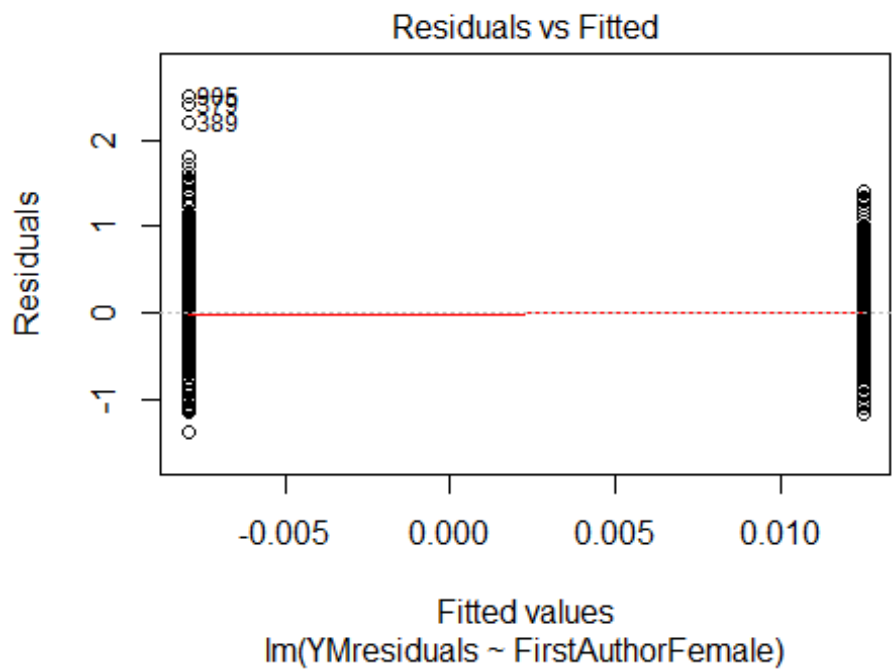
```
## Year2007          -0.11972    0.08883   -1.35   0.1780
## Year2008          -0.16871    0.09054   -1.86   0.0626 .
## Year2009          -0.02689    0.08735   -0.31   0.7582
## Year2010          -0.00901    0.07876   -0.11   0.9089
## Year2011          -0.10876    0.08792   -1.24   0.2163
## Year2012          -0.23573    0.09325   -2.53   0.0116 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.511
## Multiple R-squared:  0.0347, Adjusted R-squared:  0.0226
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
##  126 weights are ~= 1. The remaining 1242 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.233   0.874   0.947   0.904   0.985   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          7.31e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Sample size for the above analysis:  1368"
## [1] ""
## [1] ""
## [1] "####################################"
## [1] "Analysis of AJSC 3610"
## [1] "####################################"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##   67   52   25   75  165   78   68   47   56   67   74  101  105  130  121
## 2011 2012
##  135   92
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##   56   42   20   58  133   68   52   40   44   54   54   84   84  109   96
## 2011 2012
```

```
##  118   71
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##   51   39   16   52  120   65   42   34   37   48   49   78   75   98   84
## 2011 2012
##   97   60
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
##  Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 36, df = 16, p-value = 0.003
```

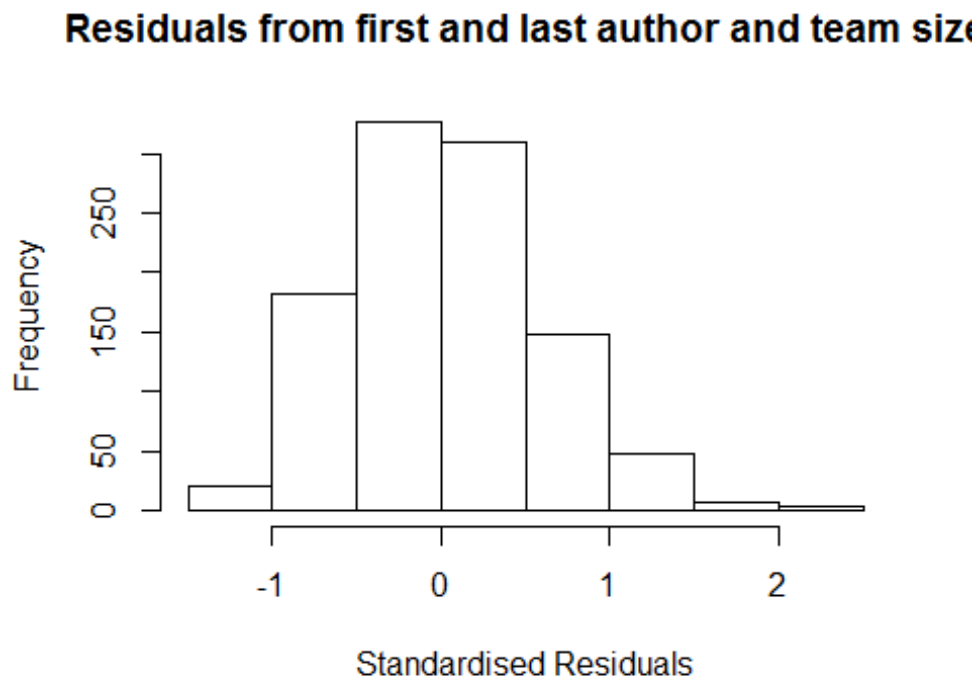



```
##
##  Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 8, df = 1, p-value = 0.005

## [1] "Female first author team size 2018 geometric mean: 3.40391489988718"
## [1] "Male first author team size 2018 geometric mean: 3.25783671487437"

## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties
```

```
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 440, p-value = 0.8
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 3.32878585280834"
## [1] "Male last author team size 2018 geometric mean: 3.30464965742378"

## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties
```

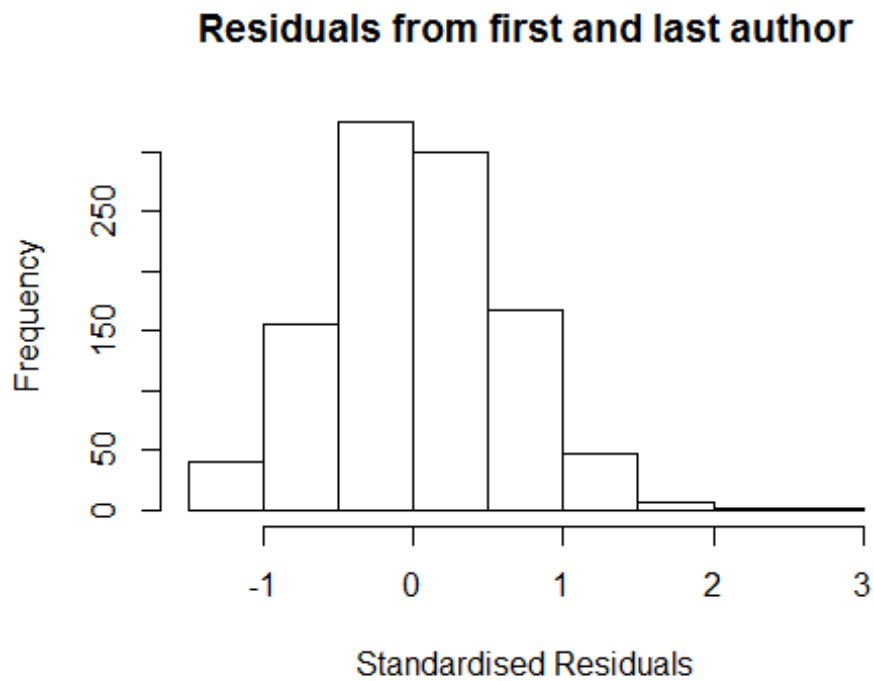



```
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 420, p-value = 0.8
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##                     GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.226  1           1.107
## LastAuthorFemale  1.179  1           1.086
## UniqueAuthors     1.449  4           1.047
## Year              1.571 16           1.014
```



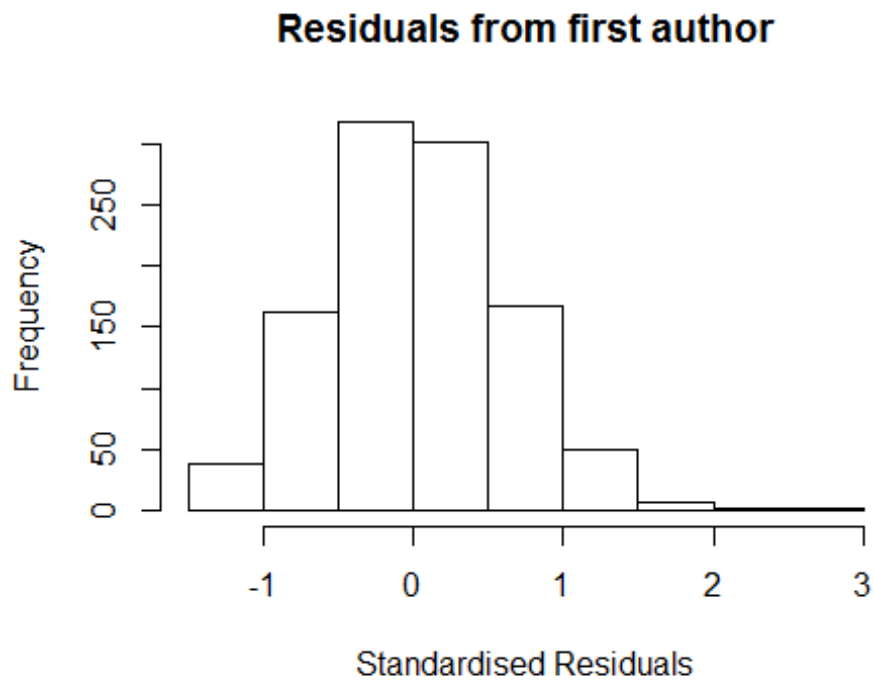

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
UniqueAuthors +
##     Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##     k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.3599 -0.3969 -0.0101  0.4029  2.2776
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         0.89510    0.13247    6.76  2.4e-11 ***
## FirstAuthorFemale1 -0.05646    0.04049   -1.39    0.164
## LastAuthorFemale1   0.07152    0.04239    1.69    0.092 .
## UniqueAuthors2      0.11025    0.06205    1.78    0.076 .
## UniqueAuthors3      0.24240    0.06106    3.97  7.7e-05 ***
## UniqueAuthors4      0.38892    0.06378    6.10  1.5e-09 ***
## UniqueAuthors5      0.50522    0.06805    7.42  2.4e-13 ***
## Year1997            0.36424    0.14962    2.43    0.015 *
## Year1998            0.30834    0.17050    1.81    0.071 .
## Year1999            0.00438    0.14673    0.03    0.976
```

```
## Year2000           -0.77211    0.14424   -5.35  1.1e-07 ***
## Year2001           -0.09102    0.14711   -0.62    0.536
## Year2002           -0.03686    0.15447   -0.24    0.811
## Year2003           -0.01806    0.15533   -0.12    0.907
## Year2004           -0.24872    0.16042   -1.55    0.121
## Year2005           -0.02222    0.15145   -0.15    0.883
## Year2006           -0.27998    0.15874   -1.76    0.078 .
## Year2007           -0.31102    0.15521   -2.00    0.045 *
## Year2008           -0.25987    0.14883   -1.75    0.081 .
## Year2009           -0.14870    0.14469   -1.03    0.304
## Year2010           -0.18718    0.14489   -1.29    0.197
## Year2011           -0.17395    0.14532   -1.20    0.232
## Year2012           -0.08009    0.15997   -0.50    0.617
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.565
## Multiple R-squared:  0.227,  Adjusted R-squared:  0.21
## Convergence in 16 IRWLS iterations
##
## Robustness weights:
##  76 weights are ~= 1. The remaining 969 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0673  0.8660  0.9460  0.9050  0.9830  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          9.57e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.157  1           1.075
## LastAuthorFemale  1.138  1           1.067
## Year              1.192 16           1.006
```



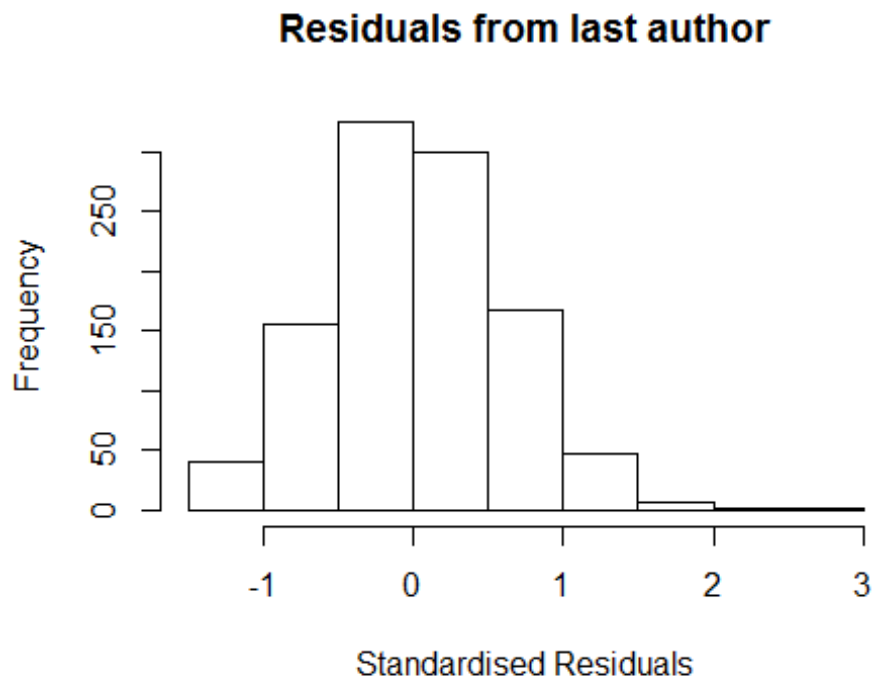

```
## [1] "List of  1 outliers with residuals above  2.5"
##       ScopusId NLCS Year OneField Fields residuals
## 379 0034035002 2.87 2000     2731      2      2.54
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.21431 -0.39658  0.00469  0.42605  2.54046
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)           1.0244     0.1537    6.67  4.3e-11 ***
## FirstAuthorFemale1   -0.0161     0.0413   -0.39   0.6965
## LastAuthorFemale1     0.1148     0.0439    2.61   0.0091 **
## Year1997              0.4436     0.1710    2.59   0.0096 **
## Year1998              0.4614     0.1915    2.41   0.0162 *
## Year1999              0.0751     0.1688    0.44   0.6564
## Year2000             -0.6949     0.1630   -4.26  2.2e-05 ***
## Year2001             -0.0428     0.1681   -0.25   0.7989
## Year2002              0.0433     0.1759    0.25   0.8056
## Year2003              0.0589     0.1780    0.33   0.7406
## Year2004             -0.1291     0.1761   -0.73   0.4639
## Year2005              0.0567     0.1729    0.33   0.7430
```

```
## Year2006            -0.1856     0.1778   -1.04   0.2967
## Year2007            -0.3225     0.1776   -1.82   0.0696 .
## Year2008            -0.2200     0.1703   -1.29   0.1967
## Year2009            -0.0925     0.1671   -0.55   0.5802
## Year2010            -0.1253     0.1669   -0.75   0.4529
## Year2011            -0.1235     0.1670   -0.74   0.4597
## Year2012             0.0558     0.1861    0.30   0.7646
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.584
## Multiple R-squared:  0.162,  Adjusted R-squared:  0.147
## Convergence in 18 IRWLS iterations
##
## Robustness weights:
##  75 weights are ~= 1. The remaining 970 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0187  0.8530  0.9500  0.9040  0.9820  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          9.57e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.104  1           1.051
## Year              1.104 16           1.003
```




```
## [1] "List of  1 outliers with residuals above  2.5"
##       ScopusId NLCS Year OneField Fields residuals
## 379 0034035002 2.87 2000     2731      2      2.54
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.14268 -0.37326  0.00634  0.42548  2.56061
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)           1.0370     0.1580    6.56  8.3e-11 ***
## FirstAuthorFemale1    0.0118     0.0405    0.29    0.771
## Year1997              0.4438     0.1739    2.55    0.011 *
## Year1998              0.4813     0.1943    2.48    0.013 *
## Year1999              0.0939     0.1710    0.55    0.583
## Year2000             -0.6756     0.1658   -4.07  5.0e-05 ***
## Year2001             -0.0202     0.1707   -0.12    0.906
## Year2002              0.0501     0.1788    0.28    0.779
## Year2003              0.0727     0.1819    0.40    0.690
## Year2004             -0.1191     0.1788   -0.67    0.505
## Year2005              0.0730     0.1756    0.42    0.678
## Year2006             -0.1889     0.1819   -1.04    0.299
```

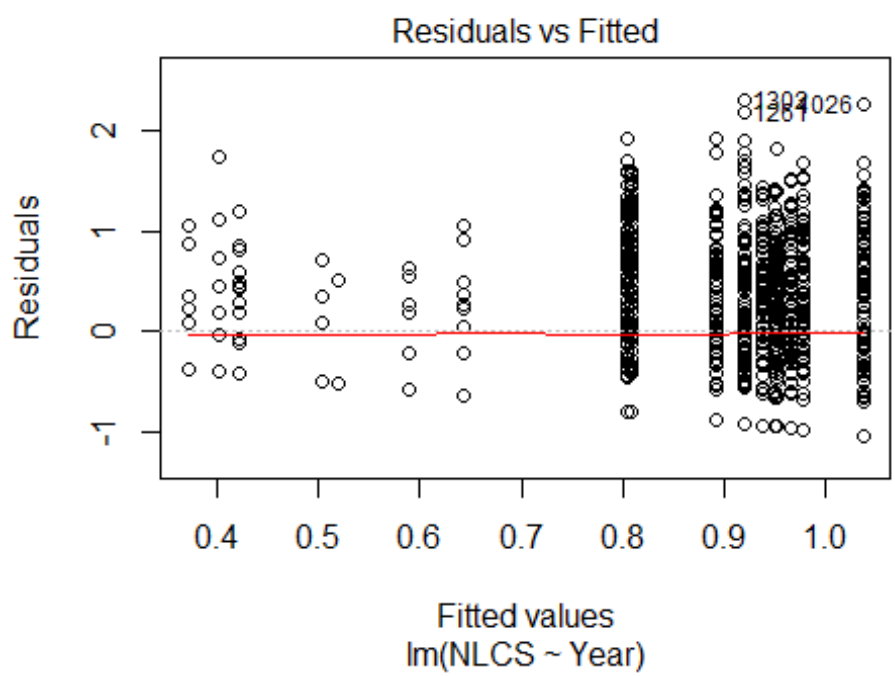
```
## Year2007            -0.3156     0.1804   -1.75    0.081 .
## Year2008            -0.2105     0.1741   -1.21    0.227
## Year2009            -0.0667     0.1692   -0.39    0.694
## Year2010            -0.1068     0.1697   -0.63    0.529
## Year2011            -0.1163     0.1703   -0.68    0.495
## Year2012             0.0762     0.1894    0.40    0.688
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.584
## Multiple R-squared:  0.156,  Adjusted R-squared:  0.142
## Convergence in 18 IRWLS iterations
##
## Robustness weights:
##  91 weights are ~= 1. The remaining 954 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0153  0.8580  0.9510  0.9020  0.9800  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          9.57e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##                   GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.093  1           1.046
## Year             1.093 16           1.003
```



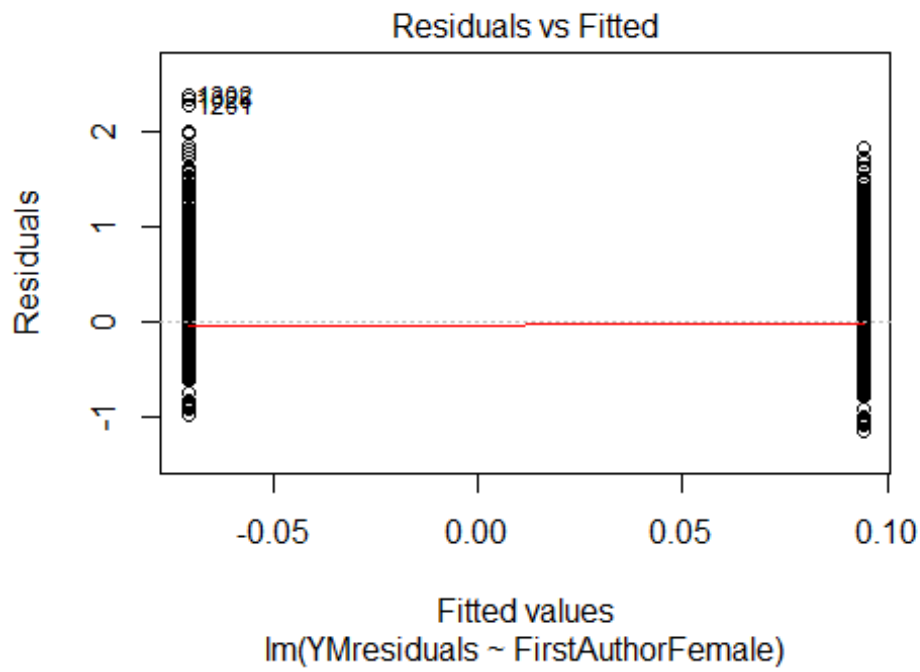

```
## [1] "List of  1 outliers with residuals above  2.5"
##       ScopusId NLCS Year OneField Fields residuals
## 379 0034035002 2.87 2000     2731      2      2.54
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.20616 -0.39267  0.00737  0.42389  2.54511
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)          1.0210     0.1527    6.69  3.7e-11 ***
## LastAuthorFemale1    0.1106     0.0431    2.57   0.0104 *
## Year1997             0.4434     0.1707    2.60   0.0095 **
## Year1998             0.4632     0.1915    2.42   0.0157 *
## Year1999             0.0745     0.1687    0.44   0.6587
## Year2000            -0.6961     0.1630   -4.27  2.1e-05 ***
## Year2001            -0.0433     0.1680   -0.26   0.7964
## Year2002             0.0437     0.1757    0.25   0.8034
## Year2003             0.0572     0.1780    0.32   0.7481
## Year2004            -0.1297     0.1758   -0.74   0.4608
## Year2005             0.0546     0.1727    0.32   0.7519
## Year2006            -0.1879     0.1779   -1.06   0.2911
```

```
## Year2007           -0.3217     0.1773   -1.81   0.0699 .
## Year2008           -0.2215     0.1700   -1.30   0.1929
## Year2009           -0.0950     0.1671   -0.57   0.5699
## Year2010           -0.1298     0.1667   -0.78   0.4362
## Year2011           -0.1262     0.1668   -0.76   0.4494
## Year2012            0.0507     0.1856    0.27   0.7848
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.585
## Multiple R-squared:  0.161,  Adjusted R-squared:  0.147
## Convergence in 18 IRWLS iterations
##
## Robustness weights:
##  80 weights are ~= 1. The remaining 965 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0189  0.8530  0.9500  0.9040  0.9810  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          9.57e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Sample size for the above analysis:  1045"
## [1] ""
## [1] ""
## [1] "####################################"
## [1] "Analysis of AJSC 3611"
## [1] "####################################"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##   27   21   22    2   14   33   32  118  108  154  149  186  177  197  174
## 2011 2012
##  261  282
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##   20   12   14    2    7    8   29  106   95  133  133  160  145  171  151
## 2011 2012
```

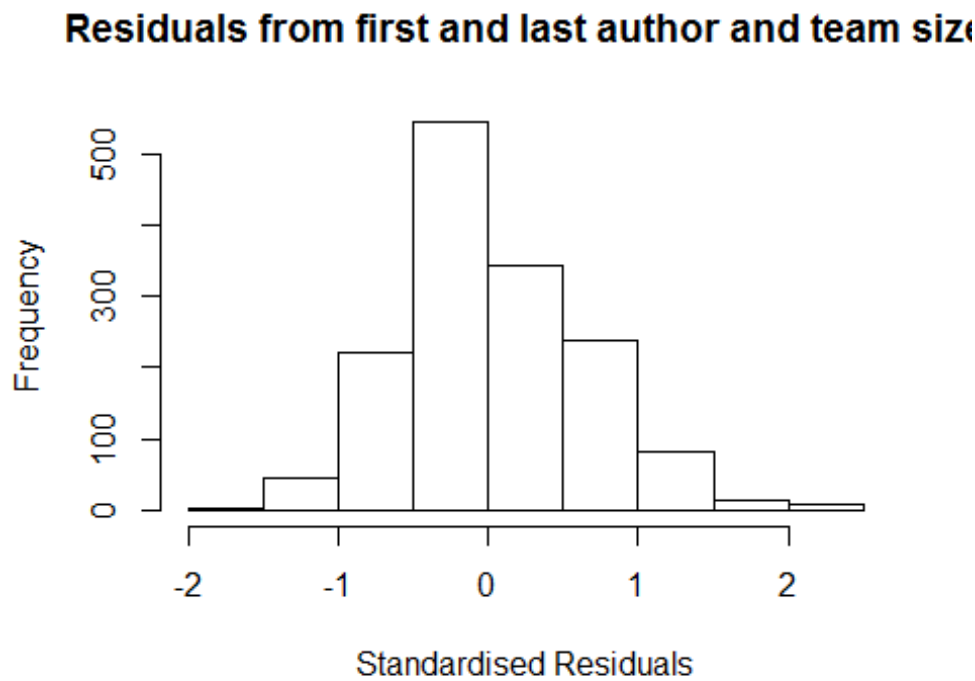
```
##  228  244
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##   19    9   14    2    6    7   25   97   87  118  120  150  136  153  131
## 2011 2012
##  210  213
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
##  Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 23, df = 16, p-value = 0.1
```




```
##
##  Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 0.35, df = 1, p-value = 0.6
```

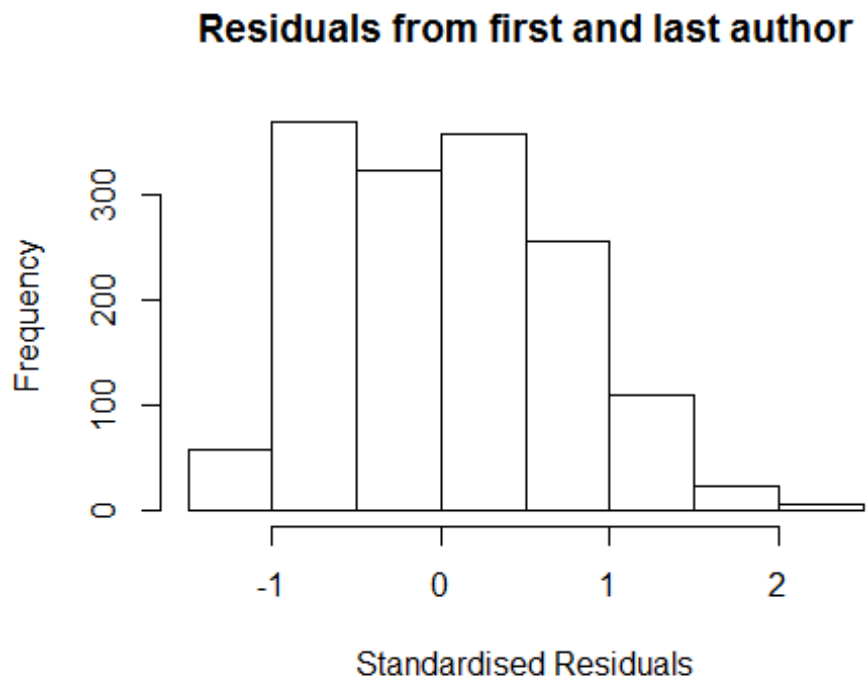



```
## [1] "Female first author team size 2018 geometric mean: 4.0361618065674"
## [1] "Male first author team size 2018 geometric mean: 3.50084592676047"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 8500, p-value = 0.3
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 3.86177274150242"
## [1] "Male last author team size 2018 geometric mean: 3.71445885453274"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 7500, p-value = 0.7
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.377  1           1.174
## LastAuthorFemale  1.310  1           1.144
## UniqueAuthors     1.362  4           1.039
## Year              1.344 16           1.009
```



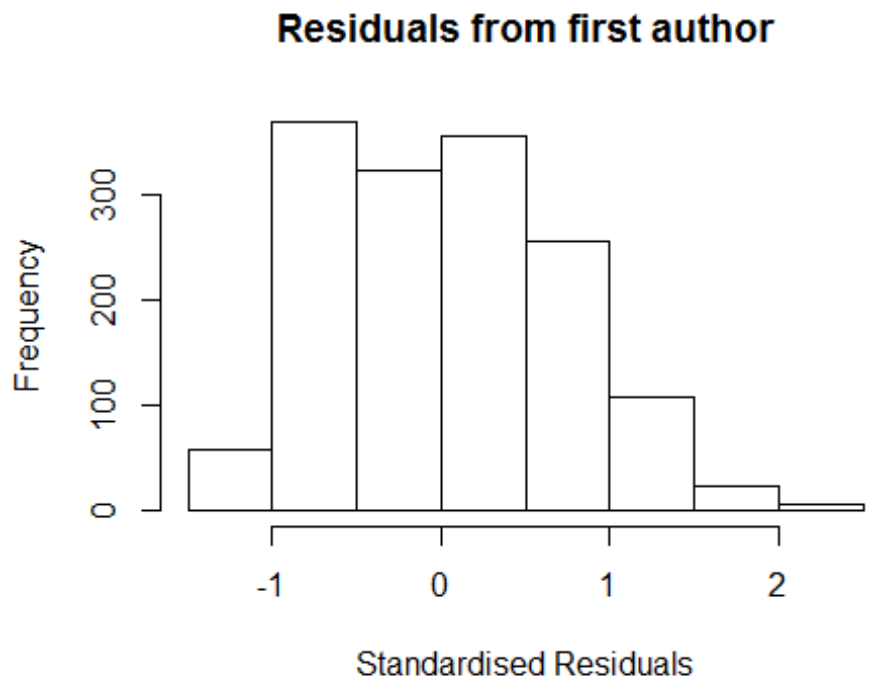

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
UniqueAuthors +
##     Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##     k.max = 1000))
##  \--> method = "MM"
## Residuals:
##    Min     1Q Median     3Q    Max
## -1.582 -0.402 -0.057  0.440  2.253
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)          0.1025     0.1413    0.73    0.468
## FirstAuthorFemale1   0.0971     0.0382    2.54    0.011 *
## LastAuthorFemale1   -0.0126     0.0377   -0.34    0.737
## UniqueAuthors2       0.4901     0.0485   10.11   <2e-16 ***
## UniqueAuthors3       0.6464     0.0506   12.77   <2e-16 ***
## UniqueAuthors4       0.7756     0.0514   15.10   <2e-16 ***
## UniqueAuthors5       1.0417     0.0500   20.82   <2e-16 ***
## Year1997            -0.2015     0.2003   -1.01    0.314
## Year1998             0.1244     0.2063    0.60    0.547
## Year1999             0.1724     0.2484    0.69    0.488
```

```
## Year2000            -0.4123     0.3068   -1.34    0.179
## Year2001             0.0225     0.2626    0.09    0.932
## Year2002            -0.1954     0.1688   -1.16    0.247
## Year2003             0.2458     0.1558    1.58    0.115
## Year2004             0.3147     0.1535    2.05    0.040 *
## Year2005             0.3404     0.1508    2.26    0.024 *
## Year2006             0.2162     0.1503    1.44    0.150
## Year2007             0.2887     0.1514    1.91    0.057 .
## Year2008             0.3158     0.1584    1.99    0.046 *
## Year2009             0.1782     0.1506    1.18    0.237
## Year2010             0.1515     0.1519    1.00    0.319
## Year2011             0.0366     0.1486    0.25    0.806
## Year2012             0.0398     0.1484    0.27    0.789
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.567
## Multiple R-squared:  0.307,  Adjusted R-squared:  0.297
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
##  103 weights are ~= 1. The remaining 1394 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0792  0.8560  0.9460  0.8960  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          6.68e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.116  1           1.056
## LastAuthorFemale  1.119  1           1.058
## Year              1.088 16           1.003
```



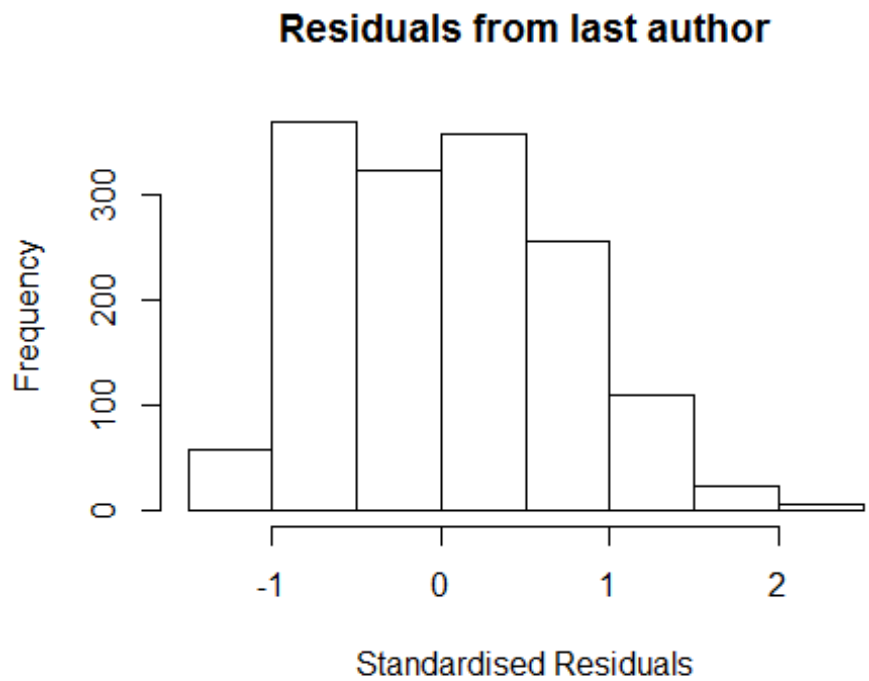

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.11993 -0.61667 -0.00329  0.53705  2.45207
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)           0.3097     0.1320    2.35  0.01910 *
## FirstAuthorFemale1    0.1935     0.0409    4.73  2.4e-06 ***
## LastAuthorFemale1     0.0133     0.0411    0.32  0.74672
## Year1997             -0.0807     0.1974   -0.41  0.68296
## Year1998              0.2937     0.2022    1.45  0.14656
## Year1999              0.2103     0.4243    0.50  0.62021
## Year2000              0.0280     0.2023    0.14  0.88993
## Year2001              0.1599     0.2172    0.74  0.46184
## Year2002             -0.0205     0.1563   -0.13  0.89545
## Year2003              0.5500     0.1491    3.69  0.00023 ***
## Year2004              0.5521     0.1535    3.60  0.00033 ***
## Year2005              0.5764     0.1501    3.84  0.00013 ***
```

```
## Year2006             0.4528     0.1480    3.06  0.00225 **
## Year2007             0.5334     0.1457    3.66  0.00026 ***
## Year2008             0.6034     0.1493    4.04  5.6e-05 ***
## Year2009             0.4672     0.1448    3.23  0.00128 **
## Year2010             0.5440     0.1428    3.81  0.00014 ***
## Year2011             0.3469     0.1409    2.46  0.01390 *
## Year2012             0.3539     0.1408    2.51  0.01208 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.756
## Multiple R-squared:  0.0522, Adjusted R-squared:  0.0407
## Convergence in 10 IRWLS iterations
##
## Robustness weights:
##  127 weights are ~= 1. The remaining 1370 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.271   0.885   0.938   0.920   0.983   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          6.68e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.038  1           1.019
## Year              1.038 16           1.001
```




```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.11510 -0.62273 -0.00535  0.53480  2.44797
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)           0.3103     0.1319    2.35  0.01879 *
## FirstAuthorFemale1    0.1977     0.0395    5.01  6.1e-07 ***
## Year1997             -0.0784     0.1967   -0.40  0.69025
## Year1998              0.2945     0.2016    1.46  0.14421
## Year1999              0.2097     0.4241    0.49  0.62102
## Year2000              0.0378     0.2006    0.19  0.85058
## Year2001              0.1616     0.2169    0.74  0.45646
## Year2002             -0.0185     0.1561   -0.12  0.90553
## Year2003              0.5524     0.1488    3.71  0.00021 ***
## Year2004              0.5552     0.1531    3.63  0.00030 ***
## Year2005              0.5795     0.1498    3.87  0.00011 ***
## Year2006              0.4560     0.1477    3.09  0.00206 **
```

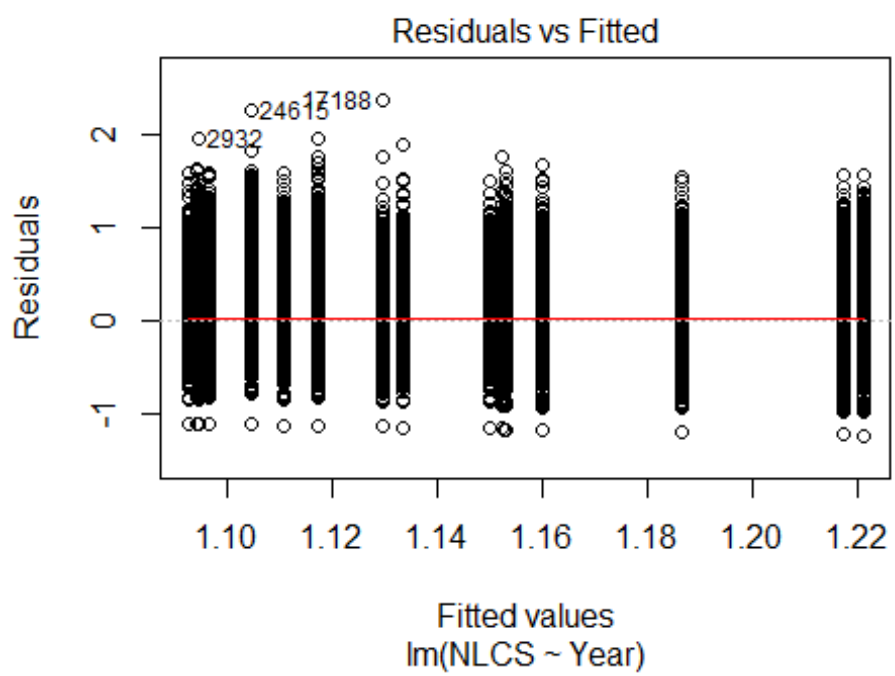
```
## Year2007             0.5370     0.1451    3.70  0.00022 ***
## Year2008             0.6071     0.1488    4.08  4.8e-05 ***
## Year2009             0.4708     0.1442    3.27  0.00112 **
## Year2010             0.5463     0.1426    3.83  0.00013 ***
## Year2011             0.3504     0.1403    2.50  0.01263 *
## Year2012             0.3572     0.1404    2.54  0.01107 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.757
## Multiple R-squared:  0.0521, Adjusted R-squared:  0.0412
## Convergence in 10 IRWLS iterations
##
## Robustness weights:
##  129 weights are ~= 1. The remaining 1368 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.275   0.885   0.938   0.920   0.983   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          6.68e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##                   GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.039  1           1.019
## Year             1.039 16           1.001
```



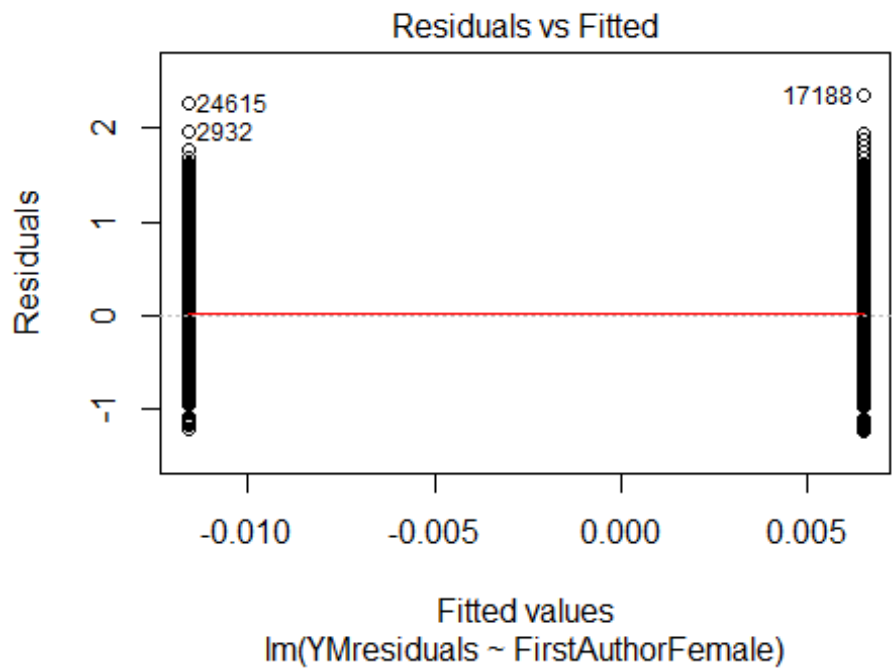

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.05182 -0.61382  0.00653  0.52078  2.39374
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)          0.3617     0.1309    2.76  0.00581 **
## LastAuthorFemale1    0.0754     0.0397    1.90  0.05771 .
## Year1997            -0.0645     0.1968   -0.33  0.74328
## Year1998             0.2595     0.2101    1.24  0.21698
## Year1999             0.1583     0.4231    0.37  0.70832
## Year2000            -0.0436     0.2099   -0.21  0.83563
## Year2001             0.1749     0.2286    0.76  0.44441
## Year2002            -0.0230     0.1595   -0.14  0.88526
## Year2003             0.5651     0.1490    3.79  0.00016 ***
## Year2004             0.5374     0.1538    3.49  0.00049 ***
## Year2005             0.5571     0.1497    3.72  0.00020 ***
## Year2006             0.4546     0.1476    3.08  0.00211 **
```

```
## Year2007            0.5304     0.1452    3.65  0.00027 ***
## Year2008            0.6148     0.1497    4.11  4.2e-05 ***
## Year2009            0.4736     0.1452    3.26  0.00113 **
## Year2010            0.5725     0.1430    4.01  6.5e-05 ***
## Year2011            0.3545     0.1412    2.51  0.01213 *
## Year2012            0.3688     0.1411    2.61  0.00907 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.765
## Multiple R-squared:  0.0374, Adjusted R-squared:  0.0263
## Convergence in 10 IRWLS iterations
##
## Robustness weights:
##  117 weights are ~= 1. The remaining 1380 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.306   0.880   0.942   0.921   0.982   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          6.68e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Sample size for the above analysis:  1497"
## [1] ""
## [1] ""
## [1] "####################################"
## [1] "Analysis of AJSC 3612"
## [1] "####################################"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 1135 1139 1279 1016 1275 1390 1163 1017 1136 1220 1214 1377 1547 1733 1753
## 2011 2012
## 1812 1809
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  878  844  949  739  855  819  992  853  956 1005 1031 1173 1316 1479 1511
## 2011 2012
```

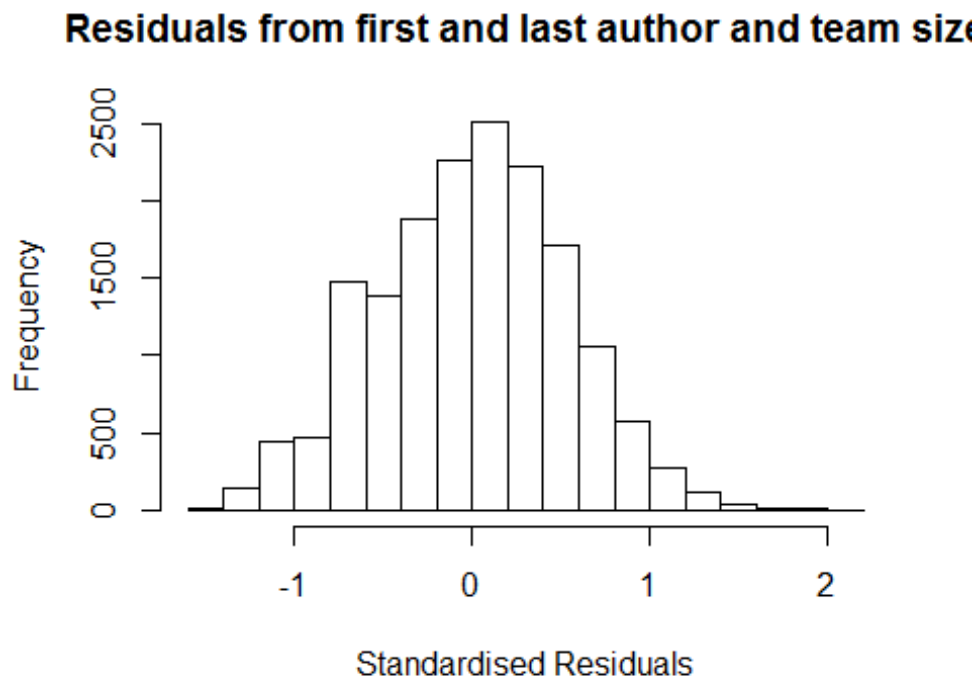
```
## 1543 1533
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  808  771  849  672  775  736  885  762  860  898  913 1051 1168 1317 1368
## 2011 2012
## 1397 1373
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
##  Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 170, df = 16, p-value <2e-16
```




```
##
##  Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 0.069, df = 1, p-value = 0.8
```

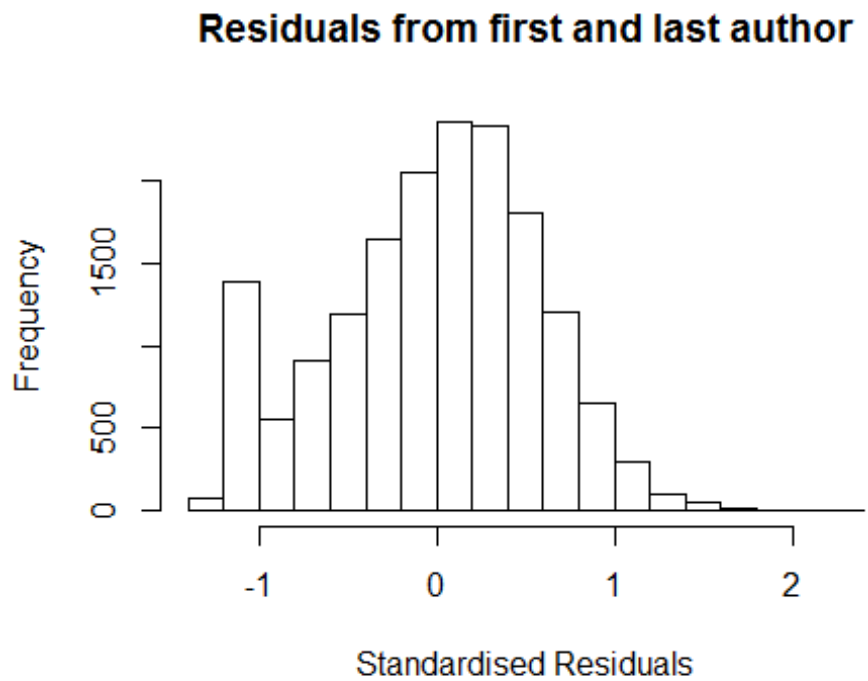



```
## [1] "Female first author team size 2018 geometric mean: 3.71869234699864"
## [1] "Male first author team size 2018 geometric mean: 4.04508832328515"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 250000, p-value = 0.007
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 3.5330789802775"
## [1] "Male last author team size 2018 geometric mean: 4.13193022022208"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 220000, p-value = 3e-06
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.089  1           1.043
## LastAuthorFemale  1.084  1           1.041
## UniqueAuthors     1.062  4           1.008
## Year              1.081 16           1.002
```



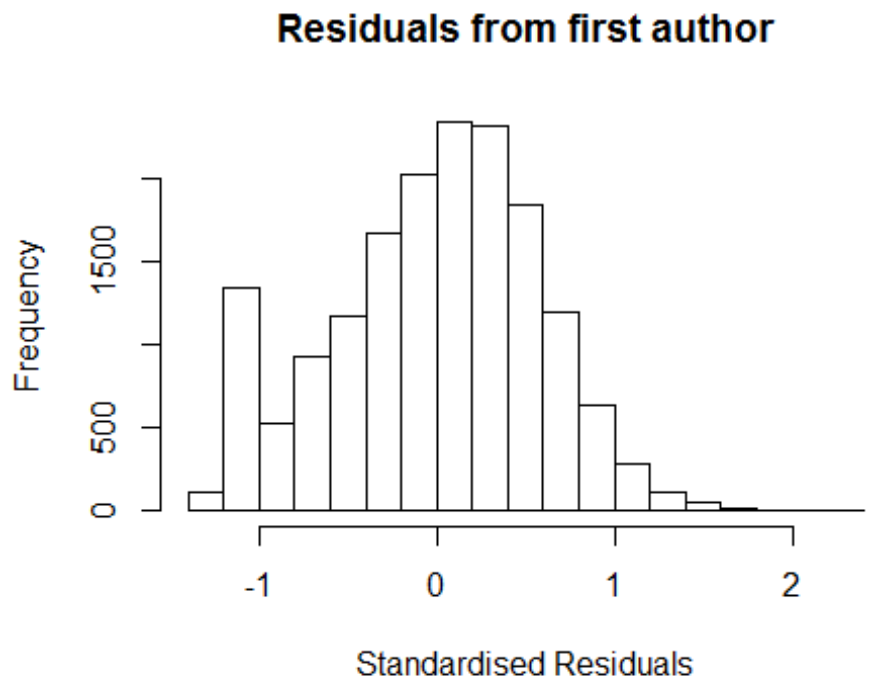

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
UniqueAuthors +
##     Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##     k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.4500 -0.3739  0.0173  0.3639  2.1354
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)         0.724390   0.024779   29.23  < 2e-16 ***
## FirstAuthorFemale1  0.000378   0.009301    0.04  0.96756
## LastAuthorFemale1  -0.028768   0.009715   -2.96  0.00307 **
## UniqueAuthors2      0.382495   0.016727   22.87  < 2e-16 ***
## UniqueAuthors3      0.518896   0.016242   31.95  < 2e-16 ***
## UniqueAuthors4      0.626086   0.016301   38.41  < 2e-16 ***
## UniqueAuthors5      0.745167   0.014895   50.03  < 2e-16 ***
## Year1997            0.000366   0.030956    0.01  0.99056
## Year1998           -0.038463   0.030375   -1.27  0.20543
## Year1999            0.008020   0.030966    0.26  0.79563
```

```
## Year2000           -0.061918   0.029569   -2.09  0.03627 *
## Year2001           -0.047025   0.030776   -1.53  0.12654
## Year2002           -0.036603   0.028935   -1.26  0.20589
## Year2003           -0.019909   0.029232   -0.68  0.49584
## Year2004           -0.062328   0.028411   -2.19  0.02826 *
## Year2005           -0.020334   0.029033   -0.70  0.48370
## Year2006           -0.030037   0.027815   -1.08  0.28020
## Year2007           -0.073134   0.027043   -2.70  0.00685 **
## Year2008           -0.090155   0.027480   -3.28  0.00104 **
## Year2009           -0.102313   0.027776   -3.68  0.00023 ***
## Year2010           -0.102818   0.027088   -3.80  0.00015 ***
## Year2011           -0.109041   0.027533   -3.96  7.5e-05 ***
## Year2012           -0.138755   0.028001   -4.96  7.3e-07 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.522
## Multiple R-squared:  0.189,  Adjusted R-squared:  0.187
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
##  1317 weights are ~= 1. The remaining 15286 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.057   0.864   0.946   0.905   0.985   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          6.02e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.050  1           1.025
## LastAuthorFemale  1.042  1           1.021
## Year              1.016 16           1.001
```



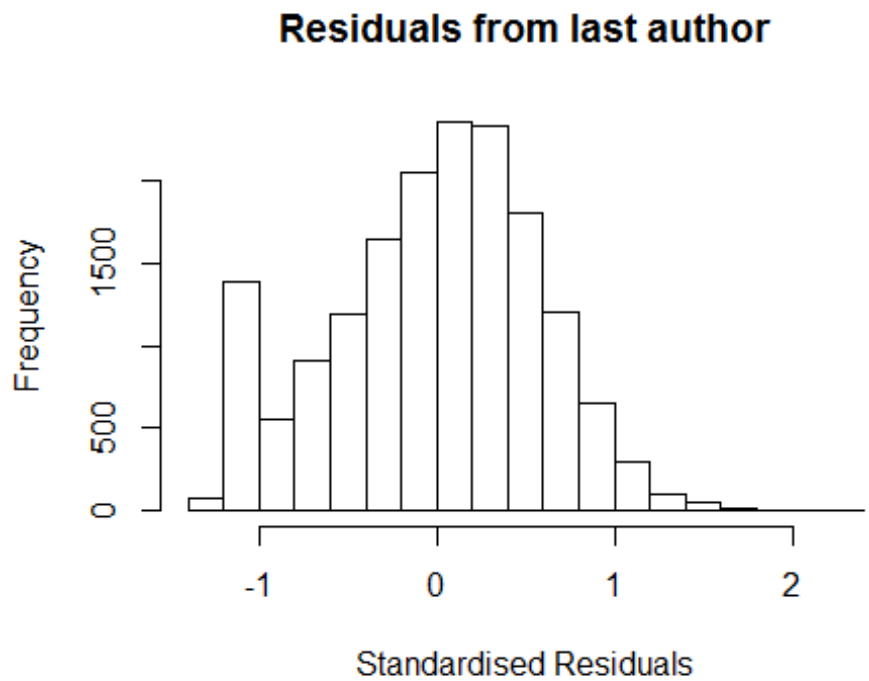

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.2485 -0.3932  0.0435  0.3968  2.3981
##
## Coefficients:
##                      Estimate Std. Error t value Pr(>|t|)
## (Intercept)          1.122251   0.026296   42.68  < 2e-16 ***
## FirstAuthorFemale1   0.002476   0.010072    0.25  0.80580
## LastAuthorFemale1   -0.067776   0.010562   -6.42  1.4e-10 ***
## Year1997             0.053898   0.035390    1.52  0.12778
## Year1998            -0.000440   0.034561   -0.01  0.98984
## Year1999             0.060907   0.035067    1.74  0.08243 .
## Year2000            -0.018425   0.033612   -0.55  0.58359
## Year2001             0.038005   0.035031    1.08  0.27798
## Year2002             0.066473   0.032425    2.05  0.04038 *
## Year2003             0.097894   0.032726    2.99  0.00278 **
## Year2004             0.056870   0.032164    1.77  0.07706 .
## Year2005             0.123799   0.032555    3.80  0.00014 ***
```

```
## Year2006            0.120656   0.030852    3.91  9.2e-05 ***
## Year2007            0.056736   0.030513    1.86  0.06299 .
## Year2008            0.033382   0.031288    1.07  0.28602
## Year2009            0.006372   0.031770    0.20  0.84103
## Year2010            0.011406   0.031173    0.37  0.71445
## Year2011            0.016820   0.031275    0.54  0.59071
## Year2012            0.000477   0.031989    0.01  0.98810
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.572
## Multiple R-squared:  0.00745,    Adjusted R-squared:  0.00637
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
##  1307 weights are ~= 1. The remaining 15296 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0395  0.8600  0.9480  0.9040  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          6.02e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.012  1           1.006
## Year              1.012 16           1.000
```




```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.2306 -0.3909  0.0435  0.3983  2.3445
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.11e+00   2.62e-02   42.33  < 2e-16 ***
## FirstAuthorFemale1 -1.69e-02   1.00e-02   -1.69  0.09158 .
## Year1997            5.31e-02   3.53e-02    1.50  0.13306
## Year1998            4.85e-05   3.45e-02    0.00  0.99888
## Year1999            6.09e-02   3.50e-02    1.74  0.08191 .
## Year2000           -1.69e-02   3.36e-02   -0.50  0.61576
## Year2001            3.74e-02   3.50e-02    1.07  0.28575
## Year2002            6.48e-02   3.24e-02    2.00  0.04593 *
## Year2003            9.48e-02   3.27e-02    2.89  0.00380 **
## Year2004            5.67e-02   3.22e-02    1.76  0.07802 .
## Year2005            1.21e-01   3.25e-02    3.71  0.00020 ***
## Year2006            1.20e-01   3.09e-02    3.87  0.00011 ***
```

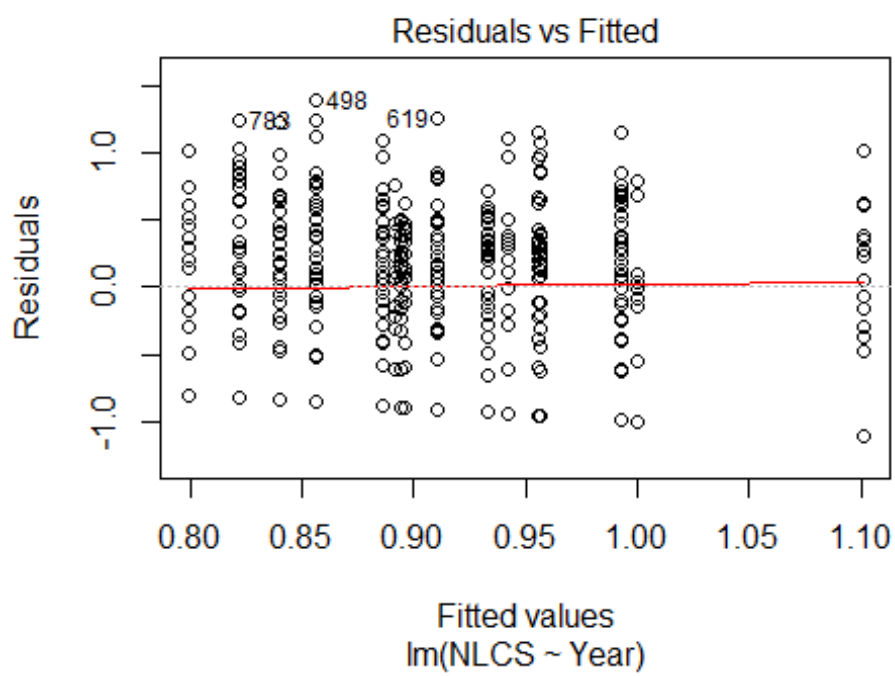
```
## Year2007            5.41e-02   3.05e-02    1.77  0.07618 .
## Year2008            3.17e-02   3.13e-02    1.01  0.31025
## Year2009            3.54e-03   3.18e-02    0.11  0.91131
## Year2010            9.12e-03   3.12e-02    0.29  0.77013
## Year2011            1.28e-02   3.13e-02    0.41  0.68323
## Year2012           -2.99e-03   3.20e-02   -0.09  0.92564
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.572
## Multiple R-squared:  0.00488,    Adjusted R-squared:  0.00386
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
##  1326 weights are ~= 1. The remaining 15277 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0551  0.8590  0.9470  0.9040  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          6.02e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##                   GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.004  1           1.002
## Year             1.004 16           1.000
```




```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.2467 -0.3933  0.0436  0.3967  2.3967
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.122814   0.026147   42.94  < 2e-16 ***
## LastAuthorFemale1 -0.067018   0.010427   -6.43  1.3e-10 ***
## Year1997            0.053869   0.035389    1.52  0.12797
## Year1998           -0.000513   0.034560   -0.01  0.98816
## Year1999            0.060811   0.035063    1.73  0.08288 .
## Year2000           -0.018441   0.033613   -0.55  0.58327
## Year2001            0.037903   0.035024    1.08  0.27919
## Year2002            0.066559   0.032425    2.05  0.04011 *
## Year2003            0.098002   0.032731    2.99  0.00276 **
## Year2004            0.056937   0.032165    1.77  0.07672 .
## Year2005            0.123889   0.032558    3.81  0.00014 ***
## Year2006            0.120740   0.030854    3.91  9.1e-05 ***
```

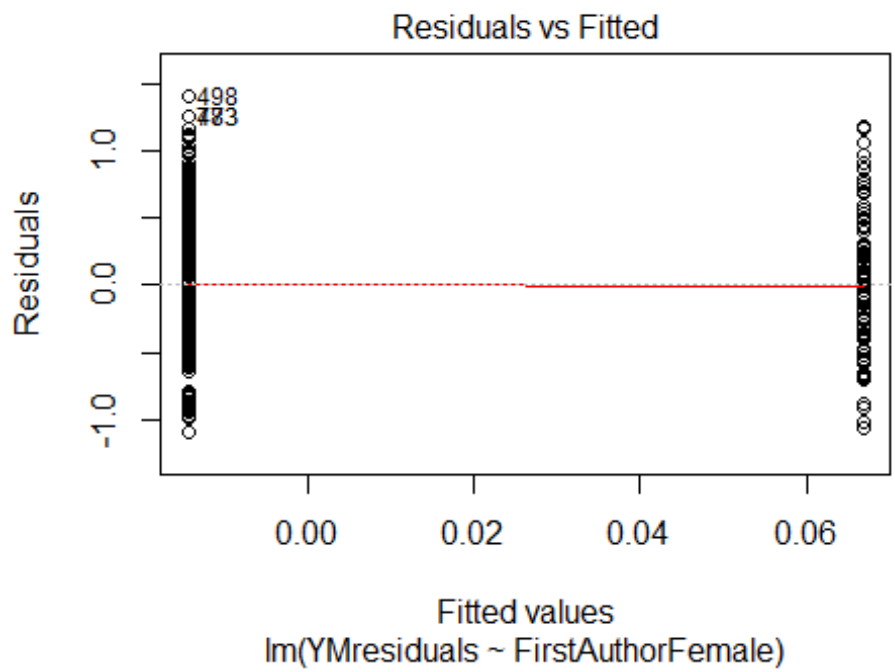
```
## Year2007           0.056880   0.030508    1.86  0.06228 .
## Year2008           0.033528   0.031289    1.07  0.28394
## Year2009           0.006586   0.031748    0.21  0.83566
## Year2010           0.011592   0.031168    0.37  0.70996
## Year2011           0.017086   0.031269    0.55  0.58478
## Year2012           0.000692   0.031974    0.02  0.98274
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.572
## Multiple R-squared:  0.00745,    Adjusted R-squared:  0.00643
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
##  1308 weights are ~= 1. The remaining 15295 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0399  0.8600  0.9480  0.9040  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          6.02e-06          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Sample size for the above analysis:  16603"
## [1] ""
## [1] ""
## [1] "####################################"
## [1] "Analysis of AJSC 3613"
## [1] "####################################"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##   30   42   26   25   34   23   19   22   25   42   25   31   62   46   73
## 2011 2012
##   58   68
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##   20   32   23   13   19   11   16   20   24   37   22   30   57   44   57
## 2011 2012
```

```
##   54   62
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##   18   30   22   13   19   11   15   16   22   32   20   24   55   41   55
## 2011 2012
##   48   57
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
##  Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 21, df = 16, p-value = 0.2
```

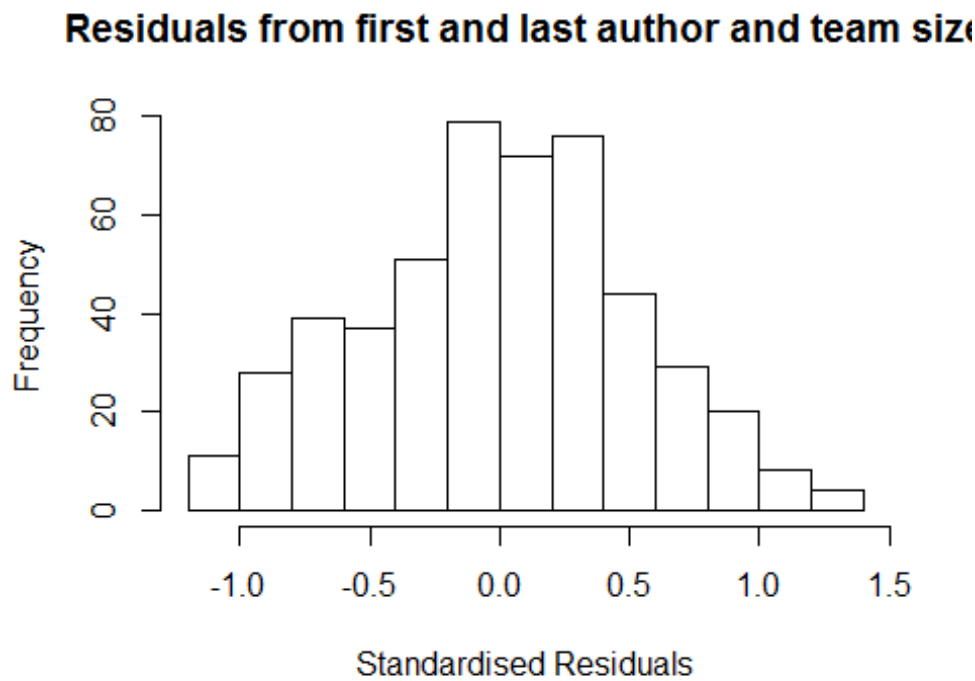



```
##
##  Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 0.034, df = 1, p-value = 0.9

## [1] "Female first author team size 2018 geometric mean: 4.14716627439691"
## [1] "Male first author team size 2018 geometric mean: 4.31096697544756"

## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties
```

```
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 150, p-value = 0.9
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 4.60577935159691"
## [1] "Male last author team size 2018 geometric mean: 4.24371199288181"

## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties
```

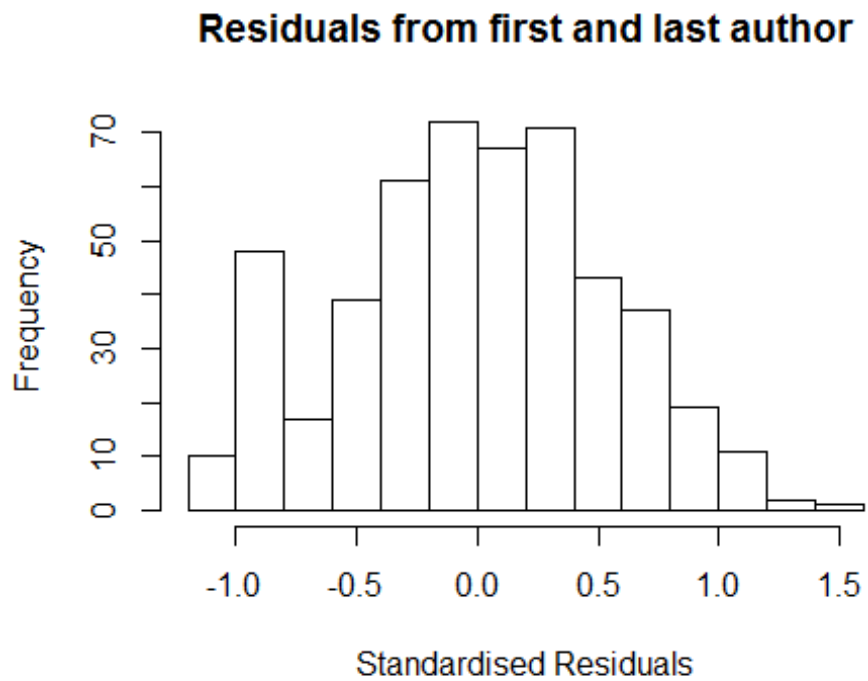



```
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 86, p-value = 0.7
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##                     GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.286  1           1.134
## LastAuthorFemale  1.211  1           1.100
## UniqueAuthors     1.798  4           1.076
## Year              2.295 16           1.026
```



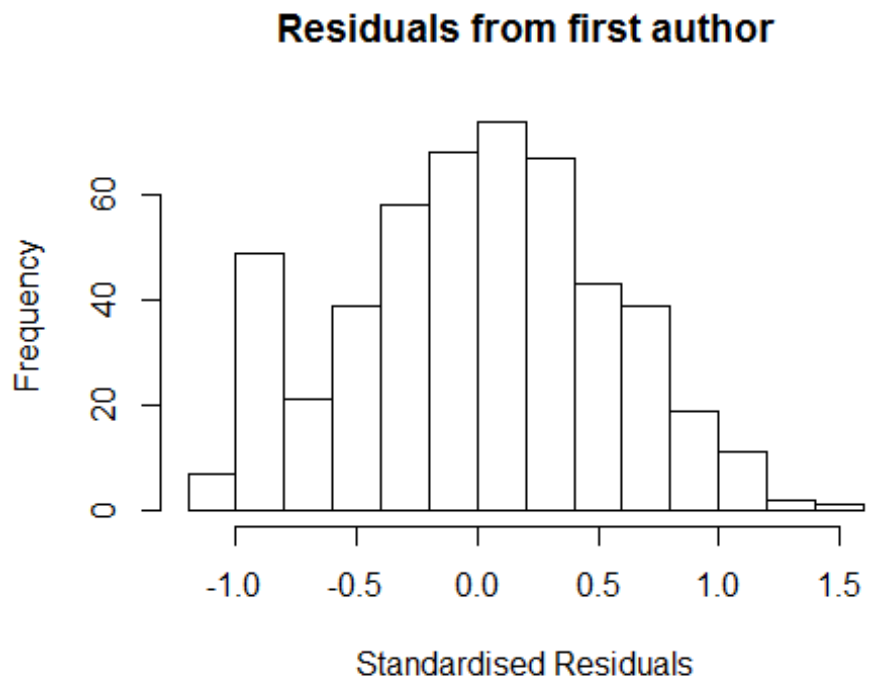

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
UniqueAuthors +
##     Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##     k.max = 1000))
##  \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.19744 -0.35398  0.00766  0.34308  1.32244
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         0.71765    0.15086    4.76  2.6e-06 ***
## FirstAuthorFemale1  0.04638    0.07383    0.63  0.53015
## LastAuthorFemale1  -0.09359    0.07319   -1.28  0.20159
## UniqueAuthors2      0.17137    0.07361    2.33  0.02033 *
## UniqueAuthors3      0.19761    0.08101    2.44  0.01508 *
## UniqueAuthors4      0.30332    0.08353    3.63  0.00031 ***
## UniqueAuthors5      0.39322    0.08831    4.45  1.1e-05 ***
## Year1997            0.08656    0.19403    0.45  0.65570
## Year1998           -0.05238    0.19871   -0.26  0.79221
## Year1999            0.03443    0.18204    0.19  0.85009
```

```
## Year2000            0.03684    0.20304    0.18  0.85608
## Year2001            0.11068    0.21472    0.52  0.60647
## Year2002            0.23429    0.22492    1.04  0.29811
## Year2003            0.00242    0.19213    0.01  0.98994
## Year2004           -0.00269    0.16205   -0.02  0.98678
## Year2005            0.00793    0.16513    0.05  0.96174
## Year2006           -0.00279    0.16848   -0.02  0.98679
## Year2007            0.14103    0.20013    0.70  0.48135
## Year2008           -0.03658    0.17112   -0.21  0.83083
## Year2009           -0.05630    0.16802   -0.34  0.73770
## Year2010           -0.01742    0.16256   -0.11  0.91471
## Year2011            0.08878    0.15903    0.56  0.57694
## Year2012           -0.14747    0.16554   -0.89  0.37347
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.522
## Multiple R-squared:  0.0728, Adjusted R-squared:  0.0299
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
##  40 weights are ~= 1. The remaining 458 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.501   0.853   0.951   0.908   0.985   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          2.01e-04          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.221  1           1.105
## LastAuthorFemale  1.203  1           1.097
## Year              1.318 16           1.009
```



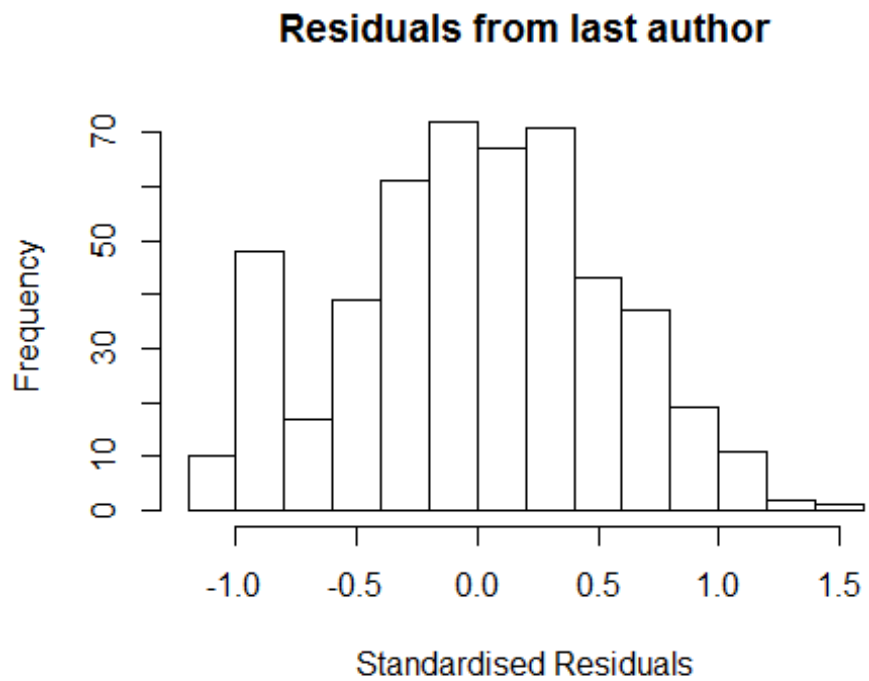

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.09404 -0.38119  0.00703  0.37478  1.46917
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         0.88968    0.15615    5.70  2.1e-08 ***
## FirstAuthorFemale1  0.08476    0.07262    1.17     0.24
## LastAuthorFemale1  -0.04982    0.07340   -0.68     0.50
## Year1997            0.04830    0.20088    0.24     0.81
## Year1998           -0.08468    0.20712   -0.41     0.68
## Year1999           -0.01344    0.19169   -0.07     0.94
## Year2000           -0.03915    0.19970   -0.20     0.84
## Year2001            0.10791    0.20861    0.52     0.61
## Year2002            0.20437    0.22161    0.92     0.36
## Year2003            0.00365    0.20005    0.02     0.99
## Year2004            0.01112    0.17001    0.07     0.95
## Year2005            0.01142    0.17485    0.07     0.95
```

```
## Year2006            0.03894    0.17857    0.22     0.83
## Year2007            0.11158    0.20968    0.53     0.59
## Year2008           -0.06703    0.18163   -0.37     0.71
## Year2009           -0.04087    0.17835   -0.23     0.82
## Year2010           -0.00282    0.17373   -0.02     0.99
## Year2011            0.13193    0.16981    0.78     0.44
## Year2012           -0.08235    0.17317   -0.48     0.63
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.555
## Multiple R-squared:  0.0213, Adjusted R-squared:  -0.0154
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
##  39 weights are ~= 1. The remaining 459 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.464   0.868   0.954   0.915   0.984   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          2.01e-04          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.155  1           1.075
## Year              1.155 16           1.005
```




```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.08830 -0.38349  0.00522  0.37000  1.42486
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         0.87997    0.15247    5.77  1.4e-08 ***
## FirstAuthorFemale1  0.07682    0.07109    1.08     0.28
## Year1997            0.05242    0.19898    0.26     0.79
## Year1998           -0.07805    0.20526   -0.38     0.70
## Year1999           -0.01261    0.18825   -0.07     0.95
## Year2000           -0.03165    0.19692   -0.16     0.87
## Year2001            0.11368    0.20848    0.55     0.59
## Year2002            0.20721    0.21976    0.94     0.35
## Year2003           -0.01250    0.20145   -0.06     0.95
## Year2004            0.01719    0.16737    0.10     0.92
## Year2005            0.01280    0.17317    0.07     0.94
## Year2006            0.04373    0.17635    0.25     0.80
```

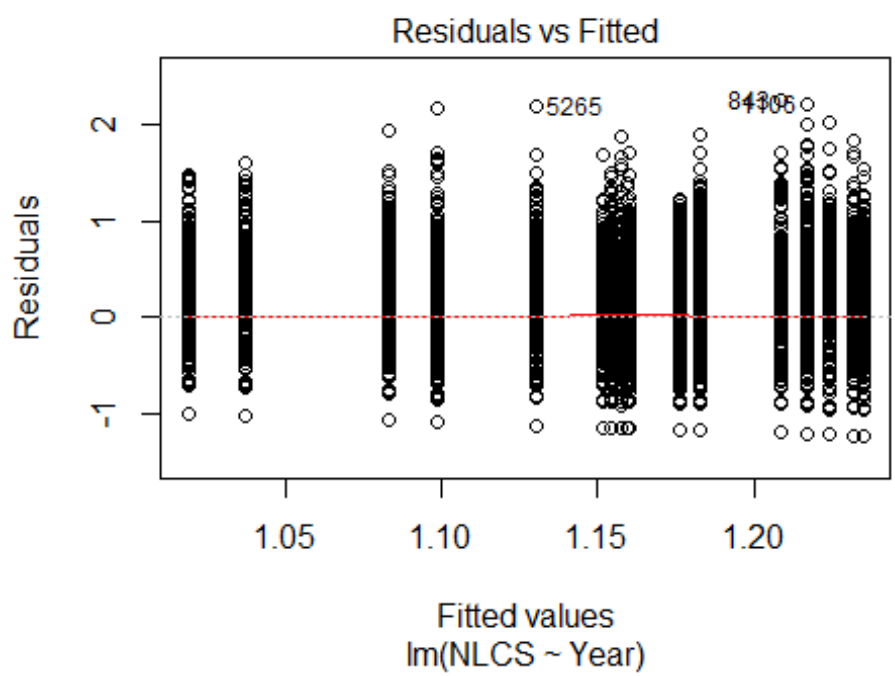
```
## Year2007            0.11613    0.20645    0.56     0.57
## Year2008           -0.06284    0.17939   -0.35     0.73
## Year2009           -0.03707    0.17620   -0.21     0.83
## Year2010            0.00463    0.17083    0.03     0.98
## Year2011            0.13150    0.16816    0.78     0.43
## Year2012           -0.08348    0.17154   -0.49     0.63
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.554
## Multiple R-squared:  0.0201, Adjusted R-squared:  -0.0146
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
##  42 weights are ~= 1. The remaining 456 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.488   0.863   0.952   0.913   0.984   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          2.01e-04          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##                   GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.142  1           1.069
## Year             1.142 16           1.004
```



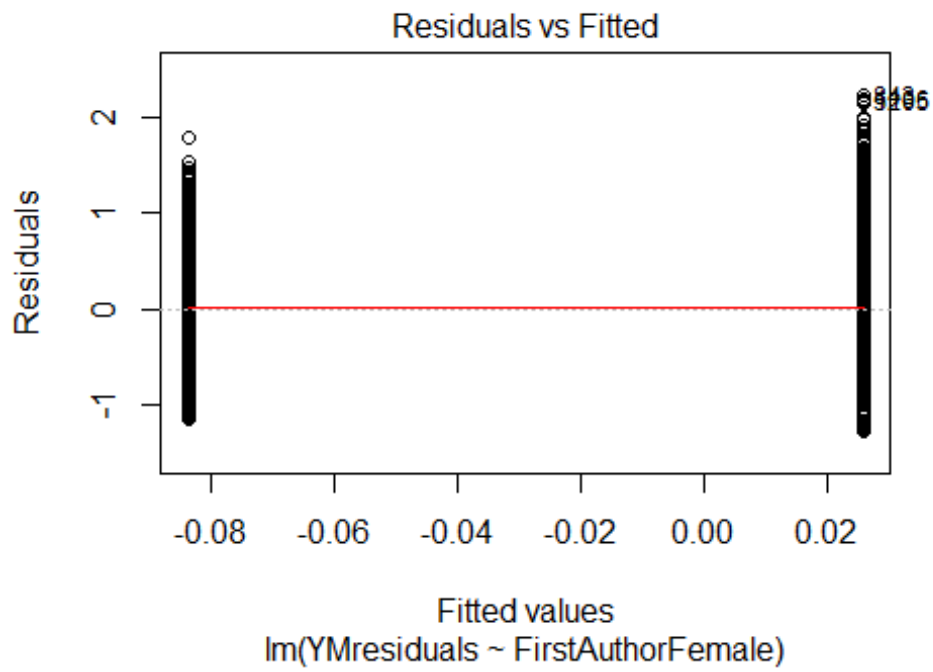

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.11267 -0.38621  0.00849  0.37245  1.43272
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         0.89750    0.15441    5.81  1.1e-08 ***
## LastAuthorFemale1 -0.03585    0.07153   -0.50     0.62
## Year1997            0.05365    0.19794    0.27     0.79
## Year1998           -0.08081    0.20659   -0.39     0.70
## Year1999           -0.00600    0.18585   -0.03     0.97
## Year2000           -0.03466    0.19696   -0.18     0.86
## Year2001            0.11685    0.20637    0.57     0.57
## Year2002            0.21518    0.21870    0.98     0.33
## Year2003            0.01474    0.20322    0.07     0.94
## Year2004            0.00643    0.16846    0.04     0.97
## Year2005            0.00737    0.17357    0.04     0.97
## Year2006            0.03705    0.17702    0.21     0.83
```

```
## Year2007           0.11248    0.20740    0.54     0.59
## Year2008          -0.05236    0.17978   -0.29     0.77
## Year2009          -0.03832    0.17632   -0.22     0.83
## Year2010           0.00823    0.17297    0.05     0.96
## Year2011           0.12631    0.16790    0.75     0.45
## Year2012          -0.07546    0.17198   -0.44     0.66
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.554
## Multiple R-squared:  0.0182, Adjusted R-squared:  -0.0166
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
##  42 weights are ~= 1. The remaining 456 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.483   0.868   0.951   0.913   0.984   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          2.01e-04          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Sample size for the above analysis:  498"
## [1] ""
## [1] ""
## [1] "####################################"
## [1] "Analysis of AJSC 3614"
## [1] "####################################"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  788  849  853  799  822  792  819  749  858  972 1168 1197 1081  723  689
## 2011 2012
##  760  752
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  405  455  505  458  381  289  530  476  571  612  768  789  720  438  433
## 2011 2012
```

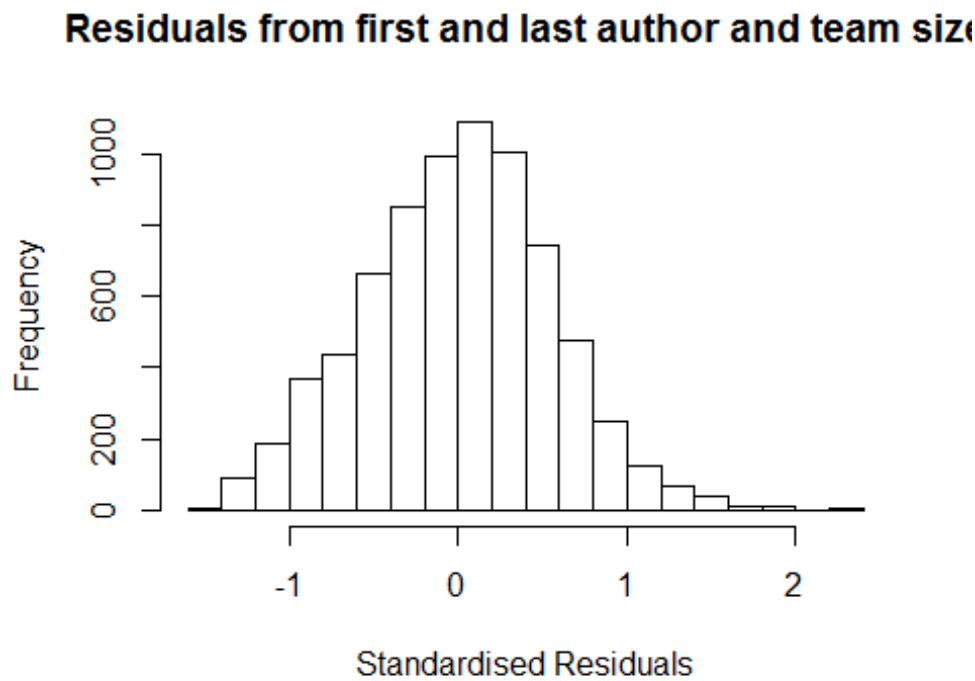
```
##  462  495
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  342  404  447  397  322  240  448  391  485  524  595  651  591  368  375
## 2011 2012
##  410  431
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
##  Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 130, df = 16, p-value <2e-16
```




```
##
##  Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 0.0011, df = 1, p-value = 1
```

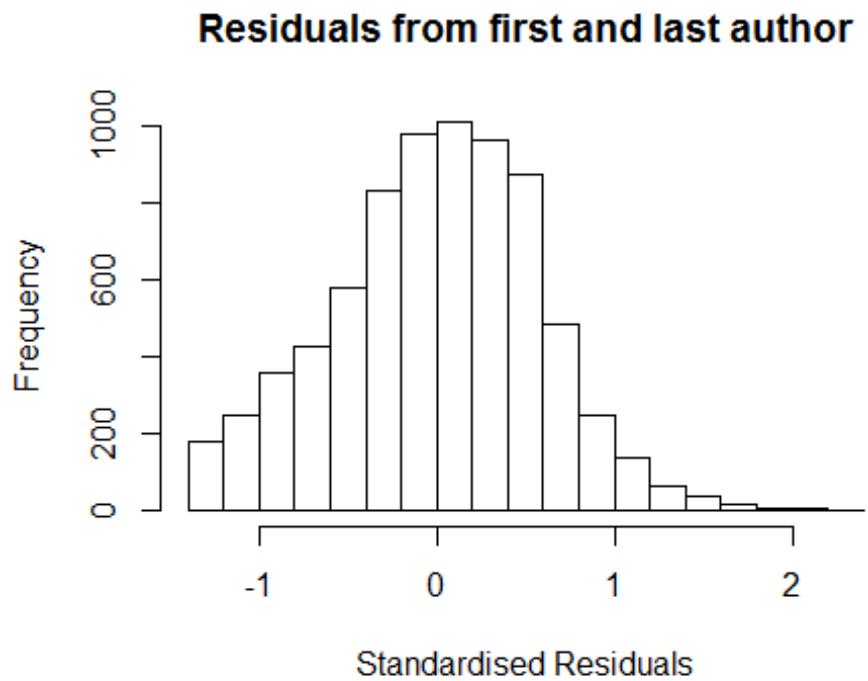



```
## [1] "Female first author team size 2018 geometric mean: 3.44250565955926"
## [1] "Male first author team size 2018 geometric mean: 4.00923965603041"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 29000, p-value = 0.08
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 2.76198953327526"
## [1] "Male last author team size 2018 geometric mean: 4.28553537853327"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 19000, p-value = 7e-08
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##                     GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.066  1           1.033
## LastAuthorFemale  1.081  1           1.040
## UniqueAuthors     1.095  4           1.011
## Year              1.098 16           1.003
```



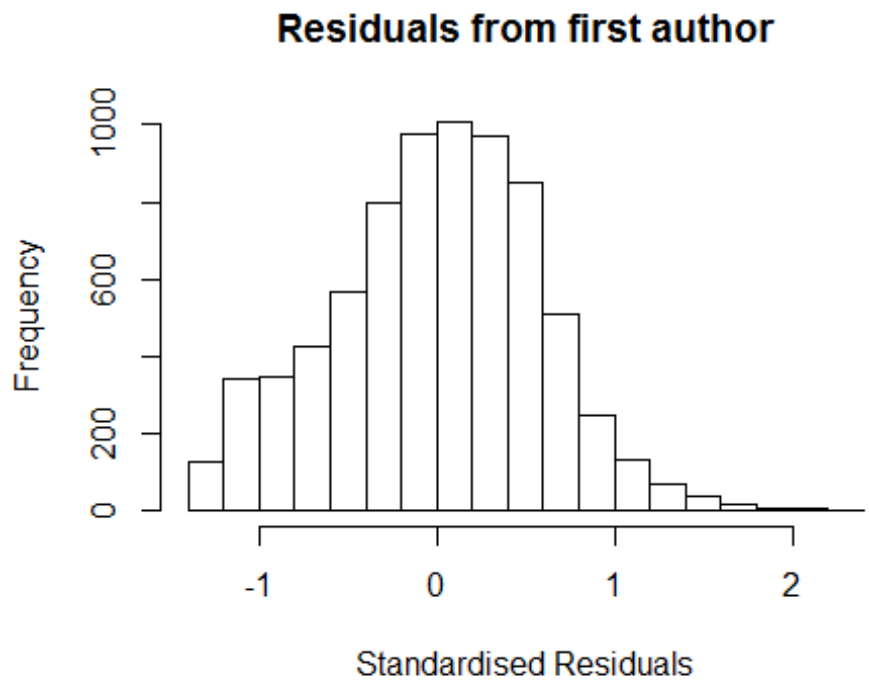

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
UniqueAuthors +
##     Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##     k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.4462 -0.3667  0.0186  0.3696  2.3658
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)          0.8386     0.0423   19.81  < 2e-16 ***
## FirstAuthorFemale1  -0.0855     0.0156   -5.47  4.7e-08 ***
## LastAuthorFemale1   -0.1313     0.0180   -7.30  3.3e-13 ***
## UniqueAuthors2       0.2889     0.0289    9.98  < 2e-16 ***
## UniqueAuthors3       0.3482     0.0270   12.92  < 2e-16 ***
## UniqueAuthors4       0.4338     0.0264   16.41  < 2e-16 ***
## UniqueAuthors5       0.5453     0.0244   22.39  < 2e-16 ***
## Year1997             0.0595     0.0484    1.23   0.2190
## Year1998             0.0196     0.0459    0.43   0.6686
## Year1999             0.0307     0.0467    0.66   0.5107
```

```
## Year2000             0.0526     0.0522    1.01   0.3135
## Year2001            -0.0669     0.0593   -1.13   0.2594
## Year2002            -0.0735     0.0448   -1.64   0.1005
## Year2003             0.0624     0.0450    1.39   0.1655
## Year2004            -0.0297     0.0427   -0.70   0.4870
## Year2005            -0.0120     0.0418   -0.29   0.7750
## Year2006            -0.0119     0.0418   -0.29   0.7755
## Year2007            -0.0158     0.0406   -0.39   0.6974
## Year2008            -0.0378     0.0411   -0.92   0.3571
## Year2009            -0.0477     0.0443   -1.08   0.2815
## Year2010            -0.1378     0.0470   -2.93   0.0034 **
## Year2011            -0.1397     0.0444   -3.15   0.0016 **
## Year2012            -0.1752     0.0440   -3.98  7.0e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.55
## Multiple R-squared:  0.114,  Adjusted R-squared:  0.111
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  604 weights are ~= 1. The remaining 6817 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0251  0.8690  0.9510  0.9060  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.35e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.030  1           1.015
## LastAuthorFemale  1.028  1           1.014
## Year              1.019 16           1.001
```



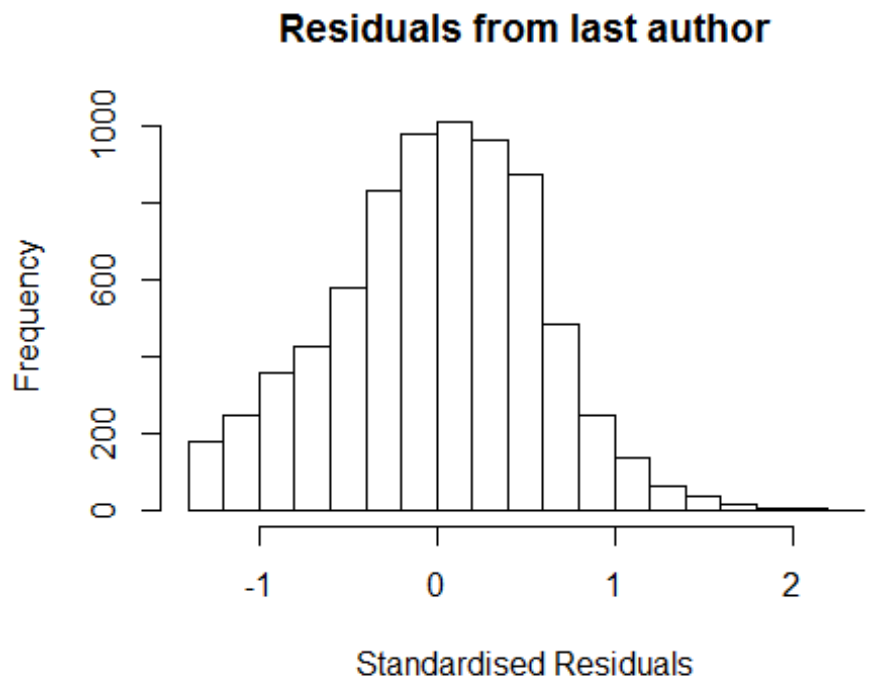

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.2925 -0.3783  0.0263  0.4013  2.2313
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)          1.22968    0.03566   34.49  < 2e-16 ***
## FirstAuthorFemale1 -0.07260    0.01648   -4.40  1.1e-05 ***
## LastAuthorFemale1  -0.17700    0.01919   -9.23  < 2e-16 ***
## Year1997            0.03135    0.05019    0.62  0.53222
## Year1998            0.00064    0.04685    0.01  0.98910
## Year1999            0.04109    0.04616    0.89  0.37340
## Year2000            0.02783    0.05282    0.53  0.59823
## Year2001           -0.05680    0.05867   -0.97  0.33298
## Year2002           -0.10278    0.04603   -2.23  0.02559 *
## Year2003            0.06285    0.04594    1.37  0.17130
## Year2004           -0.02737    0.04329   -0.63  0.52725
## Year2005           -0.00954    0.04274   -0.22  0.82335
```

```
## Year2006           -0.01438    0.04259   -0.34  0.73564
## Year2007           -0.03373    0.04116   -0.82  0.41257
## Year2008           -0.03536    0.04179   -0.85  0.39750
## Year2009           -0.00560    0.04475   -0.13  0.90039
## Year2010           -0.10884    0.04807   -2.26  0.02359 *
## Year2011           -0.13918    0.04614   -3.02  0.00257 **
## Year2012           -0.17404    0.04522   -3.85  0.00012 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.572
## Multiple R-squared:  0.0313, Adjusted R-squared:  0.0289
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  570 weights are ~= 1. The remaining 6851 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0939  0.8690  0.9500  0.9060  0.9850  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.35e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.012  1           1.006
## Year              1.012 16           1.000
```




```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.2688 -0.3834  0.0236  0.4028  2.2537
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)          1.20732    0.03572   33.80  < 2e-16 ***
## FirstAuthorFemale1 -0.10539    0.01685   -6.25  4.2e-10 ***
## Year1997             0.03119    0.05045    0.62   0.5365
## Year1998             0.00308    0.04704    0.07   0.9479
## Year1999             0.04456    0.04619    0.96   0.3347
## Year2000             0.03454    0.05329    0.65   0.5168
## Year2001            -0.04517    0.05886   -0.77   0.4429
## Year2002            -0.10716    0.04648   -2.31   0.0212 *
## Year2003             0.06149    0.04633    1.33   0.1845
## Year2004            -0.02621    0.04342   -0.60   0.5462
## Year2005            -0.01041    0.04304   -0.24   0.8089
## Year2006            -0.01576    0.04272   -0.37   0.7123
```

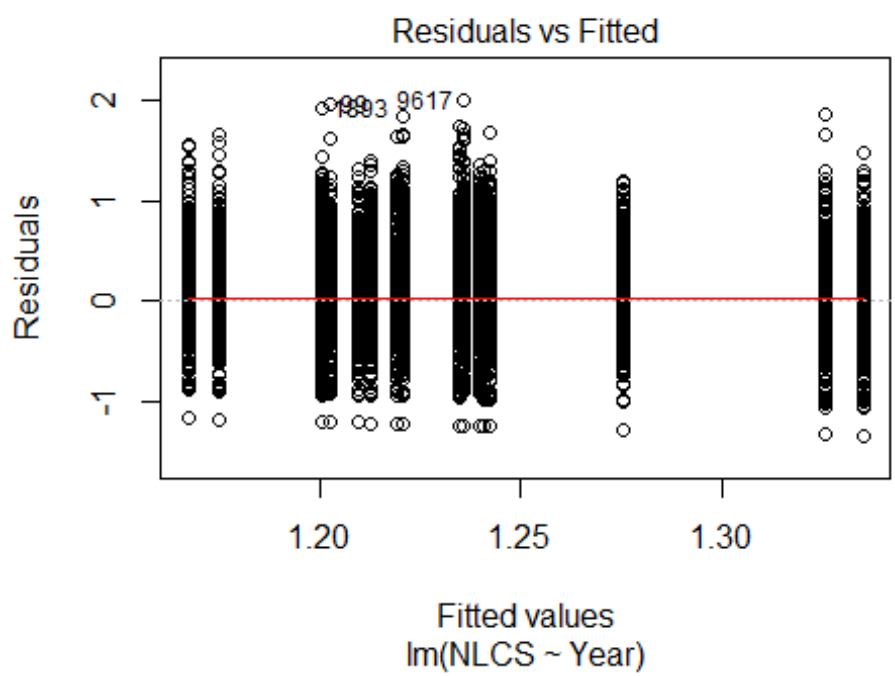
```
## Year2007           -0.03163    0.04141   -0.76   0.4449
## Year2008           -0.03627    0.04207   -0.86   0.3886
## Year2009           -0.01126    0.04512   -0.25   0.8029
## Year2010           -0.11095    0.04858   -2.28   0.0224 *
## Year2011           -0.14873    0.04661   -3.19   0.0014 **
## Year2012           -0.18019    0.04562   -3.95  7.9e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.573
## Multiple R-squared:  0.0185, Adjusted R-squared:  0.0163
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  610 weights are ~= 1. The remaining 6811 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0868  0.8660  0.9490  0.9040  0.9840  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.35e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##                   GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.009  1           1.004
## Year             1.009 16           1.000
```



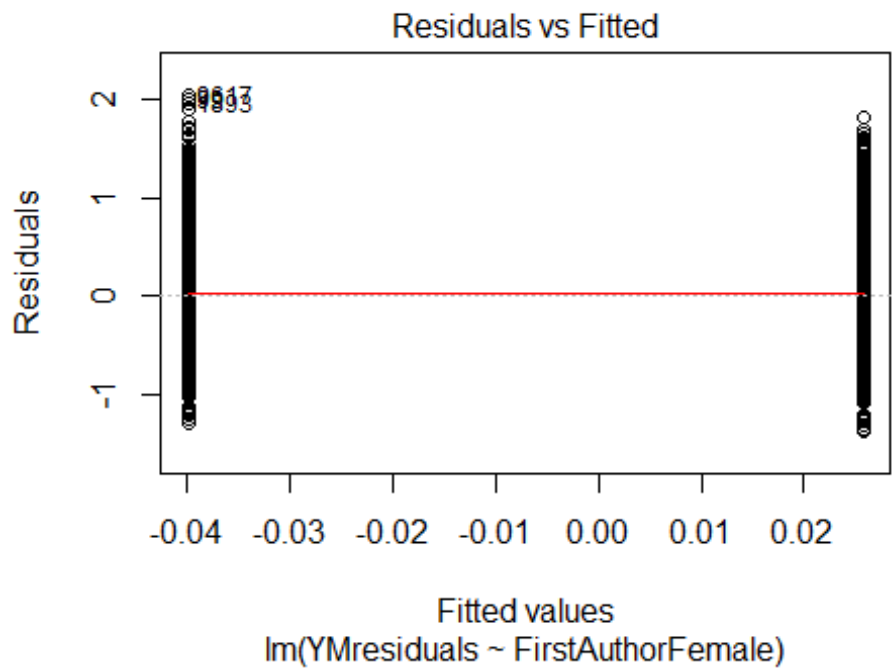

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.2784 -0.3827  0.0267  0.4013  2.2482
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.21278    0.03549   34.17   <2e-16 ***
## LastAuthorFemale1 -0.19442    0.01931  -10.07   <2e-16 ***
## Year1997            0.03387    0.05044    0.67   0.5019
## Year1998            0.00391    0.04692    0.08   0.9337
## Year1999            0.04313    0.04632    0.93   0.3518
## Year2000            0.03154    0.05301    0.59   0.5519
## Year2001           -0.05519    0.05878   -0.94   0.3478
## Year2002           -0.09997    0.04621   -2.16   0.0306 *
## Year2003            0.06565    0.04598    1.43   0.1534
## Year2004           -0.02271    0.04341   -0.52   0.6009
## Year2005           -0.00596    0.04283   -0.14   0.8894
## Year2006           -0.01012    0.04264   -0.24   0.8123
```

```
## Year2007          -0.03054    0.04121   -0.74   0.4588
## Year2008          -0.03248    0.04192   -0.77   0.4384
## Year2009          -0.00299    0.04484   -0.07   0.9469
## Year2010          -0.10864    0.04824   -2.25   0.0243 *
## Year2011          -0.14122    0.04639   -3.04   0.0023 **
## Year2012          -0.17677    0.04547   -3.89   0.0001 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.572
## Multiple R-squared:  0.0284, Adjusted R-squared:  0.0262
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  605 weights are ~= 1. The remaining 6816 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0874  0.8650  0.9490  0.9050  0.9840  0.9990
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.35e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Sample size for the above analysis:  7421"
## [1] ""
## [1] ""
## [1] "####################################"
## [1] "Analysis of AJSC 3616"
## [1] "####################################"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  455  473  395  406  500  473  423  445  441  429  490  561  580  639  678
## 2011 2012
##  690  737
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  357  381  326  332  388  305  369  390  380  363  423  492  508  545  568
## 2011 2012
```

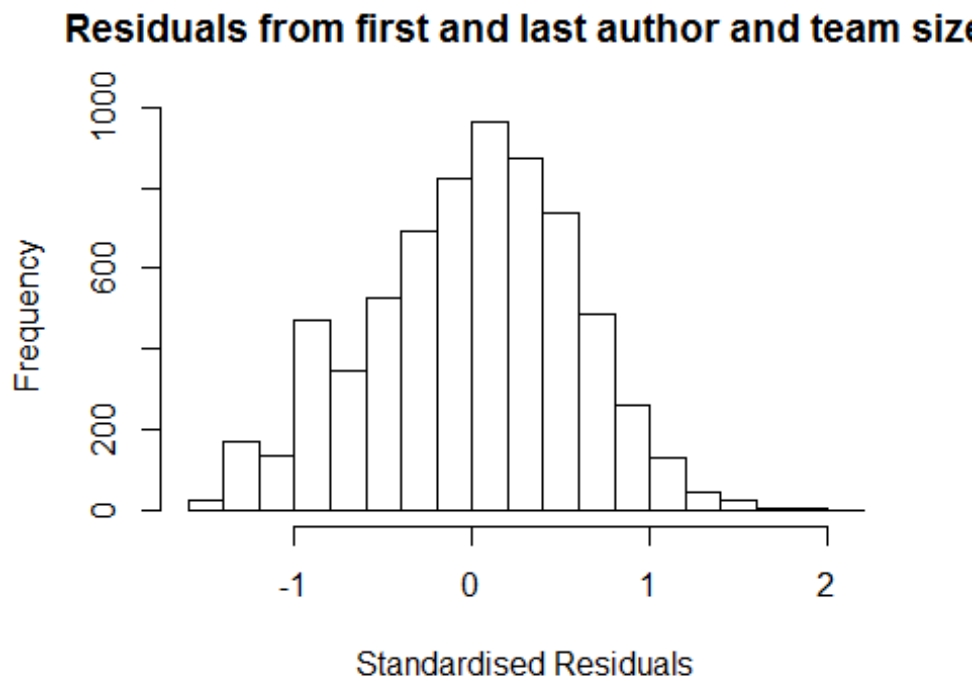
```
##  604  624
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##  334  358  299  310  362  279  331  363  351  327  382  451  466  491  514
## 2011 2012
##  542  552
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
##  Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 42, df = 16, p-value = 4e-04
```




```
##
##  Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 17, df = 1, p-value = 4e-05
```

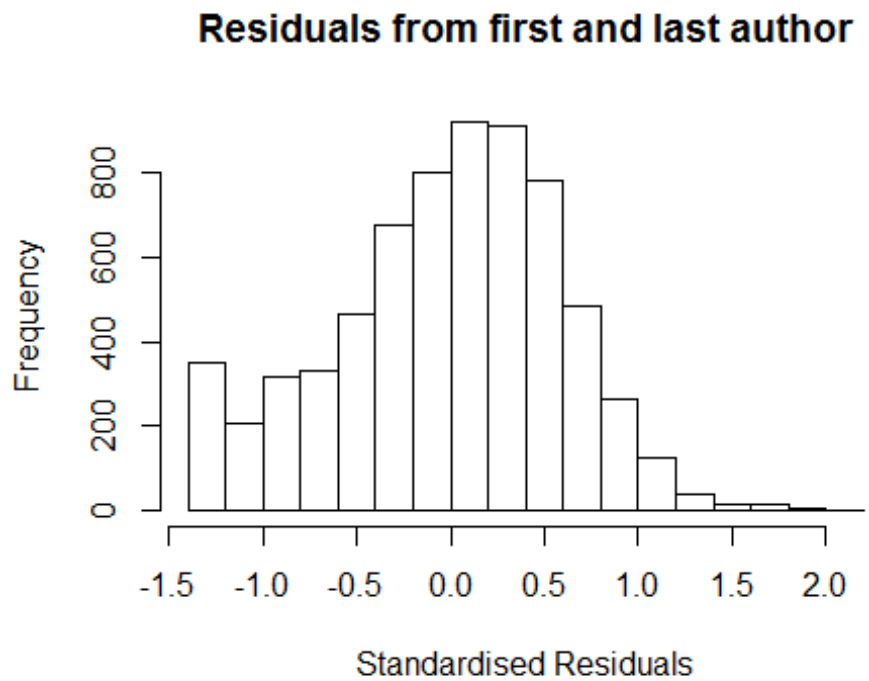



```
## [1] "Female first author team size 2018 geometric mean: 2.71427358835732"
## [1] "Male first author team size 2018 geometric mean: 2.70547378916192"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 37000, p-value = 1
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 2.54348985476112"
## [1] "Male last author team size 2018 geometric mean: 3.01855188986215"
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 36000, p-value = 6e-04
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##                     GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.085  1           1.042
## LastAuthorFemale  1.097  1           1.048
## UniqueAuthors     1.101  4           1.012
## Year              1.109 16           1.003
```



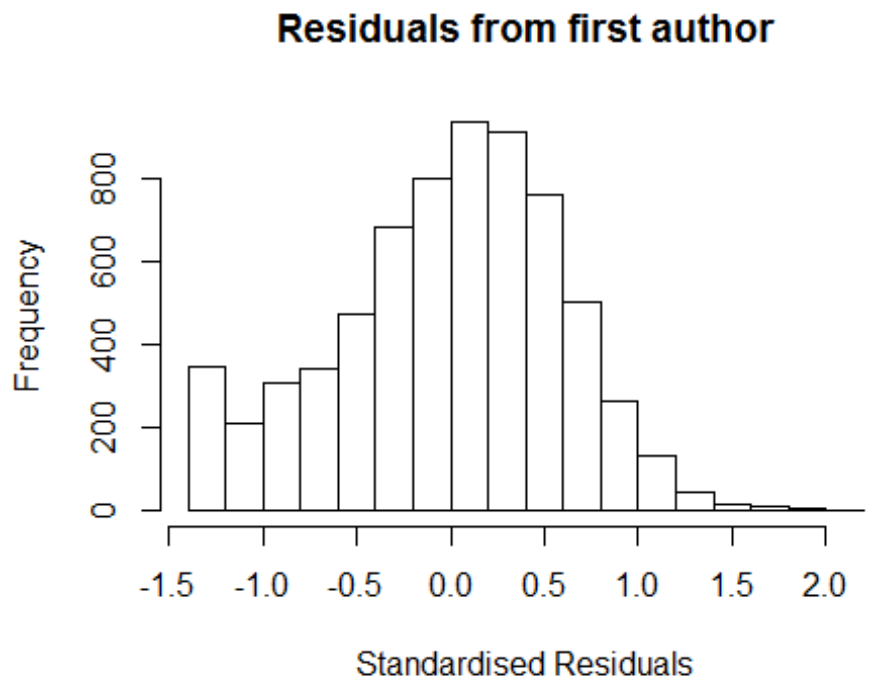

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
UniqueAuthors +
##     Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##     k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.5935 -0.3968  0.0332  0.4029  2.0160
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)          0.92163    0.03847   23.96   <2e-16 ***
## FirstAuthorFemale1   0.04493    0.01602    2.80   0.0050 **
## LastAuthorFemale1    0.04880    0.01559    3.13   0.0018 **
## UniqueAuthors2       0.30546    0.02229   13.71   <2e-16 ***
## UniqueAuthors3       0.36592    0.02395   15.28   <2e-16 ***
## UniqueAuthors4       0.40720    0.02652   15.35   <2e-16 ***
## UniqueAuthors5       0.49940    0.02566   19.46   <2e-16 ***
## Year1997             0.00202    0.04747    0.04   0.9660
## Year1998             0.01374    0.05065    0.27   0.7861
## Year1999            -0.02826    0.04798   -0.59   0.5559
```

```
## Year2000            0.04126    0.04495    0.92   0.3587
## Year2001           -0.01325    0.04624   -0.29   0.7745
## Year2002            0.07875    0.04638    1.70   0.0896 .
## Year2003           -0.00168    0.04523   -0.04   0.9703
## Year2004           -0.01293    0.04964   -0.26   0.7945
## Year2005            0.06811    0.04662    1.46   0.1441
## Year2006           -0.05356    0.04567   -1.17   0.2410
## Year2007           -0.08722    0.04530   -1.93   0.0542 .
## Year2008           -0.04762    0.04440   -1.07   0.2836
## Year2009           -0.11033    0.04381   -2.52   0.0118 *
## Year2010           -0.06540    0.04410   -1.48   0.1381
## Year2011           -0.02111    0.04410   -0.48   0.6321
## Year2012           -0.05589    0.04599   -1.22   0.2243
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.579
## Multiple R-squared:  0.0867, Adjusted R-squared:  0.0837
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
##  558 weights are ~= 1. The remaining 6154 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.200   0.865   0.948   0.906   0.985   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.49e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.072  1           1.035
## LastAuthorFemale  1.079  1           1.039
## Year              1.020 16           1.001
```



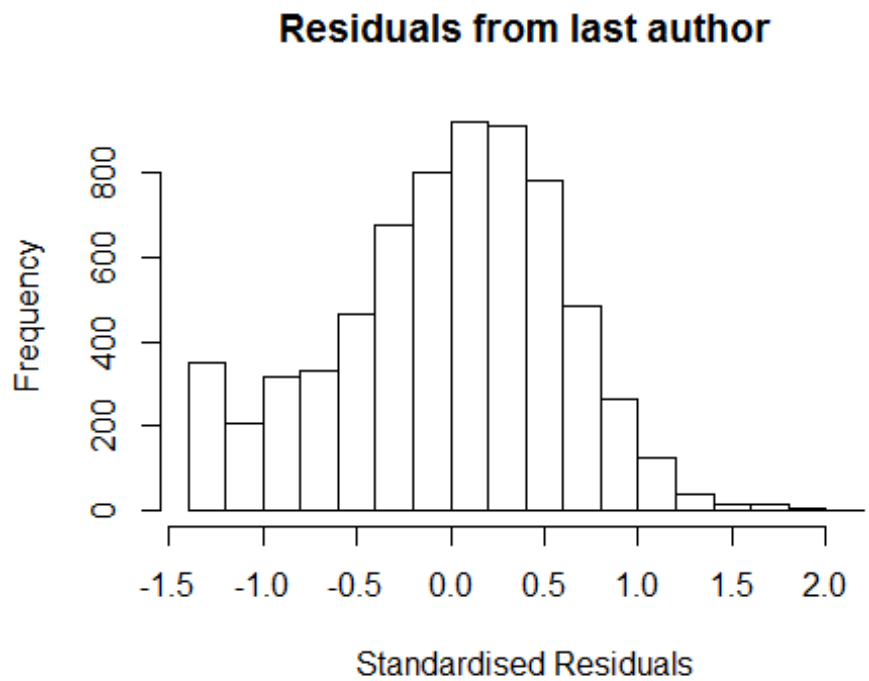

```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.3855 -0.3972  0.0502  0.4125  2.0150
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.168520   0.036888   31.68  < 2e-16 ***
## FirstAuthorFemale1  0.061874   0.016591    3.73  0.00019 ***
## LastAuthorFemale1   0.031229   0.016092    1.94  0.05234 .
## Year1997           -0.014209   0.049402   -0.29  0.77365
## Year1998            0.017302   0.052522    0.33  0.74185
## Year1999           -0.019964   0.049824   -0.40  0.68866
## Year2000            0.071607   0.046487    1.54  0.12352
## Year2001           -0.003324   0.048004   -0.07  0.94480
## Year2002            0.113341   0.047552    2.38  0.01717 *
## Year2003            0.041535   0.046300    0.90  0.36970
## Year2004            0.033109   0.049997    0.66  0.50785
## Year2005            0.123917   0.047348    2.62  0.00889 **
```

```
## Year2006           -0.010858   0.047044   -0.23  0.81747
## Year2007           -0.031144   0.047521   -0.66  0.51225
## Year2008           -0.000594   0.045594   -0.01  0.98960
## Year2009           -0.052733   0.045780   -1.15  0.24941
## Year2010           -0.012910   0.046381   -0.28  0.78075
## Year2011            0.034027   0.045575    0.75  0.45532
## Year2012            0.021209   0.047126    0.45  0.65268
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.587
## Multiple R-squared:  0.00946,    Adjusted R-squared:  0.0068
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  527 weights are ~= 1. The remaining 6185 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.214   0.862   0.948   0.902   0.985   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.49e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##                    GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.005  1           1.003
## Year              1.005 16           1.000
```




```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##    Min     1Q Median     3Q    Max
## -1.375 -0.397  0.047  0.412  2.035
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.177691   0.036687   32.10  < 2e-16 ***
## FirstAuthorFemale1  0.072305   0.016139    4.48  7.6e-06 ***
## Year1997           -0.011359   0.049337   -0.23   0.8179
## Year1998            0.019362   0.052473    0.37   0.7121
## Year1999           -0.018078   0.049840   -0.36   0.7168
## Year2000            0.073758   0.046469    1.59   0.1125
## Year2001           -0.002513   0.048042   -0.05   0.9583
## Year2002            0.116945   0.047517    2.46   0.0139 *
## Year2003            0.043229   0.046297    0.93   0.3505
## Year2004            0.036604   0.049905    0.73   0.4633
## Year2005            0.125467   0.047372    2.65   0.0081 **
## Year2006           -0.007581   0.047011   -0.16   0.8719
```

```
## Year2007           -0.029443   0.047576   -0.62   0.5360
## Year2008            0.000667   0.045625    0.01   0.9883
## Year2009           -0.051048   0.045775   -1.12   0.2648
## Year2010           -0.012633   0.046377   -0.27   0.7853
## Year2011            0.036030   0.045583    0.79   0.4293
## Year2012            0.023447   0.047127    0.50   0.6188
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.587
## Multiple R-squared:  0.00885,    Adjusted R-squared:  0.00634
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  522 weights are ~= 1. The remaining 6190 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.204   0.861   0.949   0.902   0.985   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.49e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##                   GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.012  1           1.006
## Year             1.012 16           1.000
```




```
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##     control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
##  \--> method = "MM"
## Residuals:
##    Min     1Q Median     3Q    Max
## -1.371 -0.399  0.048  0.409  1.968
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.19330    0.03611   33.05   <2e-16 ***
## LastAuthorFemale1   0.05107    0.01568    3.26   0.0011 **
## Year1997           -0.01421    0.04930   -0.29   0.7732
## Year1998            0.01801    0.05252    0.34   0.7317
## Year1999           -0.01918    0.04985   -0.38   0.7005
## Year2000            0.07468    0.04647    1.61   0.1081
## Year2001           -0.00174    0.04803   -0.04   0.9711
## Year2002            0.11366    0.04757    2.39   0.0169 *
## Year2003            0.04252    0.04641    0.92   0.3596
## Year2004            0.03357    0.05026    0.67   0.5041
## Year2005            0.12684    0.04739    2.68   0.0075 **
## Year2006           -0.00865    0.04704   -0.18   0.8540
```

```
## Year2007          -0.03025    0.04758   -0.64   0.5249
## Year2008           0.00256    0.04555    0.06   0.9552
## Year2009          -0.04915    0.04581   -1.07   0.2834
## Year2010          -0.00734    0.04643   -0.16   0.8743
## Year2011           0.03803    0.04565    0.83   0.4048
## Year2012           0.02356    0.04718    0.50   0.6176
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.587
## Multiple R-squared:  0.00727,    Adjusted R-squared:  0.00475
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  530 weights are ~= 1. The remaining 6182 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.238   0.862   0.949   0.902   0.984   0.999
## Algorithmic parameters:
##        tuning.chi                bb        tuning.psi        refine.tol
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol         solve.tol       eps.outlier             eps.x
##          1.00e-07          1.00e-07          1.49e-05          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
##   nResample      max.it    best.r.s    k.fast.s       k.max maxit.scale
##         500          50           2           1        1000         200
##   trace.lev         mts  compute.rd
##           0        1000           0
##                   psi           subsampling                   cov
##            "bisquare"         "nonsingular"         ".vcov.avar1"
## compute.outlier.stats
##                  "SM"
## seed : int(0)
## [1] "Sample size for the above analysis:  6712"
```

```r
for (iRow in 1:FieldCount){
  tryCatch({
    for (iCol in 1:10) {
      if (RegP[iRow,iCol]<0.001) {
        RegStar[iRow,iCol] <- "***"
      } else if (RegP[iRow,iCol]<0.01) {
        RegStar[iRow,iCol] <- "**"
      } else if (RegP[iRow,iCol]<0.05) {
        RegStar[iRow,iCol] <- "*"
      }
    }
  }, error = function(e) return("failed sample size"))
}
print("Stats just for gendered aticles from 1 country 1996 - 2018")
```

```
## [1] "Stats just for gendered aticles from 1 country 1996 - 2018"
```

```r
print(BasicStats)
```

```
##       1Ctry Gend 1st nth Fem1 96 Fem1 18     Chg1 FemN 96 FemN 18     ChgN
## 1000  65790        54.10  24.853  33.833   8.9798  13.162  25.218  12.0567
## 1100  70481        61.35  22.162  42.952  20.7902  19.351  29.488  10.1367
## 1101  20022        42.75  26.947  32.197   5.2499  20.093  29.167   9.0732
## 1102  47080        44.96  21.403  32.266  10.8633  16.885  23.616   6.7313
## 1103  69004        50.99  26.359  49.566  23.2070  21.473  31.405   9.9326
## 1104  68309        60.45  18.322  38.580  20.2587  17.104  26.646   9.5417
## 1105 133737        69.25  22.415  39.887  17.4721  19.413  28.132   8.7192
## 1106  51760        39.57  35.316  51.386  16.0699  26.582  35.490   8.9076
## 1107  34399        51.61  13.228  32.449  19.2211  13.071  23.106  10.0352
## 1108  12735        52.31  18.644  36.538  17.8944  16.271  30.385  14.1134
## 1109  39515        57.77  23.885  39.451  15.5664  18.930  25.105   6.1759
## 1110  55221        55.65  24.506  36.241  11.7354  20.763  31.086  10.3237
## 1111  33107        41.02  15.909  30.873  14.9644  14.876  22.590   7.7143
## 1200   5356        79.11  28.333  49.782  21.4483  26.667  51.965  25.2984
## 1201  59825        75.59  32.905  50.069  17.1631  32.847  45.496  12.6490
## 1202  40850        84.70  24.896  39.892  14.9953  25.864  39.250  13.3860
## 1203  38522        79.74  43.006  57.963  14.9575  39.306  55.950  16.6439
## 1204   7847        79.66  33.708  46.804  13.0958  33.708  40.183   6.4748
## 1205   2412        85.61  28.000  39.560  11.5604  28.000  40.659  12.6593
## 1206   1137        81.62  33.333  68.132  34.7985  33.333  54.945  21.6117
## 1207  14167        82.24  27.907  38.128  10.2206  25.914  35.821   9.9073
## 1208  31212        87.43  40.065  48.645   8.5797  39.414  48.784   9.3702
## 1209   1891        84.40  35.294  64.151  28.8568  38.235  58.491  20.2553
## 1210   5975        88.20  37.615  47.231   9.6166  39.450  44.625   5.1759
## 1211  37958        85.86  22.439  27.459   5.0198  20.854  27.309   6.4554
## 1212  21914        84.76  26.906  32.471   5.5648  27.354  32.863   5.5085
## 1213  16291        84.90  45.902  55.393   9.4910  46.721  54.450   7.7290
## 1300 107157        55.23  28.313  41.398  13.0849  16.871  28.147  11.2757
## 1301   1313        55.29     NaN  33.913      NaN     NaN  29.565     NaN
## 1302   9305        63.63  44.091  45.270   1.1794  35.455  37.162   1.7076
## 1303 266485        50.05  28.810  35.859   7.0490  15.077  23.239   8.1620
## 1304  62518        49.34  24.104  28.619   4.5147  13.844  20.265   6.4207
## 1305  61364        45.56  25.919  32.184   6.2645  14.361  22.917   8.5559
## 1306 101918        53.99  30.254  45.072  14.8181  18.476  33.283  14.8063
## 1307 164124        50.10  34.941  43.437   8.4962  18.910  27.725   8.8143
## 1308  45725        53.43  27.921  44.017  16.0955  20.962  30.536   9.5743
## 1309  39905        58.29  39.669  51.605  11.9365  20.662  37.693  17.0308
## 1310  46902        54.09  37.741  56.757  19.0157  22.681  41.164  18.4829
## 1311 141184        48.95  31.681  44.767  13.0867  19.715  31.834  12.1188
## 1312 239591        50.85  33.383  41.927   8.5432  17.994  27.982   9.9880
## 1313  60070        51.55  22.469  40.275  17.8061  15.333  24.729   9.3962
## 1314  92824        56.16  24.586  44.930  20.3448  17.035  28.529  11.4938
## 1315  22725        52.72  26.104  32.360   6.2551  16.466  21.124   4.6577
## 1400  20154        71.52  23.211  30.318   7.1068  22.258  30.988   8.7307
## 1401   8176        75.66  38.835  34.231  -4.6042  33.657  38.077   4.4200
```

```
## 1402  18755        69.48  13.098  24.320  11.2220  13.929  23.469   9.5401
## 1403  32095        70.91  26.634  36.011   9.3768  25.327  32.671   7.3447
## 1404   7685        58.07  23.500  27.222   3.7222  20.000  22.778   2.7778
## 1405  21331        72.93  18.704  27.933   9.2290  18.704  23.743   5.0390
## 1406  23941        67.62  22.575  42.999  20.4236  25.573  38.897  13.3236
## 1407  21174        78.75  28.169  41.828  13.6585  26.408  36.422  10.0137
## 1408  44174        68.10  20.755  34.304  13.5491  20.755  31.582  10.8276
## 1409   8540        62.21  31.579  40.179   8.5996  31.579  37.500   5.9211
## 1410   5103        71.12  16.279  29.858  13.5788  17.442  31.280  13.8378
## 1500  52265        43.57  14.634  24.113   9.4788   8.068  18.790  10.7228
## 1501   2027        43.36  14.286  26.582  12.2966   5.714  22.785  17.0705
## 1502  38011        46.25  20.091  27.586   7.4949  12.557  19.675   7.1184
## 1503  45809        46.98  16.667  26.944  10.2778   9.211  17.407   8.1969
## 1504   2180        69.91  28.947  51.923  22.9757  10.526  36.538  26.0121
## 1505  28589        48.36  16.957  24.675   7.7188   8.986  15.368   6.3825
## 1506   3473        42.67   9.459  13.462   4.0021  12.162   9.615  -2.5468
## 1507   8933        40.06   4.516  18.692  14.1755   7.742  12.150   4.4076
## 1508   5781        44.25  17.391  23.293   5.9019  15.942  15.261  -0.6810
## 1600 146383        46.89  16.990  28.691  11.7009  10.240  19.805   9.5645
## 1601   3190        48.93  19.672  35.393  15.7211  16.393  21.910   5.5167
## 1602  42275        51.87  19.580  33.529  13.9491  12.611  21.454   8.8431
## 1603  22614        39.72  17.293  29.455  12.1613  12.782  19.273   6.4908
## 1604  34261        51.82  21.451  27.413   5.9624  11.192  20.721   9.5290
## 1605  92490        50.16  14.946  28.146  13.1995  13.148  17.200   4.0525
## 1606 115039        48.39  16.314  24.240   7.9260   8.302  15.889   7.5865
## 1607  31310        48.09  15.625  32.161  16.5365  12.674  23.438  10.7639
## 1700  19159        56.59  15.503  19.817   4.3140  16.184  20.000   3.8160
## 1701   2051        54.51  60.000  20.635 -39.3651  60.000  23.810 -36.1905
## 1702  21216        56.47  12.137  22.549  10.4123  10.769  17.647   6.8778
## 1703  25465        49.48  11.777  19.570   7.7930  10.744  20.000   9.2562
## 1704  15255        46.79   8.933  16.889   7.9559  10.422  12.000   1.5782
## 1705  29122        48.32   8.908  21.190  12.2820  10.756  24.535  13.7790
## 1706  85820        47.93  14.075  24.955  10.8802  12.245  24.503  12.2578
## 1707  10647        48.48  11.852  18.079   6.2272   8.148  12.429   4.2812
## 1708  20780        44.69   9.239  10.660   1.4208  10.145   9.645  -0.5003
## 1709  10691        60.49  23.333  40.613  17.2797  14.667  40.996  26.3295
## 1710  31172        53.01  21.093  25.873   4.7803  17.916  27.143   9.2267
## 1711  17344        41.14   7.513  16.096   8.5829   5.440   8.904   3.4637
## 1712  51930        49.02   9.813  14.506   4.6934  10.187  14.969   4.7819
## 1800   4196        68.06  20.536  28.221   7.6851  19.643  22.699   3.0565
## 1801    543        70.72     NaN  33.333      NaN     NaN  26.667      NaN
## 1802  10121        56.61  19.665  26.000   6.3347  19.247  23.500   4.2531
## 1803  23808        54.78  14.286  23.243   8.9575  14.593  21.081   6.4881
## 1804  21711        51.69  11.530  22.569  11.0389  10.421  19.450   9.0283
## 1900  26128        50.55  18.417  29.646  11.2288  15.396  22.419   7.0232
## 1901  31291        47.44  13.456  29.561  16.1055  11.009  18.938   7.9285
## 1902  51518        47.58  12.717  31.514  18.7963  10.000  21.092  11.0918
## 1903   5743        54.12  11.215  20.732   9.5168  14.019  19.512   5.4935
## 1904  33709        52.99  14.977  34.245  19.2683  16.381  27.705  11.3245
## 1905   1419        58.28   6.452  20.000  13.5484   6.452  16.667  10.2151
```

```
## 1906  43100        48.84  15.571  29.072  13.5005  12.336  18.039   5.7028
## 1907  21645        59.22  15.722  27.063  11.3406  14.260  19.967   5.7074
## 1908  46914        46.22  14.875  29.038  14.1624  11.401  18.108   6.7070
## 1909  12083        43.56   6.410  18.879  12.4688   7.692  16.519   8.8269
## 1910  37101        51.72  15.082  40.217  25.1354  15.082  28.882  13.8000
## 1911  22029        49.21  15.709  35.431  19.7224  16.092  23.310   7.2181
## 1912  71446        48.09  11.965  23.725  11.7602  10.911  17.941   7.0298
## 1913   2163        63.94  10.000  30.208  20.2083  10.000  19.792   9.7917
## 2000  10256        71.00  15.476  23.618   8.1419  20.833  23.869   3.0360
## 2001   5713        75.30  44.828  31.741 -13.0870  37.931  27.645 -10.2860
## 2002  89078        71.57  15.099  24.522   9.4226  14.755  23.717   8.9615
## 2003  35437        68.24  12.712  19.098   6.3860  13.030  18.714   5.6844
## 2100  12973        48.51   7.895  24.220  16.3254  13.158  18.349   5.1907
## 2101   1970        59.49   4.545  21.008  16.4629   0.000  23.529  23.5294
## 2102  26735        41.08   9.716  15.932   6.2161   6.161  13.940   7.7791
## 2103  18341        43.73   8.430  18.924  10.4941   6.105  15.936   9.8316
## 2104  15454        36.50   6.011  17.436  11.4250   7.650  15.385   7.7343
## 2105  27282        43.83  13.585  27.493  13.9078  10.566  21.491  10.9248
## 2200  39100        49.70  12.469  22.970  10.5004   8.998  19.722  10.7240
## 2201   7617        57.57  14.371  24.823  10.4514  13.772  17.730   3.9580
## 2202  26770        49.15   6.329  12.233   5.9039   6.487  10.680   4.1923
## 2203   7360        45.37   7.692  16.766   9.0742   6.593  13.772   7.1790
## 2204  47367        50.13  16.570  31.565  14.9952  12.139  22.173  10.0338
## 2205  34888        46.17   9.946  16.875   6.9288   9.409  16.964   7.5557
## 2206   8326        33.57   5.263   8.955   3.6921   1.316  10.448   9.1320
## 2207  30454        44.29   8.389  10.400   2.0107   6.711  11.200   4.4886
## 2208 147579        40.27   7.297  13.244   5.9470   6.321  11.785   5.4631
## 2209  36485        45.17  16.000  19.664   3.6643  10.000  15.468   5.4676
## 2210 102929        35.58   8.423  14.680   6.2570   7.597  12.463   4.8660
## 2211  74428        32.60  10.137  15.451   5.3143   6.027  14.152   8.1242
## 2212   9608        51.16   9.412  26.761  17.3488   9.412  16.901   7.4896
## 2213  19445        55.10  14.241  25.081  10.8404  10.443  23.948  13.5052
## 2214   4150        47.64  13.740  40.541  26.8001  15.267  33.784  18.5166
## 2215  16901        49.67   5.221  15.385  10.1637   5.622  13.257   7.6345
## 2216   3072        68.46  16.949  35.625  18.6758  22.034  35.625  13.5911
## 2300  42615        59.62  18.607  40.037  21.4298  15.307  31.718  16.4114
## 2301   8054        72.98  28.655  41.481  12.8265  28.655  30.864   2.2092
## 2302   7345        58.71   9.929  30.417  20.4876  12.766  22.917  10.1507
## 2303  93046        62.85  20.468  39.968  19.5002  16.591  27.194  10.6029
## 2304  47616        55.56  18.910  36.887  17.9762  14.103  27.301  13.1980
## 2305  23307        51.66  15.190  33.427  18.2371  13.382  26.124  12.7420
## 2306  11051        66.41  15.190  39.755  24.5649  12.025  29.597  17.5719
## 2307  30336        59.60  25.888  49.912  24.0239  16.124  37.257  21.1324
## 2308  36759        68.34  15.823  36.283  20.4599  15.665  30.285  14.6205
## 2309  22040        72.95  19.438  41.093  21.6555  16.628  28.571  11.9438
## 2310  28583        53.87  18.750  36.224  17.4745  18.333  26.531   8.1973
## 2311  19175        53.83  19.251  35.342  16.0907  18.449  28.664  10.2153
## 2312  52921        50.87  14.370  30.579  16.2084  13.189  23.554  10.3647
## 2400  24202        57.55  29.167  41.121  11.9540  15.417  29.914  14.4971
## 2401    404        65.59     NaN  56.667      NaN     NaN  46.667     NaN
```

```
## 2402  22338        50.05  27.434  38.318  10.8841  15.265  25.467  10.2018
## 2403 107365        55.75  34.483  50.852  16.3691  20.145  32.329  12.1833
## 2404  57679        56.87  34.120  50.496  16.3753  19.573  33.411  13.8376
## 2405  19249        58.21  27.686  52.118  24.4319  18.595  33.149  14.5541
## 2406  35157        57.64  33.873  50.976  17.1032  20.423  31.670  11.2469
## 2500 110869        35.88  11.842  20.241   8.3991   7.785  15.831   8.0460
## 2501   3145        45.82  29.412  43.284  13.8718  27.941  38.060  10.1185
## 2502  25503        48.23  16.393  34.511  18.1174  11.475  25.136  13.6605
## 2503  28103        37.80  15.439  22.932   7.4929   8.789  16.541   7.7528
## 2504  96284        32.26   7.286  17.093   9.8063   5.913  15.330   9.4170
## 2505  89477        39.84  18.067  25.104   7.0371   9.476  18.031   8.5548
## 2506  26879        34.57  12.834  20.650   7.8159   6.952  17.017  10.0653
## 2507  38090        40.01  18.825  21.862   3.0367  13.084  17.630   4.5464
## 2508  56115        38.32  12.785  22.996  10.2108   8.524  16.698   8.1749
## 2600  37457        61.13  11.662  13.699   2.0368  12.342  13.790   1.4479
## 2601   3530        62.63  22.667  23.478   0.8116  13.333  23.478  10.1449
## 2602  15052        65.94  11.019  12.771   1.7513  12.672  11.472  -1.2003
## 2603  15488        50.72  11.111  10.667  -0.4444   9.722   9.067  -0.6556
## 2604  74597        48.42  11.291  14.734   3.4421  10.142  13.292   3.1495
## 2605  18785        44.43  13.746  14.347   0.6012   8.591  12.848   4.2569
## 2606   5962        46.80   5.155  15.000   9.8454   6.186  11.818   5.6326
## 2607   9768        59.02  14.103  18.750   4.6474  12.821  18.092   5.2716
## 2608  10446        65.36  11.792  14.011   2.2185   8.962  13.736   4.7740
## 2609   2630        76.39   9.302  14.943   5.6402  10.465  12.644   2.1786
## 2610  16531        47.11   6.516  10.490   3.9739   5.666   6.294   0.6280
## 2611  34686        45.67  10.139  22.618  12.4787   9.344  20.885  11.5415
## 2612   7461        43.13   9.091  14.286   5.1948   5.303  10.526   5.2233
## 2613  48391        48.43  11.642  20.534   8.8921  11.029  16.861   5.8320
## 2614  18710        54.29  11.716  15.650   3.9337  11.386  11.406   0.0197
## 2700 206205        66.02  27.594  46.116  18.5224  22.782  38.070  15.2884
## 2701  50005        67.15  37.670  56.809  19.1390  30.874  43.445  12.5713
## 2702  11986        63.25  38.596  47.467   8.8702  26.316  36.800  10.4842
## 2703  26221        64.82  19.923  43.116  23.1922  19.540  32.261  12.7211
## 2704   9325        58.34  39.474  52.195  12.7214  23.246  35.122  11.8763
## 2705 123508        63.63  13.583  28.211  14.6283  10.126  20.309  10.1828
## 2706  27317        68.85  16.986  39.977  22.9909  13.288  32.118  18.8308
## 2707   8168        67.45  25.352  51.634  26.2819  23.944  46.732  22.7884
## 2708  30674        65.80  30.687  52.941  22.2542  23.053  42.210  19.1564
## 2709    519        65.70  22.222  50.000  27.7778  11.111  50.000  38.8889
## 2710   3972        58.71  39.726  59.615  19.8894  16.438  50.000  33.5616
## 2711  28350        73.49  21.875  39.560  17.6854  17.308  32.747  15.4396
## 2712  49803        57.35  38.471  56.336  17.8651  22.816  40.550  17.7348
## 2713  38263        66.20  37.438  61.655  24.2170  27.454  46.290  18.8360
## 2714   6449        77.22  32.632  49.275  16.6438  36.842  41.739   4.8970
## 2715  37075        56.22  19.272  35.984  16.7123  16.488  20.315   3.8267
## 2716  32515        54.31  42.586  56.990  14.4047  27.186  41.262  14.0758
## 2717  28879        71.03  50.210  59.670   9.4592  47.546  48.397   0.8509
## 2718  12887        63.33  46.957  50.115   3.1584  43.478  41.034  -2.4438
## 2719  47677        73.67  30.866  52.719  21.8528  32.491  44.654  12.1631
## 2720  42405        53.99  27.671  46.776  19.1048  18.564  30.612  12.0483
```

```
## 2721  15385        56.65  22.222  32.515  10.2931  14.646  20.859   6.2124
## 2722  12289        55.54  41.379  42.984   1.6051  28.448  36.080   7.6319
## 2723  51235        58.30  34.016  48.853  14.8368  19.234  32.848  13.6142
## 2724  29000        64.07  29.970  48.786  18.8153  20.920  37.064  16.1445
## 2725  65855        63.28  31.122  50.967  19.8445  22.704  37.671  14.9672
## 2726  31234        64.12  30.000  46.485  16.4855  24.203  32.990   8.7868
## 2727  17325        61.54  26.486  42.670  16.1831  21.622  25.602   3.9801
## 2728 115454        63.64  24.104  37.875  13.7716  17.800  25.720   7.9204
## 2729  49821        63.04  30.386  70.416  40.0296  27.439  54.904  27.4650
## 2730 125196        59.72  29.502  47.807  18.3045  20.267  36.597  16.3304
## 2731  48938        58.02  23.834  38.915  15.0818  17.546  29.107  11.5612
## 2732  78969        70.38  15.515  24.520   9.0057  13.963  16.770   2.8070
## 2733  36077        68.20  21.144  35.933  14.7889  16.169  26.950  10.7807
## 2734  41447        64.08  30.172  48.857  18.6852  22.159  38.863  16.7049
## 2735 101366        69.23  39.072  60.430  21.3585  32.689  46.103  13.4144
## 2736  53521        61.04  27.838  47.501  19.6632  23.100  36.396  13.2961
## 2737  51890        59.72  21.668  38.335  16.6667  14.683  27.064  12.3811
## 2738 128012        74.14  37.492  58.369  20.8775  34.373  46.723  12.3499
## 2739 138291        70.51  37.361  60.519  23.1585  34.475  50.677  16.2025
## 2740  43362        63.57  20.460  35.012  14.5525  13.793  26.167  12.3740
## 2741 103948        55.78  21.651  32.970  11.3185  16.510  23.349   6.8385
## 2742  30146        71.97  39.447  56.413  16.9658  36.099  43.680   7.5813
## 2743  14124        59.55  29.032  67.111  38.0789  20.968  45.333  24.3656
## 2744     19        84.21     NaN     NaN      NaN     NaN     NaN     NaN
## 2745  10839        56.38  34.568  55.326  20.7586  23.868  44.330  20.4616
## 2746 162964        66.67  12.478  30.006  17.5276  11.806  18.862   7.0559
## 2747  21763        50.57  21.479  42.834  21.3550  13.028  27.524  14.4963
## 2748  34443        62.82   9.237  33.258  24.0215   8.032  17.228   9.1963
## 2800  90019        57.41  33.482  42.874   9.3915  18.480  29.246  10.7665
## 2801   4327        66.74  24.000  47.490  23.4903  24.000  35.907  11.9073
## 2802  28707        70.85  32.227  55.037  22.8098  25.588  44.139  18.5514
## 2803  15164        70.75  30.366  53.090  22.7238  20.681  38.193  17.5127
## 2804  44113        51.51  35.320  44.534   9.2144  20.910  31.183  10.2727
## 2805  25446        67.78  33.138  50.655  17.5172  29.032  40.611  11.5791
## 2806  11111        65.06  37.700  58.209  20.5093  28.115  39.104  10.9895
## 2807   6281        58.91  43.878  66.986  23.1081  30.612  49.282  18.6701
## 2808  48075        61.72  28.226  43.650  15.4246  19.489  30.488  10.9992
## 2809  26423        55.22  29.226  35.294   6.0682  22.749  27.336   4.5868
## 2900  34172        75.45  76.731  80.433   3.7022  71.283  72.419   1.1361
## 2901   5443        77.00  85.714  83.108  -2.6062  80.000  76.351  -3.6486
## 2902  19855        71.91  51.929  71.799  19.8700  47.774  64.482  16.7072
## 2903   6216        69.79  65.217  75.519  10.3013  65.217  69.295   4.0772
## 2904   2781        70.69  69.565  59.259 -10.3060  69.565  47.407 -22.1578
## 2905   5960        74.45  68.056  79.279  11.2237  63.889  72.523   8.6336
## 2906   9061        71.02  83.929  83.099  -0.8300  79.464  77.465  -1.9995
## 2907   9688        70.63  55.224  60.182   4.9585  55.970  55.927  -0.0431
## 2908   5345        77.34     NaN  81.369      NaN     NaN  72.624     NaN
## 2909   9477        71.91  52.308  68.856  16.5488  50.769  59.124   8.3549
## 2910   5832        83.64  36.000  44.086   8.0860  34.667  43.011   8.3441
## 2911  15999        74.66  55.691  72.678  16.9865  53.659  68.124  14.4653
```

```
## 2912   9003        71.08  59.130  83.108  23.9777  59.130  80.743  21.6128
## 2913   4915        75.40  63.855  85.646  21.7905  59.036  74.641  15.6050
## 2914   5216        74.18  76.389  68.342  -8.0472  73.611  64.824  -8.7870
## 2915   1473        87.17     NaN  74.227      NaN     NaN  63.918      NaN
## 2916  26601        63.46  52.967  69.442  16.4747  35.604  52.693  17.0890
## 2917   4855        72.71  91.667  83.333  -8.3333  91.667  71.053 -20.6140
## 2918      0           NA      NA      NA       NA      NA      NA       NA
## 2919   6093        75.61  86.517  88.679   2.1624  87.640  81.132  -6.5084
## 2920   2000        73.45  75.000  78.125   3.1250  62.500  75.000  12.5000
## 2921   8229        79.39  74.317  68.992  -5.3247  71.038  59.948 -11.0899
## 2922   1264        81.96 100.000  96.078  -3.9216 100.000  89.216 -10.7843
## 2923   3023        73.57     NaN  84.733      NaN     NaN  86.260      NaN
## 3000   8318        63.84  32.161  50.309  18.1478  23.116  46.605  23.4894
## 3001   1855        70.89  35.897  47.436  11.5385  29.487  39.744  10.2564
## 3002  45714        50.98  14.497  35.238  20.7411  13.905  20.476   6.5709
## 3003  40528        49.53  22.374  38.992  16.6173  16.743  24.074   7.3313
## 3004  90894        57.14  26.387  43.868  17.4809  18.358  30.103  11.7448
## 3005  46898        60.09  28.617  52.298  23.6809  21.743  37.883  16.1404
## 3100 103376        38.23   9.302  19.308  10.0059   6.977  14.930   7.9532
## 3101  48090        31.98   8.481  11.952   3.4710   6.007  11.346   5.3394
## 3102  15352        48.65  16.332  22.910   6.5786  15.075  20.124   5.0485
## 3103  40390        49.98  11.919  21.638   9.7194  10.562  19.211   8.6493
## 3104 180052        33.53  11.036  16.258   5.2221   6.554  13.391   6.8373
## 3105  32113        38.30  12.127  19.840   7.7130   8.946  13.981   5.0350
## 3106  56100        42.07   7.359  12.803   5.4440   8.016   8.478   0.4617
## 3107  62097        39.61   7.717  14.554   6.8369   6.660  11.628   4.9683
## 3108  15821        50.62  18.625  25.948   7.3229  12.894  18.950   6.0565
## 3109  22218        43.76   7.619   8.020   0.4010   6.190   7.769   1.5789
## 3110  25828        34.59  13.068  22.876   9.8076   9.659  21.569  11.9095
## 3200  51816        77.06  35.367  53.152  17.7844  34.184  47.907  13.7229
## 3201   6959        74.68  31.048  49.603  18.5548  33.468  41.667   8.1989
## 3202  42059        76.13  36.973  55.273  18.2998  35.243  47.117  11.8736
## 3203  66427        78.66  37.918  58.698  20.7795  36.927  47.185  10.2578
## 3204  77959        76.76  48.920  66.569  17.6487  46.811  58.885  12.0741
## 3205  39749        76.11  32.936  50.949  18.0128  29.106  43.950  14.8438
## 3206  15672        75.11  39.950  51.520  11.5705  30.653  40.034   9.3805
## 3207  54138        77.32  43.573  58.015  14.4416  44.120  53.958   9.8379
## 3300  23641        79.44  29.619  44.081  14.4621  32.405  42.830  10.4250
## 3301  39951        78.67  43.562  57.437  13.8749  43.175  53.259  10.0838
## 3302   8879        79.56  25.000  45.513  20.5128  25.893  41.239  15.3465
## 3303  20946        71.81  24.708  42.894  18.1856  25.097  42.125  17.0282
## 3304 172107        78.38  42.388  59.036  16.6478  42.054  57.012  14.9575
## 3305  51367        73.03  21.341  38.947  17.6059  21.748  36.541  14.7934
## 3306  49545        73.97  47.243  63.851  16.6077  43.358  56.164  12.8060
## 3307   7290        65.65  23.077  35.470  12.3932  21.154  29.487   8.3333
## 3308  55103        81.47  27.464  42.339  14.8748  27.163  40.587  13.4235
## 3309  35994        75.94  44.924  56.618  11.6932  43.952  57.904  13.9519
## 3310  44040        79.70  44.305  58.149  13.8440  40.648  56.350  15.7018
## 3311   5853        74.32  20.000  34.884  14.8837  14.545  34.012  19.4662
## 3312 121701        80.38  31.747  45.726  13.9789  32.003  44.047  12.0437
```

```
## 3313   8572        53.34  13.571  25.220  11.6485  18.571  25.220   6.6485
## 3314  30524        79.85  41.348  54.752  13.4044  40.582  53.306  12.7239
## 3315  32041        77.64  41.591  53.928  12.3369  40.145  53.367  13.2223
## 3316  43474        82.08  36.893  50.786  13.8932  35.761  51.076  15.3156
## 3317   8978        73.72  39.819  58.021  18.2024  43.891  53.743   9.8519
## 3318  17178        81.33  67.630  73.050   5.4196  64.451  68.440   3.9888
## 3319  11474        75.30  50.186  68.806  18.6198  48.327  61.676  13.3484
## 3320  32318        81.56  17.978  29.823  11.8459  17.790  31.458  13.6682
## 3321  15978        80.34  27.003  38.697  11.6938  27.893  36.037   8.1441
## 3322  12731        76.36  25.879  37.345  11.4665  28.754  38.053   9.2991
## 3400  38373        59.71  36.830  63.149  26.3183  29.799  41.777  11.9775
## 3401    736        45.11  50.000  55.882   5.8824  50.000  50.000   0.0000
## 3402   3262        24.52  40.909  72.000  31.0909  25.000  44.000  19.0000
## 3403   5331        30.16  30.645  71.084  40.4392  16.129  50.602  34.4734
## 3404   4220        59.00  40.800  73.077  32.2769  32.800  49.038  16.2385
## 3500  16080        55.30  20.427  35.890  15.4627  18.293  31.902  13.6092
## 3501    439        54.21     NaN  53.125      NaN     NaN  50.000     NaN
## 3502      0           NA      NA      NA       NA      NA      NA      NA
## 3503     21        52.38     NaN 100.000      NaN     NaN   0.000     NaN
## 3504   7358        60.13  12.054  26.014  13.9599  10.714  21.959  11.2452
## 3505   1786        49.16   0.000  34.483  34.4828  50.000  20.690 -29.3103
## 3506   1437        54.98  13.514  28.125  14.6115  13.514  15.625   2.1115
## 3600   8159        70.44  37.255  58.427  21.1721  39.542  50.562  11.0193
## 3601   2962        80.18  58.750  76.336  17.5859  62.500  65.649   3.1489
## 3602   1147        76.98  14.286  40.426  26.1398  14.286  34.043  19.7568
## 3603   1465        79.66  61.111  66.250   5.1389  55.556  56.250   0.6944
## 3604    370        57.84     NaN  40.000      NaN     NaN  80.000     NaN
## 3605   3943        60.49  25.000  49.064  24.0637  25.000  39.326  14.3258
## 3606    356        85.67     NaN  50.000      NaN     NaN  50.000     NaN
## 3607   7153        54.82  26.846  45.794  18.9488  13.423  35.514  22.0912
## 3608    144        90.97     NaN     NaN      NaN     NaN     NaN     NaN
## 3609   2744        76.49  79.787  81.633   1.8454  80.851  68.367 -12.4837
## 3610   1920        71.93  23.529  38.333  14.8039  23.529  35.000  11.4706
## 3611   4274        75.48  26.316  53.386  27.0707  10.526  43.028  32.5016
## 3612  34588        73.78  30.693  43.344  12.6514  27.723  36.856   9.1334
## 3613   1126        79.22  11.111  20.930   9.8191  16.667   9.302  -7.3643
## 3614  19906        52.22  26.023  42.745  16.7217  16.959  30.000  13.0409
## 3615      0           NA      NA      NA       NA      NA      NA      NA
## 3616  13278        76.14  54.790  71.854  17.0639  49.701  62.583  12.8822
```

```r
print(paste("Stats just for gendered aticles from 1 country 1996 -",
LastYearForCitationAnalysis))
```

```
## [1] "Stats just for gendered aticles from 1 country 1996 - 2012"
```

```r
print(BasicStats2)
```

```
##      1CtyMFto14 Fem1All Fem1 96 Fem1 14      Chg1 FemNAll  FemN96 FemN14
## 1000      23663  27.279  24.853  31.364   6.51070  16.321  13.162 19.351
## 1100      21942  32.978  22.162  40.785  18.62334  22.874  19.351 25.504
## 1101       6988  27.519  26.947  30.682   3.73478  22.295  20.093 25.000
```

```
## 1102      14104  22.830  21.403  26.856   5.45300  17.463  16.885 19.982
## 1103      24860  32.796  26.359  38.471  12.11162  24.618  21.473 27.367
## 1104      29126  26.200  18.322  32.908  14.58607  19.663  17.104 21.758
## 1105      62287  29.733  22.415  34.422  12.00712  22.250  19.413 23.528
## 1106      14339  42.137  35.316  50.412  15.09526  30.902  26.582 37.603
## 1107      12622  21.510  13.228  26.572  13.34342  16.661  13.071 21.234
## 1108       4863  27.884  18.644  35.714  17.07022  21.036  16.271 28.571
## 1109      16758  27.080  23.885  32.837   8.95217  18.791  18.930 20.372
## 1110      22608  29.485  24.506  35.249  10.74366  22.461  20.763 26.593
## 1111       9403  23.312  15.909  27.891  11.98207  15.516  14.876 20.408
## 1200       2415  37.971  28.333  45.608  17.27477  39.089  26.667 45.608
## 1201      30252  39.706  32.905  44.118  11.21216  37.009  32.847 37.804
## 1202      21590  32.594  24.896  36.658  11.76209  32.631  25.864 36.259
## 1203      19886  48.909  43.006  51.929   8.92348  48.059  39.306 49.732
## 1204       3721  32.975  33.708  33.787   0.07960  32.196  33.708 34.060
## 1205       1292  28.715  28.000  33.557   5.55705  28.947  28.000 34.228
## 1206        431  44.780  33.333  58.974  25.64103  44.780  33.333 58.974
## 1207       7155  32.676  27.907  36.328   8.42115  31.572  25.914 34.766
## 1208      16494  43.252  40.065  44.540   4.47493  43.416  39.414 44.755
## 1209        954  40.461  35.294  45.000   9.70588  39.832  38.235 48.000
## 1210       3118  39.897  37.615  44.479   6.86385  40.475  39.450 42.945
## 1211      19504  21.944  22.439  23.764   1.32485  21.365  20.854 23.058
## 1212      11147  26.770  26.906  27.808   0.90264  27.173  27.354 27.624
## 1213       7893  48.575  45.902  49.454   3.55251  48.714  46.721 49.236
## 1300      28281  34.712  28.313  40.368  12.05488  22.609  16.871 25.784
## 1301        159  28.302     NaN  25.000       NaN  17.610     NaN 25.000
## 1302       3959  44.632  44.091  51.087   6.99605  32.811  35.455 33.696
## 1303     103817  31.870  28.810  34.709   5.89927  17.854  15.077 19.650
## 1304      22650  28.464  24.104  33.422   9.31792  16.243  13.844 19.668
## 1305      19164  30.098  25.919  32.230   6.31087  18.717  14.361 20.450
## 1306      38905  37.982  30.254  42.983  12.72947  23.097  18.476 29.092
## 1307      64067  38.822  34.941  41.343   6.40253  21.899  18.910 25.108
## 1308      18981  33.291  27.921  36.847   8.92573  21.722  20.962 24.913
## 1309      17679  42.989  39.669  48.718   9.04918  27.021  20.662 31.339
## 1310      18725  43.578  37.741  53.679  15.93788  27.407  22.681 34.699
## 1311      48404  36.266  31.681  39.356   7.67584  22.477  19.715 25.303
## 1312      91611  36.563  33.383  39.903   6.51961  20.917  17.994 23.311
## 1313      21368  29.497  22.469  33.315  10.84592  18.055  15.333 21.216
## 1314      39158  30.765  24.586  36.467  11.88185  18.313  17.035 20.255
## 1315       9060  28.157  26.104  28.864   2.75964  17.064  16.466 19.181
## 1400      10489  25.532  23.211  25.957   2.74600  25.379  22.258 25.674
## 1401       4510  35.477  38.835  30.702  -8.13320  34.812  33.657 29.532
## 1402       9110  19.989  13.098  25.477  12.37913  19.660  13.929 24.796
## 1403      15766  27.553  26.634  29.301   2.66688  26.976  25.327 29.379
## 1404       3337  25.442  23.500  27.835   4.33505  23.105  20.000 24.742
## 1405      11429  23.764  18.704  27.977   9.27307  23.362  18.704 26.255
## 1406      11739  30.829  22.575  35.749  13.17384  29.926  25.573 30.072
## 1407      11710  34.509  28.169  40.968  12.79873  34.091  26.408 40.323
## 1408      20883  26.217  20.755  30.270   9.51521  25.676  20.755 30.141
## 1409       3409  35.817  31.579  38.944   7.36495  34.380  31.579 30.033
```

```
## 1410       2493  22.343  16.279  28.022  11.74291  21.861  17.442 28.022
## 1500      14183  18.318  14.634  24.251   9.61653  11.535   8.068 15.441
## 1501        616  15.584  14.286   9.091  -5.19481  10.065   5.714  9.091
## 1502      10983  26.614  20.091  30.567  10.47599  16.052  12.557 19.983
## 1503      14925  21.159  16.667  25.667   9.00000  10.238   9.211 12.500
## 1504       1125  29.689  28.947  39.189  10.24182  22.667  10.526 29.730
## 1505      10482  23.039  16.957  27.298  10.34153  11.429   8.986 13.928
## 1506       1141  18.230   9.459  15.152   5.69206  10.692  12.162  6.061
## 1507       2308   8.449   4.516  12.097   7.58065   8.146   7.742 11.290
## 1508       1483  17.937  17.391  17.172  -0.21959  12.272  15.942 12.121
## 1600      45777  22.902  16.990  26.693   9.70287  13.904  10.240 16.732
## 1601       1003  26.022  19.672  31.915  12.24276  18.046  16.393 29.787
## 1602      15942  26.628  19.580  31.298  11.71880  15.983  12.611 19.574
## 1603       6390  21.440  17.293  27.112   9.81875  13.380  12.782 17.092
## 1604      12938  22.747  21.451  25.935   4.48380  12.398  11.192 16.472
## 1605      35402  19.680  14.946  23.229   8.28299  12.909  13.148 13.314
## 1606      40391  19.106  16.314  20.433   4.11925  11.309   8.302 13.976
## 1607      10715  22.930  15.625  25.286   9.66071  14.298  12.674 18.429
## 1700       7700  17.857  15.503  22.747   7.24423  17.208  16.184 21.030
## 1701        587  21.806  60.000  18.966 -41.03448  22.487  60.000 20.690
## 1702       9128  17.452  12.137  21.507   9.37060  16.499  10.769 20.956
## 1703       9082  14.072  11.777  15.689   3.91229  12.068  10.744 14.370
## 1704       5673  10.541   8.933   9.732   0.79854  10.788  10.422 14.765
## 1705      10725  16.662   8.908  19.669  10.76186  15.925  10.756 19.669
## 1706      28421  17.051  14.075  18.625   4.55090  15.123  12.245 18.028
## 1707       4103  12.137  11.852  12.717   0.86491  10.821   8.148 16.185
## 1708       7217  11.279   9.239  11.538   2.29933  10.586  10.145 11.538
## 1709       4056  29.043  23.333  32.927   9.59350  26.603  14.667 34.451
## 1710      12284  23.030  21.093  27.660   6.56682  21.459  17.916 25.532
## 1711       5264   9.821   7.513  12.014   4.50118   7.903   5.440  8.834
## 1712      19933  12.562   9.813  13.679   3.86602  12.136  10.187 12.953
## 1800       1908  21.069  20.536  27.439   6.90331  21.069  19.643 28.049
## 1801        173  30.058     NaN  25.455       NaN  17.919     NaN 10.909
## 1802       4328  22.066  19.665  26.214   6.54832  19.016  19.247 20.874
## 1803       9631  17.288  14.286  18.954   4.66853  16.800  14.593 20.098
## 1804       7984  14.717  11.530  19.512   7.98226  13.364  10.421 12.383
## 1900       9257  20.784  18.417  22.924   4.50632  18.138  15.396 18.272
## 1901      11074  21.329  13.456  27.692  14.23665  14.728  11.009 19.487
## 1902      16925  18.003  12.717  22.252   9.53438  13.235  10.000 19.326
## 1903       2071  16.127  11.215  19.136   7.92085  15.886  14.019 18.519
## 1904      12385  23.351  14.977  29.454  14.47708  18.870  16.381 25.059
## 1905        518  15.830   6.452  16.279   9.82746  12.355   6.452 23.256
## 1906      15652  21.109  15.571  28.769  13.19742  14.177  12.336 19.908
## 1907       9086  18.787  15.722  22.913   7.19050  15.948  14.260 17.476
## 1908      15862  19.903  14.875  26.502  11.62701  13.769  11.401 19.635
## 1909       3462  10.976   6.410  14.692   8.28169   9.157   7.692  8.057
## 1910      14287  23.217  15.082  30.133  15.05137  17.064  15.082 22.000
## 1911       7893  22.476  15.709  29.213  13.50467  16.445  16.092 20.899
## 1912      25815  17.439  11.965  22.164  10.19831  12.787  10.911 16.095
## 1913        828  18.599  10.000  23.894  13.89381  17.391  10.000 21.239
```

```
## 2000       4539  18.462  15.476  19.910   4.43331  19.277  20.833 19.005
## 2001       2801  27.776  44.828  28.634 -16.19322  27.562  37.931 29.956
## 2002      45106  17.889  15.099  21.156   6.05682  17.667  14.755 20.357
## 2003      17211  16.797  12.712  17.285   4.57306  16.867  13.030 17.048
## 2100       3039  21.224   7.895  24.304  16.40953  13.656  13.158 14.657
## 2101        547  13.346   4.545  10.294   5.74866  10.420   0.000 10.294
## 2102       6732  11.319   9.716  12.839   3.12342   8.393   6.161  9.403
## 2103       5191  12.291   8.430  11.483   3.05302   9.131   6.105  8.373
## 2104       3334   9.868   6.011  11.246   5.23527   7.888   7.650  7.599
## 2105       6130  16.476  13.585  21.931   8.34613  12.985  10.566 13.517
## 2200      14789  16.249  12.469  21.078   8.60899  13.875   8.998 16.667
## 2201       2987  19.953  14.371  25.592  11.22116  18.179  13.772 28.436
## 2202       9512   8.715   6.329  12.126   5.79713   7.811   6.487  9.302
## 2203       2026   9.329   7.692  10.553   2.86046   8.243   6.593 11.055
## 2204      13968  24.871  16.570  28.178  11.60721  15.829  12.139 17.014
## 2205      10253  13.196   9.946  17.578   7.63189  11.265   9.409 14.583
## 2206       2056   7.442   5.263   3.883  -1.37966   6.615   1.316  7.767
## 2207       9463   8.655   8.389   9.363   0.97329   7.323   6.711  9.363
## 2208      44959   9.640   7.297  13.561   6.26407   7.992   6.321 10.771
## 2209      11522  13.652  16.000  15.960  -0.03955  11.101  10.000 11.723
## 2210      23914  10.128   8.423  12.513   4.09069   8.246   7.597 11.219
## 2211      15771  12.669  10.137  15.962   5.82546   9.156   6.027 12.676
## 2212       3355  10.641   9.412  15.842   6.42982  10.134   9.412 14.851
## 2213       6733  19.174  14.241  23.270   9.02993  17.199  10.443 19.811
## 2214       1499  21.948  13.740  37.705  23.96446  20.947  15.267 34.426
## 2215       4949   9.416   5.221  15.566  10.34515   8.406   5.622 10.849
## 2216       1153  24.458  16.949  35.833  18.88418  24.892  22.034 35.000
## 2300      18898  26.945  18.607  32.821  14.21471  20.907  15.307 23.321
## 2301       3824  31.145  28.655  34.109   5.45356  30.021  28.655 28.295
## 2302       2698  24.611   9.929  28.906  18.97717  18.903  12.766 17.969
## 2303      39287  28.755  20.468  34.713  14.24522  20.908  16.591 26.060
## 2304      18065  29.017  18.910  35.603  16.69270  19.247  14.103 21.575
## 2305       8148  21.306  15.190  26.756  11.56598  16.078  13.382 20.067
## 2306       4292  26.212  15.190  32.512  17.32244  20.783  12.025 24.631
## 2307      11773  34.885  25.888  45.608  19.72053  23.936  16.124 31.419
## 2308      15495  24.982  15.823  30.511  14.68806  21.962  15.665 26.032
## 2309       9628  29.040  19.438  33.333  13.89539  21.853  16.628 25.889
## 2310       9667  26.689  18.750  31.266  12.51551  20.306  18.333 21.092
## 2311       6667  23.909  19.251  30.017  10.76612  18.584  18.449 21.640
## 2312      18244  20.719  14.370  24.398  10.02826  16.098  13.189 21.411
## 2400       7131  34.034  29.167  40.995  11.82859  23.166  15.417 28.791
## 2401         97  48.454     NaN  62.500       NaN  26.804     NaN 37.500
## 2402       8439  35.063  27.434  40.801  13.36783  20.370  15.265 26.230
## 2403      45384  41.206  34.483  45.415  10.93225  23.158  20.145 25.516
## 2404      22891  41.488  34.120  48.667  14.54633  22.585  19.573 27.697
## 2405       7701  39.320  27.686  47.109  19.42289  20.491  18.595 25.680
## 2406      13812  41.927  33.873  45.540  11.66733  22.126  20.423 25.386
## 2500      24211  16.579  11.842  20.846   9.00422  11.718   7.785 15.724
## 2501       1014  25.247  29.412  36.364   6.95187  24.162  27.941 18.182
## 2502       7326  28.569  16.393  31.169  14.77539  18.127  11.475 20.071
```

```
## 2503       7362  20.035  15.439  25.046   9.60695  12.986   8.789 17.811
## 2504      20875  13.849   7.286  18.155  10.86929   9.657   5.913 11.601
## 2505      25971  20.003  18.067  22.870   4.80271  11.775   9.476 15.539
## 2506       6253  16.840  12.834  21.263   8.42893  11.163   6.952 16.632
## 2507      10681  21.487  18.825  23.547   4.72184  13.416  13.084 15.499
## 2508      14623  18.847  12.785  21.247   8.46132  11.708   8.524 14.838
## 2600      15313  12.695  11.662  14.558   2.89612  12.721  12.342 13.643
## 2601       1462  17.031  22.667  10.526 -12.14035  15.185  13.333 12.281
## 2602       6924  12.565  11.019  13.320   2.30080  13.504  12.672 14.314
## 2603       5347  12.755  11.111  13.126   2.01538  11.895   9.722 12.172
## 2604      25642  12.842  11.291  13.940   2.64862  11.130  10.142 10.889
## 2605       5687  13.645  13.746  12.555  -1.19064  11.605   8.591 12.775
## 2606       1680   9.107   5.155  10.345   5.19019   9.226   6.186 16.379
## 2607       3913  14.132  14.103  17.045   2.94289  13.800  12.821 16.667
## 2608       4542  13.232  11.792  13.703   1.91017  12.748   8.962 13.994
## 2609       1399   7.791   9.302  11.957   2.65420   7.720  10.465  8.696
## 2610       6114   8.096   6.516  11.985   5.46944   7.082   5.666  5.993
## 2611       9937  15.135  10.139  22.618  12.47864  12.609   9.344 19.686
## 2612       2321  11.676   9.091  11.616   2.52525   8.875   5.303  8.586
## 2613      15997  15.947  11.642  17.602   5.96027  13.540  11.029 13.657
## 2614       7792  12.012  11.716  13.832   2.11603  12.205  11.386 12.472
## 2700      86658  37.040  27.594  41.489  13.89538  30.964  22.782 33.073
## 2701      21664  45.813  37.670  50.312  12.64195  36.600  30.874 39.189
## 2702       5491  39.465  38.596  43.416   4.81988  27.245  26.316 33.808
## 2703      11011  28.753  19.923  34.816  14.89225  23.304  19.540 27.223
## 2704       3929  40.494  39.474  46.099   6.62561  25.834  23.246 30.851
## 2705      54613  20.349  13.583  24.198  10.61542  13.462  10.126 14.289
## 2706      13171  27.576  16.986  35.692  18.70569  18.336  13.288 23.101
## 2707       3535  42.914  25.352  44.648  19.29621  38.303  23.944 37.309
## 2708      12998  44.607  30.687  51.974  21.28727  31.790  23.053 37.098
## 2709        290  30.000  22.222  20.000  -2.22222  32.069  11.111 10.000
## 2710       1727  47.481  39.726  51.754  12.02836  29.936  16.438 41.228
## 2711      13021  27.433  21.875  34.768  12.89264  22.441  17.308 27.453
## 2712      19175  44.271  38.471  52.069  13.59830  28.490  22.816 34.095
## 2713      15753  49.927  37.438  54.700  17.26267  36.330  27.454 40.940
## 2714       2902  37.802  32.632  40.290   7.65828  37.388  36.842 41.449
## 2715      13663  26.327  19.272  32.703  13.43075  17.954  16.488 16.937
## 2716      11217  44.513  42.586  49.279   6.69373  30.703  27.186 34.324
## 2717      13958  53.353  50.210  55.832   5.62153  43.853  47.546 43.739
## 2718       4117  40.223  46.957  45.472  -1.48408  35.098  43.478 38.976
## 2719      19410  42.102  30.866  45.122  14.25605  37.661  32.491 39.536
## 2720      15141  35.308  27.671  36.502   8.83084  22.482  18.564 25.319
## 2721       5700  26.719  22.222  36.327  14.10512  16.474  14.646 18.762
## 2722       4393  40.974  41.379  45.179   3.79975  27.908  28.448 32.507
## 2723      20374  39.079  34.016  43.849   9.83318  24.448  19.234 26.190
## 2724      12555  39.164  29.970  43.128  13.15764  25.520  20.920 29.976
## 2725      26555  41.567  31.122  49.485  18.36253  26.296  22.704 30.815
## 2726      13366  39.788  30.000  46.768  16.76768  27.241  24.203 29.798
## 2727       7411  30.482  26.486  36.268   9.78112  21.374  21.622 20.423
## 2728      45865  30.138  24.104  36.769  12.66540  21.908  17.800 25.667
```

```
## 2729      20652  48.107  30.386  62.828  32.44198  36.171  27.439 46.241
## 2730      47819  39.135  29.502  44.348  14.84616  26.019  20.267 31.737
## 2731      20477  30.629  23.834  36.090  12.25597  22.367  17.546 22.566
## 2732      33060  20.130  15.515  24.095   8.58036  17.157  13.963 17.828
## 2733      16811  25.775  21.144  31.504  10.36015  20.368  16.169 24.425
## 2734      17014  38.850  30.172  44.518  14.34656  28.529  22.159 33.488
## 2735      44018  47.817  39.072  55.254  16.18196  37.019  32.689 42.629
## 2736      20779  36.340  27.838  43.394  15.55604  27.311  23.100 30.293
## 2737      21825  27.858  21.668  30.587   8.91963  16.948  14.683 18.519
## 2738      59457  45.862  37.492  52.798  15.30677  37.937  34.373 42.483
## 2739      60815  48.731  37.361  54.759  17.39869  41.894  34.475 47.084
## 2740      18321  26.330  20.460  32.235  11.77569  17.450  13.793 20.182
## 2741      38125  23.137  21.651  26.203   4.55238  16.567  16.510 19.390
## 2742      14470  47.540  39.447  51.796  12.34882  40.712  36.099 41.341
## 2743       5635  43.940  29.032  54.338  25.30564  31.074  20.968 40.183
## 2744         15  26.667     NaN     NaN       NaN  20.000     NaN    NaN
## 2745       4144  41.361  34.568  50.299  15.73150  28.185  23.868 31.737
## 2746      69555  19.204  12.478  25.619  13.14062  13.842  11.806 16.459
## 2747       7671  28.966  21.479  37.002  15.52302  17.781  13.028 23.340
## 2748      14176  20.873   9.237  28.179  18.94175  12.239   8.032 16.581
## 2800      37080  34.725  33.482  39.982   6.49942  21.993  18.480 24.909
## 2801       1338  35.800  24.000  36.287  12.28692  26.383  24.000 29.114
## 2802      13434  41.589  32.227  50.817  18.58969  31.383  25.588 36.792
## 2803       7110  39.297  30.366  47.917  17.55017  27.848  20.681 31.597
## 2804      15368  37.936  35.320  37.646   2.32653  23.887  20.910 24.942
## 2805      10313  38.098  33.138  40.220   7.08256  29.031  29.032 30.119
## 2806       4841  43.379  37.700  47.802  10.10252  31.584  28.115 35.165
## 2807       2285  55.098  43.878  66.102  22.22414  39.475  30.612 44.068
## 2808      18386  35.532  28.226  40.056  11.83006  24.171  19.489 27.709
## 2809      10656  31.250  29.226  35.213   5.98730  22.100  22.749 20.495
## 2900      18336  73.151  76.731  81.212   4.48060  67.261  71.283 74.627
## 2901       2470  81.862  85.714  86.312   0.59750  75.628  80.000 77.186
## 2902       9838  67.412  51.929  71.204  19.27541  60.012  47.774 63.743
## 2903       2823  73.114  65.217  72.075   6.85808  70.351  65.217 64.906
## 2904       1173  49.872  69.565  46.018 -23.54752  43.137  69.565 33.628
## 2905       2959  72.356  68.056  75.299   7.24325  68.672  63.889 66.135
## 2906       4066  84.333  83.929  80.628  -3.30030  79.316  79.464 73.560
## 2907       4613  58.097  55.224  61.930   6.70641  55.734  55.970 56.836
## 2908       2282  77.651     NaN  75.581       NaN  72.918     NaN 73.837
## 2909       4438  63.767  52.308  68.644  16.33638  56.354  50.769 57.062
## 2910       2897  47.705  36.000  45.673   9.67308  45.564  34.667 44.471
## 2911       8294  67.784  55.691  71.141  15.44988  64.782  53.659 68.624
## 2912       4244  76.909  59.130  80.000  20.86957  73.351  59.130 76.667
## 2913       2173  80.258  63.855  82.464  18.60903  73.033  59.036 72.512
## 2914       2448  76.348  76.389  80.707   4.31851  72.508  73.611 78.457
## 2915        671  59.911     NaN  63.208       NaN  56.781     NaN 56.604
## 2916      10226  58.165  52.967  65.878  12.91104  44.113  35.604 48.954
## 2917       1897  86.400  91.667  84.848  -6.81818  76.647  91.667 76.515
## 2918          0      NA      NA      NA        NA      NA      NA     NA
## 2919       2864  86.557  86.517  86.544   0.02749  80.028  87.640 78.899
```

```
## 2920       1109  55.636  75.000  53.271 -21.72897  54.734  62.500 49.533
## 2921       4069  67.855  74.317  68.857  -5.45980  62.792  71.038 60.286
## 2922        541  85.952 100.000  92.857  -7.14286  75.601 100.000 81.429
## 2923       1322  93.419     NaN  94.030       NaN  90.545     NaN 89.552
## 3000       3186  40.176  32.161  47.101  14.94065  32.298  23.116 40.217
## 3001        727  36.314  35.897  45.000   9.10256  31.774  29.487 21.667
## 3002      17595  22.575  14.497  28.238  13.74130  14.896  13.905 15.199
## 3003      13638  27.856  22.374  32.130   9.75553  19.468  16.743 21.931
## 3004      36902  34.394  26.387  41.359  14.97199  22.755  18.358 26.787
## 3005      19524  37.800  28.617  48.319  19.70161  26.665  21.743 33.480
## 3100      26497  11.228   9.302  13.468   4.16550   8.843   6.977 11.723
## 3101      11439   9.686   8.481  12.791   4.31013   8.244   6.007  9.136
## 3102       5505  17.620  16.332  19.101   2.76947  14.151  15.075 10.393
## 3103      14354  17.131  11.919  19.355   7.43623  13.083  10.562 14.776
## 3104      42113  13.737  11.036  16.515   5.47926   9.816   6.554 12.755
## 3105       8572  14.022  12.127  16.156   4.02923  10.044   8.946 12.925
## 3106      18628  10.473   7.359  11.817   4.45854   8.605   8.016  8.342
## 3107      16821  11.301   7.717  14.555   6.83797   8.091   6.660 10.644
## 3108       5826  19.653  18.625  24.892   6.26713  13.680  12.894 19.481
## 3109       7176   8.891   7.619  12.107   4.48749   7.302   6.190  8.232
## 3110       6785  18.379  13.068  22.706   9.63824  12.395   9.659 17.890
## 3200      27714  40.499  35.367  46.616  11.24835  38.385  34.184 44.378
## 3201       3663  36.719  31.048  35.968   4.91999  36.555  33.468 36.759
## 3202      20468  43.214  36.973  48.011  11.03793  40.620  35.243 46.866
## 3203      32950  47.199  37.918  53.937  16.01893  41.791  36.927 45.556
## 3204      37929  56.743  48.920  60.692  11.77207  51.467  46.811 54.110
## 3205      20085  38.531  32.936  43.238  10.30180  34.693  29.106 37.841
## 3206       7960  39.146  39.950  42.341   2.39104  32.613  30.653 41.136
## 3207      26328  48.276  43.573  53.214   9.64090  45.864  44.120 48.730
## 3300      12079  34.307  29.619  39.210   9.59152  34.953  32.405 38.476
## 3301      20773  48.683  43.562  50.409   6.84662  47.186  43.175 49.150
## 3302       4406  33.454  25.000  36.412  11.41161  32.978  25.893 35.356
## 3303      10360  31.602  24.708  37.241  12.53321  31.467  25.097 34.483
## 3304      85932  49.377  42.388  53.135  10.74694  48.214  42.054 51.913
## 3305      25049  29.814  21.341  34.535  13.19307  29.135  21.748 34.084
## 3306      21720  54.282  47.243  59.334  12.09119  49.641  43.358 54.028
## 3307       3217  28.567  23.077  37.879  14.80186  25.676  21.154 33.838
## 3308      27983  33.395  27.464  35.288   7.82335  33.449  27.163 35.472
## 3309      19257  50.828  44.924  52.929   8.00417  49.800  43.952 51.357
## 3310      22767  50.288  44.305  52.938   8.63327  49.291  40.648 50.853
## 3311       2327  23.550  20.000  27.304   7.30375  23.807  14.545 28.669
## 3312      62635  36.738  31.747  40.336   8.58952  36.460  32.003 39.657
## 3313       2865  19.721  13.571  22.269   8.69748  18.743  18.571 22.689
## 3314      15415  46.623  41.348  49.379   8.03148  45.787  40.582 46.384
## 3315      14336  45.968  41.591  50.485   8.89349  44.908  40.145 49.838
## 3316      20270  42.654  36.893  45.857   8.96339  42.570  35.761 44.955
## 3317       4254  50.799  39.819  55.046  15.22687  48.966  43.891 47.401
## 3318       8565  72.528  67.630  72.539   4.90880  70.893  64.451 71.891
## 3319       5446  56.849  50.186  60.164   9.97840  52.607  48.327 53.183
## 3320      16741  22.149  17.978  26.913   8.93557  22.388  17.790 26.719
```

```
## 3321       8313  34.717  27.003  38.692  11.68913  34.560  27.893 39.782
## 3322       6206  34.096  25.879  43.846  17.96756  34.434  28.754 40.577
## 3400      16604  46.073  36.830  55.989  19.15875  33.721  29.799 38.385
## 3401        196  44.388  50.000  60.000  10.00000  43.367  50.000 66.667
## 3402        523  49.331  40.909  45.455   4.54545  39.006  25.000 45.455
## 3403       1117  40.107  30.645  48.864  18.21848  28.021  16.129 35.227
## 3404       1810  52.486  40.800  52.846  12.04553  40.497  32.800 42.276
## 3500       6868  24.228  20.427  28.760   8.33307  19.642  18.293 26.121
## 3501         42  40.476     NaN  33.333       NaN  21.429     NaN 14.286
## 3502          0      NA      NA      NA        NA      NA      NA     NA
## 3503          2   0.000     NaN     NaN       NaN   0.000     NaN    NaN
## 3504       2986  17.850  12.054  23.370  11.31599  13.530  10.714 18.478
## 3505        596  21.141   0.000  13.636  13.63636  15.436  50.000 18.182
## 3506        608  25.000  13.514  46.154  32.64033  16.776  13.514 19.231
## 3600       5189  45.982  37.255  57.732  20.47706  43.322  39.542 58.763
## 3601       1456  63.736  58.750  69.697  10.94697  55.701  62.500 53.030
## 3602        566  26.148  14.286  31.707  17.42160  28.269  14.286 36.585
## 3603        740  56.351  61.111  45.161 -15.94982  54.054  55.556 35.484
## 3604        201  24.876     NaN     NaN       NaN  19.403     NaN    NaN
## 3605        956  42.469  25.000  53.459  28.45912  35.669  25.000 41.509
## 3606        208  60.096     NaN  68.421       NaN  59.135     NaN 68.421
## 3607       2727  34.177  26.846  39.303  12.45785  25.119  13.423 24.378
## 3608        123  57.724     NaN     NaN       NaN  56.911     NaN    NaN
## 3609       1368  75.585  79.787  81.818   2.03095  69.079  80.851 69.091
## 3610       1045  38.660  23.529  56.667  33.13725  33.014  23.529 40.000
## 3611       1497  42.953  26.316  50.235  23.91895  41.216  10.526 45.070
## 3612      16603  36.150  30.693  41.661  10.96753  31.585  27.723 36.781
## 3613        498  17.068  11.111  22.807  11.69591  17.068  16.667 26.316
## 3614       7421  24.161  26.023  32.715   6.69123  17.733  16.959 22.274
## 3615          0      NA      NA      NA        NA      NA      NA     NA
## 3616       6712  60.667  54.790  62.681   7.89074  55.781  49.701 56.884
##           ChgN 1CtyMF96 1CtyMF14
## 1000   6.18888     1360     1540
## 1100   6.15217      925     3972
## 1101   4.90654      642      264
## 1102   3.09701      841     1091
## 1103   5.89369     1453     1648
## 1104   4.65393     1561     2036
## 1105   4.11532     2998     5113
## 1106  11.02065      790     1093
## 1107   8.16282      635      843
## 1108  12.30024      295      336
## 1109   1.44246     1009     1075
## 1110   5.83009     1416     1444
## 1111   5.53213      484      588
## 1200  18.94144       60      296
## 1201   4.95678     1714     2550
## 1202  10.39490      723     2005
## 1203  10.42569      865     1866
## 1204   0.35208       89      367
```

```
## 1205   6.22819       25      149
## 1206  25.64103        9       39
## 1207   8.85200      301      768
## 1208   5.34156      307     1859
## 1209   9.76471       34      100
## 1210   3.49524      109      326
## 1211   2.20386      820     1982
## 1212   0.27005      446     1086
## 1213   2.51450      244      916
## 1300   8.91217     1547     4786
## 1301       NaN        0       36
## 1302  -1.75889      220      368
## 1303   4.57251     7362     6056
## 1304   5.82378     1423     1505
## 1305   6.08926     1142     1511
## 1306  10.61505     1851     2829
## 1307   6.19742     4791     3708
## 1308   3.95069     1164     1148
## 1309  10.67657     1268     1053
## 1310  12.01764     1089     1196
## 1311   5.58766     3368     3387
## 1312   5.31774     5991     5774
## 1313   5.88284     1037     1777
## 1314   3.21984     2172     2276
## 1315   2.71477      498      537
## 1400   3.41621      629      705
## 1401  -4.12479      309      342
## 1402  10.86633      481      734
## 1403   4.05262      612     1273
## 1404   4.74227      200      194
## 1405   7.55140      679      697
## 1406   4.49927      567      828
## 1407  13.91413      568      930
## 1408   9.38667     1166     1556
## 1409  -1.54594      152      303
## 1410  10.58012      172      182
## 1500   7.37297     1066     1101
## 1501   3.37662       35       11
## 1502   7.42599      438     1181
## 1503   3.28947      912     1200
## 1504  19.20341       38       74
## 1505   4.94207      690      718
## 1506  -6.10156       74       66
## 1507   3.54839      155      124
## 1508  -3.82082       69       99
## 1600   6.49146     2578     3574
## 1601  13.39379       61       47
## 1602   6.96302      904     1032
## 1603   4.31038      266      509
## 1604   5.28025      965      856
```

```
## 1605   0.16670     2335     2118
## 1606   5.67418     2409     2540
## 1607   5.75496      576      700
## 1700   4.84606      587      466
## 1701 -39.31034        5       58
## 1702  10.18665      585      544
## 1703   3.62570      484      682
## 1704   4.34326      403      298
## 1705   8.91312      595      605
## 1706   5.78299     1421     2008
## 1707   8.03682      270      173
## 1708   1.39353      552      390
## 1709  19.78455      150      328
## 1710   7.61578      787      799
## 1711   3.39351      386      283
## 1712   2.76610     1335      965
## 1800   8.40592      112      164
## 1801       NaN        0       55
## 1802   1.62692      239      206
## 1803   5.50511      651      612
## 1804   1.96145      451      533
## 1900   2.87674      695      602
## 1901   8.47801      654      585
## 1902   9.32624      920     1128
## 1903   4.49983      107      162
## 1904   8.67873      641      842
## 1905  16.80420       31       43
## 1906   7.57225      989      869
## 1907   3.21613      547      515
## 1908   8.23454      921      932
## 1909   0.36456      234      211
## 1910   6.91803      915      750
## 1911   4.80692      522      445
## 1912   5.18364     1613     1516
## 1913  11.23894       30      113
## 2000  -1.82881      168      442
## 2001  -7.97509      116      227
## 2002   5.60135     2616     3252
## 2003   4.01848      944     1267
## 2100   1.49888       38      539
## 2101  10.29412       22       68
## 2102   3.24212      422      553
## 2103   2.26855      344      418
## 2104  -0.05149      183      329
## 2105   2.95120      265      725
## 2200   7.66911     2045      816
## 2201  14.66356      167      211
## 2202   2.81498      632      602
## 2203   4.46187       91      199
## 2204   4.87539      519     1487
```

```
## 2205   5.17473      744      768
## 2206   6.45120       76      103
## 2207   2.65114      596      502
## 2208   4.44957     2974     2581
## 2209   1.72316       50      708
## 2210   3.62196     1211     1854
## 2211   6.64866      730     1278
## 2212   5.43972      255      202
## 2213   9.36828      316      636
## 2214  19.15905      131       61
## 2215   5.22657      249      424
## 2216  12.96610       59      120
## 2300   8.01348     1091     1042
## 2301  -0.36040      171      258
## 2302   5.20279      141      256
## 2303   9.46954     1754     2924
## 2304   7.47250      936     1219
## 2305   6.68533      553      598
## 2306  12.60523      158      406
## 2307  15.29466      676      888
## 2308  10.36763      632     1429
## 2309   9.26125      427      900
## 2310   2.75848      480      806
## 2311   3.19129      374      573
## 2312   8.22181     1016     1205
## 2400  13.37480      240      844
## 2401       NaN        0       24
## 2402  10.96402      452      549
## 2403   5.37130     2755     2759
## 2404   8.12347     1313     1650
## 2405   7.08523      484      588
## 2406   4.96252      803     1166
## 2500   7.93874      912     2245
## 2501  -9.75936       68       11
## 2502   8.59543      244      847
## 2503   9.02216      421      539
## 2504   5.68752      947     1724
## 2505   6.06317     1583     1596
## 2506   9.67971      374      475
## 2507   2.41514      749      671
## 2508   6.31398      657     1139
## 2600   1.30121     1029     1312
## 2601  -1.05263       75      114
## 2602   1.64194      363      503
## 2603   2.44962      288      419
## 2604   0.74702     1479     1901
## 2605   4.18427      291      454
## 2606  10.19374       97      116
## 2607   3.84615      156      264
## 2608   5.03190      212      343
```

```
## 2609  -1.76946       86       92
## 2610   0.32679      353      267
## 2611  10.34193      503      955
## 2612   3.28283      132      198
## 2613   2.62764      816     1318
## 2614   1.08552      606      441
## 2700  10.29133     2660     8460
## 2701   8.31540     1030     1924
## 2702   7.49204      342      281
## 2703   7.68320      522      922
## 2704   7.60545      228      282
## 2705   4.16337     2459     3835
## 2706   9.81327      730      961
## 2707  13.36521      142      327
## 2708  14.04482      655     1089
## 2709  -1.11111       18       20
## 2710  24.78971       73      114
## 2711  10.14498      416     1162
## 2712  11.27959      824     1619
## 2713  13.48581      601     1468
## 2714   4.60717       95      345
## 2715   0.44871      467     1110
## 2716   7.13801      526     1110
## 2717  -3.80630      713     1166
## 2718  -4.50188      115      508
## 2719   7.04491      554     2327
## 2720   6.75557      571     1252
## 2721   4.11601      198      501
## 2722   4.05861      232      363
## 2723   6.95628     1123     1512
## 2724   9.05642      674      844
## 2725   8.11137     1176     2330
## 2726   5.59508      690      990
## 2727  -1.19909      370      568
## 2728   7.86691     2427     3900
## 2729  18.80203      984     1676
## 2730  11.47061     1949     4052
## 2731   5.02005      986     1294
## 2732   3.86509     1418     2984
## 2733   8.25562      804     1130
## 2734  11.32975     1223     1505
## 2735   9.94079     2068     3826
## 2736   7.19328     1013     1809
## 2737   3.83570     1403     1566
## 2738   8.11036     3014     5235
## 2739  12.60948     2599     5316
## 2740   6.38910      870     1427
## 2741   2.87970     2023     2950
## 2742   5.24180      687     1253
## 2743  19.21491      248      438
```

```
## 2744       NaN        0        0
## 2745   7.86821      243      334
## 2746   4.65274     2829     5699
## 2747  10.31149      284      527
## 2748   8.54863      498     1164
## 2800   6.42904     2013     2196
## 2801   5.11392       25      237
## 2802  11.20372      723     1041
## 2803  10.91659      382      576
## 2804   4.03147      923     1283
## 2805   1.08712      341     1089
## 2806   7.04982      313      364
## 2807  13.45555       98      177
## 2808   8.22025      744     1790
## 2809  -2.25363      633      727
## 2900   3.34423      881     1139
## 2901  -2.81369       35      263
## 2902  15.96897      337      764
## 2903  -0.31173       92      265
## 2904 -35.93690       23      113
## 2905   2.24657       72      251
## 2906  -5.90408      112      382
## 2907   0.86631      134      373
## 2908       NaN        0      344
## 2909   6.29292      195      354
## 2910   9.80449       75      416
## 2911  14.96562      246      596
## 2912  17.53623      115      420
## 2913  13.47570       83      211
## 2914   4.84548       72      311
## 2915       NaN        0      106
## 2916  13.34920      455     1099
## 2917 -15.15152       36      264
## 2918        NA       NA       NA
## 2919  -8.74137       89      327
## 2920 -12.96729       24      107
## 2921 -10.75254      183      350
## 2922 -18.57143        1       70
## 2923       NaN        0      134
## 3000  17.10181      199      276
## 3001  -7.82051       78       60
## 3002   1.29329     1014     1158
## 3003   5.18864      657     1108
## 3004   8.42898     2217     2546
## 3005  11.73693     1251     1368
## 3100   4.74627      129     1834
## 3101   3.12915      566      602
## 3102  -4.68212      398      356
## 3103   4.21426     1032      961
## 3104   6.20071     2365     3191
```

```
## 3105   3.97885      503      588
## 3106   0.32584      761     1007
## 3107   3.98484      946     1381
## 3108   6.58654      349      462
## 3109   2.04197      420      413
## 3110   8.23082      352      436
## 3200  10.19342     1606     1832
## 3201   3.29115      248      253
## 3202  11.62324      925     1835
## 3203   8.62904     1614     2959
## 3204   7.29936     1991     3236
## 3205   8.73497     1175     1538
## 3206  10.48271      398      581
## 3207   4.60983     1097     2520
## 3300   6.07097      682     1089
## 3301   5.97519     1033     1589
## 3302   9.46334      112      379
## 3303   9.38548      514      725
## 3304   9.85863     3291     7607
## 3305  12.33612      984     1998
## 3306  10.66958      798     2073
## 3307  12.68454      156      198
## 3308   8.30870     1329     2712
## 3309   7.40466      926     1400
## 3310  10.20522      957     2110
## 3311  14.12349       55      293
## 3312   7.65378     2731     5593
## 3313   4.11765      140      238
## 3314   5.80229      653     1369
## 3315   9.69373      553     1547
## 3316   9.19440      618     2329
## 3317   3.50921      221      327
## 3318   7.44032      346      772
## 3319   4.85561      269      487
## 3320   8.92829      534     1542
## 3321  11.88884      337      734
## 3322  11.82293      313      520
## 3400   8.58565      896     1102
## 3401  16.66667        4       15
## 3402  20.45455       44       44
## 3403  19.09824       62       88
## 3404   9.47642      125      123
## 3500   7.82869      328      379
## 3501       NaN        0       21
## 3502        NA       NA       NA
## 3503       NaN        0        0
## 3504   7.76398      224      184
## 3505 -31.81818        2       44
## 3506   5.71726       74       26
## 3600  19.22040      306       97
```

```
## 3601  -9.46970       80      132
## 3602  22.29965        7       41
## 3603 -20.07168       36       62
## 3604       NaN        0        0
## 3605  16.50943       24      159
## 3606       NaN        0       19
## 3607  10.95529      149      201
## 3608       NaN        0        0
## 3609 -11.76015       94      110
## 3610  16.47059       51       60
## 3611  34.54411       19      213
## 3612   9.05800      808     1373
## 3613   9.64912       18       57
## 3614   5.31472      342      431
## 3615        NA       NA       NA
## 3616   7.18346      334      552
```

```r
print("GenderTeamSize in 2018")
```

```
## [1] "GenderTeamSize in 2018"
```

```r
print(GenderTeamSize)
```

```
##      Articles FirstF FirstM    FirstP LastF LastM     LastP
## 1000     2403  4.748  4.566 2.259e-01 4.608 4.633 8.269e-01
## 1100     2852  4.418  3.914 1.056e-05 4.207 4.089 2.913e-01
## 1101      264  2.709  2.281 2.483e-02 2.739 2.288 6.891e-02
## 1102     1156  3.656  3.603 7.415e-01 3.420 3.684 1.563e-01
## 1103     1729  3.411  3.020 6.192e-05 3.147 3.236 4.563e-01
## 1104     1944  3.737  3.499 1.702e-01 3.734 3.537 3.487e-02
## 1105     4781  3.477  3.171 2.157e-06 3.372 3.258 1.212e-01
## 1106     1082  3.777  3.560 2.496e-01 3.614 3.701 6.674e-01
## 1107      792  3.641  3.418 6.253e-01 3.629 3.448 2.581e-01
## 1108      260  3.385  3.172 3.605e-01 3.105 3.312 3.514e-01
## 1109      948  3.714  3.138 9.257e-04 3.715 3.241 2.874e-03
## 1110     1261  3.553  3.276 4.998e-02 3.354 3.383 7.436e-01
## 1111      664  4.060  3.397 2.524e-03 3.676 3.564 8.384e-01
## 1200      229  1.197  1.256 8.204e-01 1.272 1.179 2.132e-01
## 1201     2187  2.224  2.044 4.467e-03 2.081 2.175 1.183e-01
## 1202     2028  1.132  1.100 9.662e-02 1.122 1.107 9.175e-01
## 1203     1689  1.641  1.517 4.890e-03 1.598 1.574 4.915e-01
## 1204      438  1.663  1.758 6.427e-01 1.510 1.865 1.405e-03
## 1205       91  1.019  1.026 8.337e-01 1.038 1.013 3.605e-01
## 1206       91  1.698  1.228 1.164e-02 1.435 1.658 2.845e-01
## 1207      737  1.741  1.295 8.860e-12 1.671 1.339 4.193e-07
## 1208     1439  1.066  1.056 4.767e-01 1.069 1.053 2.659e-01
## 1209      106  1.436  1.423 8.469e-01 1.369 1.524 2.486e-01
## 1210      307  1.220  1.180 1.698e-01 1.186 1.209 7.082e-01
## 1211     2003  1.315  1.135 9.567e-11 1.326 1.132 1.945e-10
## 1212     1275  1.258  1.169 1.408e-02 1.276 1.160 1.749e-03
## 1213      955  1.129  1.105 1.742e-01 1.112 1.127 8.030e-01
```

```
## 1300     4679  4.848  4.589 8.992e-03 4.562 4.747 8.491e-02
## 1301      115  3.940  4.076 7.300e-01 3.704 4.174 7.764e-01
## 1302      296  5.421  5.493 7.281e-01 5.485 5.446 6.235e-01
## 1303     3550  4.783  4.291 9.634e-06 4.524 4.443 4.123e-01
## 1304     1209  4.240  4.025 4.103e-01 4.221 4.052 7.302e-01
## 1305     1392  4.445  3.955 2.532e-03 3.930 4.161 1.525e-01
## 1306     2638  7.316  7.066 6.617e-01 7.003 7.266 1.880e-02
## 1307     2514  5.157  4.938 2.969e-01 4.647 5.188 8.335e-04
## 1308      727  5.124  5.150 7.883e-01 4.855 5.268 7.484e-02
## 1309      841  4.880  4.409 5.539e-02 4.601 4.673 6.634e-01
## 1310      962  5.136  5.124 5.065e-01 5.053 5.187 4.053e-01
## 1311     3201  5.053  4.463 4.585e-08 4.666 4.742 9.160e-01
## 1312     4360  5.086  4.640 1.704e-05 4.837 4.817 8.203e-01
## 1313     1383  5.838  6.050 3.087e-01 5.875 5.993 4.206e-01
## 1314     2012  4.520  4.296 2.319e-01 4.293 4.436 3.370e-01
## 1315      445  4.088  4.052 9.806e-01 4.386 3.982 7.923e-02
## 1400      597  1.817  1.910 3.130e-01 1.835 1.903 4.319e-01
## 1401      260  1.890  1.809 6.757e-01 1.888 1.806 6.092e-01
## 1402      588  1.865  1.911 5.704e-01 1.861 1.912 6.217e-01
## 1403     1108  2.028  1.868 5.610e-02 1.913 1.930 4.970e-01
## 1404      180  2.154  2.133 8.816e-01 2.043 2.168 5.308e-01
## 1405      716  1.995  1.849 9.137e-02 1.892 1.888 9.326e-01
## 1406      707  2.049  2.033 8.449e-01 1.924 2.118 7.117e-03
## 1407      777  2.186  2.070 2.465e-01 2.134 2.109 8.958e-01
## 1408     1580  2.108  1.881 5.420e-04 1.999 1.937 5.408e-01
## 1409      336  1.937  2.080 2.628e-01 1.877 2.113 5.847e-02
## 1410      211  1.657  1.594 7.791e-01 1.669 1.587 6.558e-01
## 1500     1240  3.971  3.459 1.678e-03 3.807 3.525 1.685e-01
## 1501       79  2.599  3.525 1.129e-01 2.185 3.655 5.693e-03
## 1502      986  4.915  4.586 1.819e-01 4.372 4.751 4.751e-02
## 1503     1080  4.810  4.343 1.055e-02 4.532 4.450 9.625e-01
## 1504       52  4.619  3.584 9.980e-02 3.874 4.217 6.452e-01
## 1505      462  4.787  4.561 4.600e-01 4.864 4.572 6.636e-01
## 1506       52  5.495  3.242 2.111e-02 3.987 3.430 6.813e-01
## 1507      321  3.149  2.735 3.412e-02 2.538 2.848 2.556e-01
## 1508      249  3.902  3.540 3.677e-01 3.276 3.687 1.582e-01
## 1600     3590  4.653  4.241 2.193e-04 4.199 4.395 4.210e-02
## 1601      178  3.483  3.498 7.729e-01 2.777 3.725 6.318e-03
## 1602      853  4.157  3.495 8.648e-04 3.942 3.641 2.269e-01
## 1603      275  4.379  3.938 1.599e-01 4.266 4.016 4.845e-01
## 1604      777  4.620  4.098 3.960e-03 4.643 4.134 9.770e-03
## 1605     1343  4.405  4.225 2.893e-01 4.130 4.306 2.896e-01
## 1606     2335  3.823  3.394 1.094e-04 3.869 3.427 4.515e-04
## 1607      768  4.282  3.624 7.347e-04 4.126 3.736 6.248e-02
## 1700      545  2.223  2.112 5.601e-01 2.196 2.118 4.999e-01
## 1701       63  2.680  2.543 6.751e-01 2.698 2.532 4.858e-01
## 1702      102  2.370  2.297 9.410e-01 2.250 2.327 9.461e-01
## 1703      465  2.968  2.505 3.386e-02 3.046 2.487 6.841e-03
## 1704      225  2.640  2.720 8.471e-01 2.051 2.811 1.294e-02
## 1705      269  2.173  2.511 1.354e-01 2.210 2.513 1.767e-01
```

```
## 1706     1106  2.632  2.799 2.781e-01 2.577 2.817 1.178e-01
## 1707      177  3.240  2.819 4.478e-01 2.600 2.935 3.759e-01
## 1708      197  3.042  2.734 6.123e-01 2.948 2.747 4.394e-01
## 1709      261  2.382  2.394 8.678e-01 2.137 2.582 4.048e-02
## 1710      630  2.130  2.571 1.571e-03 2.022 2.630 3.679e-06
## 1711      292  3.194  2.633 2.021e-02 2.775 2.711 6.457e-01
## 1712      648  2.715  2.640 5.806e-01 2.557 2.668 7.255e-01
## 1800      163  2.466  2.216 3.907e-01 2.197 2.311 4.888e-01
## 1801       15  2.352  1.943 3.618e-01 1.861 2.153 6.269e-01
## 1802      200  2.244  2.268 6.334e-01 2.257 2.264 8.557e-01
## 1803      555  2.137  2.163 8.555e-01 2.054 2.185 2.554e-01
## 1804      545  2.722  2.013 3.997e-06 2.527 2.074 1.603e-03
## 1900      678  3.135  2.924 2.250e-01 2.808 3.039 2.815e-01
## 1901      433  3.275  2.998 4.152e-01 2.938 3.110 6.269e-01
## 1902     1209  3.776  3.223 2.806e-04 3.363 3.395 8.611e-01
## 1903      164  3.492  3.316 6.530e-01 3.521 3.312 7.704e-01
## 1904      841  2.945  2.664 1.503e-01 2.688 2.784 4.837e-01
## 1905       60  3.378  2.266 9.229e-02 2.681 2.411 4.465e-01
## 1906      571  3.539  3.057 4.405e-02 3.468 3.132 1.454e-01
## 1907      606  3.346  2.904 1.373e-02 3.001 3.022 5.604e-01
## 1908      613  3.394  2.897 2.136e-02 3.231 2.991 3.067e-01
## 1909      339  2.711  2.705 9.282e-01 2.713 2.705 8.230e-01
## 1910      644  3.873  3.296 2.952e-02 3.841 3.393 3.856e-02
## 1911      429  3.613  2.962 1.477e-02 3.328 3.134 3.882e-01
## 1912     1020  3.755  2.926 9.475e-07 3.553 3.014 2.096e-03
## 1913       96  4.291  2.745 8.036e-04 3.273 3.109 8.043e-01
## 2000      398  1.566  1.499 4.606e-01 1.561 1.500 4.726e-01
## 2001      293  1.676  1.662 9.695e-01 1.574 1.703 1.996e-01
## 2002     2981  1.890  1.738 1.209e-03 1.827 1.758 1.853e-01
## 2003     1042  1.942  1.776 3.488e-02 1.946 1.776 1.910e-02
## 2100      545  3.414  3.537 5.178e-01 3.236 3.571 1.767e-01
## 2101      119  2.296  2.203 7.554e-01 2.198 2.230 9.436e-01
## 2102      703  3.302  3.126 2.110e-01 3.293 3.131 3.559e-01
## 2103      502  3.256  3.243 7.982e-01 3.084 3.277 6.096e-01
## 2104      390  3.079  3.569 1.966e-01 3.220 3.527 5.987e-01
## 2105     1033  3.266  3.298 9.673e-01 2.949 3.389 7.465e-03
## 2200      431  3.123  2.538 1.252e-02 2.694 2.654 8.313e-01
## 2201      282  2.618  2.555 7.691e-01 2.286 2.637 1.735e-01
## 2202      515  3.036  2.490 7.294e-03 2.772 2.526 2.486e-01
## 2203      167  2.962  2.594 3.239e-01 2.558 2.668 7.659e-01
## 2204     1565  4.661  4.355 3.092e-02 4.291 4.496 1.565e-01
## 2205     1120  2.842  2.640 6.123e-02 2.800 2.648 1.486e-01
## 2206      134  2.994  2.492 2.305e-01 2.850 2.499 3.019e-01
## 2207      250  2.622  2.555 6.870e-01 2.477 2.573 7.688e-01
## 2208      891  3.384  3.056 1.102e-01 3.670 3.029 6.113e-04
## 2209      834  2.761  2.672 4.010e-01 2.495 2.726 1.362e-01
## 2210     2030  3.163  2.944 9.304e-02 3.084 2.960 3.088e-01
## 2211     1385  3.532  3.079 5.407e-03 3.682 3.065 1.226e-04
## 2212      284  3.813  2.792 1.114e-04 4.061 2.860 1.820e-04
## 2213      618  2.870  2.540 2.415e-02 2.702 2.593 5.079e-01
```

```
## 2214       74  1.848  1.750 7.019e-01 1.765 1.801 8.426e-01
## 2215      611  2.975  2.580 6.845e-02 2.767 2.618 7.668e-01
## 2216      160  1.225  1.468 2.516e-02 1.244 1.455 3.842e-02
## 2300     1094  3.368  2.909 3.910e-04 3.043 3.105 6.867e-01
## 2301      405  2.445  2.212 1.381e-01 2.231 2.340 5.712e-01
## 2302      240  3.734  3.307 1.289e-01 3.687 3.359 2.249e-01
## 2303     3100  3.565  3.383 5.736e-02 3.526 3.428 2.049e-01
## 2304     1304  4.242  4.027 2.902e-01 4.050 4.126 5.784e-01
## 2305      712  4.158  3.263 4.502e-07 4.109 3.356 9.692e-05
## 2306      571  3.135  3.099 7.695e-01 3.109 3.115 9.323e-01
## 2307     1130  4.760  3.970 2.121e-05 4.644 4.178 8.788e-03
## 2308     1684  2.764  2.592 6.681e-02 2.660 2.650 8.159e-01
## 2309     1134  3.382  3.275 6.762e-01 3.396 3.288 4.512e-01
## 2310      980  4.101  3.813 1.916e-01 4.020 3.877 4.116e-01
## 2311      614  3.986  3.677 1.664e-01 3.756 3.794 7.803e-01
## 2312     1452  3.540  3.208 9.220e-02 3.453 3.262 1.475e-01
## 2400     1160  4.537  3.974 1.556e-03 4.060 4.257 2.154e-01
## 2401       30  5.467  5.368 9.163e-01 4.500 6.387 1.746e-01
## 2402      428  4.549  4.664 5.326e-01 4.133 4.799 8.147e-02
## 2403     2289  5.911  5.654 4.403e-01 5.933 5.713 2.740e-01
## 2404     1715  5.100  4.751 3.282e-02 4.804 4.985 3.514e-01
## 2405      543  5.114  4.711 3.875e-01 5.053 4.851 8.281e-01
## 2406      922  5.416  5.176 4.634e-01 5.160 5.361 4.091e-01
## 2500     2653  4.198  3.761 3.823e-04 3.937 3.829 4.749e-01
## 2501      134  2.513  3.354 1.584e-02 2.608 3.199 1.109e-01
## 2502      736  5.310  4.735 3.607e-02 5.145 4.855 4.595e-01
## 2503      532  4.503  3.851 1.030e-02 4.694 3.865 6.361e-03
## 2504     1135  4.247  3.600 6.804e-04 4.365 3.595 1.000e-04
## 2505      721  3.996  3.610 1.824e-01 3.902 3.661 2.998e-01
## 2506      523  3.777  3.129 6.435e-03 3.806 3.150 5.953e-03
## 2507      709  3.926  3.596 7.945e-02 3.785 3.640 4.956e-01
## 2508     1048  3.811  3.663 6.366e-01 3.757 3.684 5.371e-01
## 2600     1095  1.947  1.461 2.425e-08 1.867 1.470 6.430e-07
## 2601      115  1.709  1.435 1.665e-01 1.638 1.454 2.867e-01
## 2602      462  1.782  1.383 1.345e-04 1.784 1.388 3.644e-04
## 2603      375  2.091  1.458 3.189e-04 1.877 1.483 2.445e-02
## 2604     1595  2.305  1.989 1.167e-04 2.311 1.993 1.876e-03
## 2605      467  2.824  2.310 8.449e-03 2.754 2.327 7.054e-02
## 2606      220  2.802  2.300 7.619e-02 2.441 2.359 5.192e-01
## 2607      304  2.199  1.680 1.506e-03 2.295 1.668 1.224e-04
## 2608      364  1.699  1.407 1.929e-02 1.755 1.401 1.260e-02
## 2609       87  1.347  1.174 1.389e-01 1.335 1.180 5.059e-01
## 2610      286  2.206  1.803 5.339e-02 1.857 1.841 9.188e-01
## 2611     1039  3.004  2.618 2.439e-03 2.986 2.630 1.480e-03
## 2612      133  1.905  1.656 2.792e-01 1.724 1.686 9.129e-01
## 2613     1198  2.849  2.187 4.590e-09 2.686 2.239 1.756e-04
## 2614      377  2.566  1.982 1.876e-03 2.104 2.059 9.632e-01
## 2700     2871  3.751  3.571 4.645e-02 3.471 3.769 9.570e-03
## 2701     1968  4.567  4.122 4.501e-03 4.337 4.393 4.906e-01
## 2702      375  3.742  3.390 1.839e-01 3.405 3.641 4.399e-01
```

```
## 2703      995  4.763  4.174 1.798e-02 4.759 4.265 4.168e-02
## 2704      205  5.800  4.738 6.898e-02 5.081 5.368 2.986e-01
## 2705     3885  5.569  5.575 5.234e-01 5.613 5.563 4.585e-01
## 2706      878  4.558  4.422 6.529e-01 4.410 4.507 6.498e-01
## 2707      306  2.972  2.541 8.385e-02 2.844 2.679 5.073e-01
## 2708     1258  3.892  3.706 2.072e-01 3.793 3.810 7.885e-01
## 2709        2  2.000  2.000       NaN 2.000 2.000       NaN
## 2710      104  4.776  4.358 5.116e-01 4.414 4.799 4.719e-01
## 2711     1365  3.858  4.060 1.928e-01 3.596 4.179 3.013e-05
## 2712     1381  5.677  5.336 3.805e-01 5.437 5.587 1.511e-01
## 2713     1523  5.366  4.557 5.425e-05 5.172 4.930 3.996e-01
## 2714      345  3.229  2.621 3.790e-03 2.870 2.930 8.130e-01
## 2715     1270  4.682  4.629 8.040e-01 4.322 4.735 2.012e-01
## 2716     1030  5.680  5.179 1.591e-02 5.548 5.397 2.702e-01
## 2717     1029  4.535  4.648 3.945e-01 4.294 4.865 6.140e-03
## 2718      870  4.602  4.745 9.569e-01 4.279 4.967 3.635e-03
## 2719     2703  3.751  3.059 5.302e-12 3.480 3.348 2.915e-01
## 2720     1225  5.627  5.780 4.025e-01 5.154 5.971 5.315e-03
## 2721      489  5.191  5.317 7.250e-01 4.471 5.511 1.836e-02
## 2722      449  4.381  4.049 2.087e-01 4.151 4.209 8.378e-01
## 2723     1787  5.801  5.705 8.361e-01 5.770 5.743 8.450e-01
## 2724      947  5.209  4.921 2.290e-01 4.666 5.307 2.520e-03
## 2725     2482  5.465  4.829 1.586e-05 5.254 5.078 3.379e-01
## 2726     1067  5.391  4.662 5.935e-04 5.106 4.931 2.767e-01
## 2727      457  5.409  4.946 2.591e-01 4.843 5.244 2.585e-01
## 2728     4895  5.244  5.193 4.281e-01 5.139 5.238 2.038e-01
## 2729     1876  4.466  4.174 1.667e-01 4.295 4.480 1.088e-01
## 2730     4514  6.447  6.636 1.554e-02 6.015 6.872 2.609e-12
## 2731     1254  4.366  4.026 1.101e-01 4.319 4.089 8.016e-02
## 2732     3858  4.326  4.520 1.076e-02 4.021 4.569 7.984e-08
## 2733     1436  4.201  4.078 4.528e-01 3.950 4.187 2.882e-02
## 2734     1531  3.697  3.610 5.685e-01 3.503 3.750 1.138e-01
## 2735     4837  4.320  4.376 2.642e-01 4.099 4.561 2.453e-10
## 2736     1981  4.707  4.416 6.046e-02 4.478 4.595 4.172e-01
## 2737     1393  4.894  5.054 1.434e-01 4.859 5.043 3.062e-01
## 2738     5813  4.163  3.788 4.063e-05 3.930 4.068 1.446e-03
## 2739     6127  4.206  3.766 4.726e-10 3.989 4.065 1.749e-01
## 2740     1628  5.465  5.467 2.831e-01 5.611 5.416 4.119e-01
## 2741     3482  4.860  4.905 9.703e-01 4.632 4.972 5.465e-02
## 2742     1076  3.835  4.043 2.517e-01 3.585 4.209 1.163e-04
## 2743      450  4.982  4.248 3.847e-02 4.566 4.866 7.267e-02
## 2744       NA     NA     NA        NA    NA    NA        NA
## 2745      291  5.366  5.250 6.476e-01 5.206 5.402 5.729e-01
## 2746     7242  4.991  4.849 6.859e-01 4.973 4.873 2.009e-01
## 2747      614  6.799  6.653 5.679e-01 6.581 6.766 7.363e-01
## 2748     1335  5.203  5.561 1.033e-03 5.307 5.467 1.479e-01
## 2800     2561  4.621  4.183 6.700e-04 4.403 4.350 7.649e-01
## 2801      259  5.243  4.518 1.060e-01 4.665 4.956 2.270e-01
## 2802     1092  4.005  3.783 1.483e-01 3.967 3.854 4.252e-01
## 2803      631  5.078  4.883 6.199e-01 4.853 5.070 1.284e-01
```

```
## 2804     1116  5.608  4.959 9.641e-04 5.293 5.214 9.394e-01
## 2805     1145  3.879  3.978 4.967e-01 3.751 4.054 9.996e-02
## 2806      335  4.655  4.012 7.816e-02 4.525 4.280 4.187e-01
## 2807      209  4.847  5.474 1.448e-01 4.747 5.353 2.616e-01
## 2808     1945  5.125  4.947 5.765e-01 5.073 5.002 5.994e-01
## 2809      578  4.555  3.858 4.511e-03 4.339 4.001 2.078e-01
## 2900     1385  2.845  3.229 4.896e-03 2.721 3.499 4.318e-09
## 2901      296  3.137  3.954 2.819e-02 2.951 4.509 1.911e-06
## 2902      656  2.420  3.113 2.271e-04 2.178 3.580 1.981e-14
## 2903      241  2.120  3.113 1.304e-03 1.969 3.403 4.873e-07
## 2904      135  4.027  4.050 9.659e-01 3.772 4.290 2.211e-01
## 2905      222  2.754  2.762 7.484e-01 2.520 3.487 2.045e-03
## 2906      355  2.085  2.391 2.236e-01 1.974 2.787 1.563e-04
## 2907      329  2.595  4.315 1.246e-08 2.439 4.444 4.754e-12
## 2908      263  2.596  3.976 2.659e-04 2.428 4.146 6.755e-08
## 2909      411  3.465  2.708 1.429e-03 3.335 3.036 1.945e-01
## 2910      279  1.858  1.500 5.376e-03 1.880 1.493 6.900e-03
## 2911      549  2.716  3.275 6.958e-03 2.575 3.572 2.822e-07
## 2912      296  2.023  2.080 7.099e-01 1.965 2.341 4.083e-02
## 2913      209  3.018  2.549 3.741e-01 2.927 3.002 7.941e-01
## 2914      199  2.197  3.078 3.480e-03 2.227 2.904 1.778e-02
## 2915       97  1.925  1.964 8.726e-01 1.765 2.278 1.567e-02
## 2916     1021  4.582  4.484 9.684e-01 4.606 4.492 4.925e-01
## 2917      228  3.903  5.062 7.615e-02 3.576 5.620 1.422e-04
## 2918       NA     NA     NA        NA    NA    NA        NA
## 2919      212  2.738  2.408 4.078e-01 2.605 3.142 1.053e-01
## 2920       32  2.128  3.120 1.280e-01 2.078 3.193 8.120e-02
## 2921      387  2.784  2.630 7.231e-01 2.649 2.870 2.933e-01
## 2922      102  2.429  1.732 3.307e-01 2.322 3.117 1.860e-01
## 2923      131  2.456  2.669 5.501e-01 2.482 2.525 8.245e-01
## 3000      324  3.676  3.478 6.864e-01 3.636 3.525 9.684e-01
## 3001       78  4.577  4.088 3.914e-01 4.589 4.140 5.750e-01
## 3002      840  5.233  5.916 1.701e-02 5.010 5.849 1.294e-02
## 3003      972  4.420  4.513 3.561e-01 4.416 4.496 8.498e-01
## 3004     2332  4.807  4.638 5.436e-01 4.484 4.813 9.846e-03
## 3005     1436  4.864  4.539 3.656e-02 4.729 4.693 7.582e-01
## 3100     2284  4.057  3.194 5.838e-09 3.911 3.254 7.049e-06
## 3101      661  3.119  2.627 1.841e-02 3.204 2.621 2.906e-02
## 3102      323  3.080  2.935 5.086e-01 2.908 2.983 8.361e-01
## 3103      989  3.511  2.810 7.977e-06 3.435 2.843 3.643e-04
## 3104     2651  3.616  3.124 1.907e-05 3.523 3.152 1.824e-03
## 3105      751  4.077  3.404 5.307e-04 3.780 3.489 1.827e-01
## 3106      578  2.757  2.655 4.403e-01 2.321 2.702 1.522e-01
## 3107     1333  3.853  3.236 1.092e-03 3.791 3.262 1.259e-02
## 3108      343  5.675  4.522 2.013e-02 6.674 4.440 5.769e-05
## 3109      399  2.300  2.105 5.147e-01 2.154 2.118 8.411e-01
## 3110      306  3.461  3.708 3.439e-01 3.693 3.638 6.023e-01
## 3200     1983  2.984  2.531 4.905e-09 2.835 2.697 7.826e-02
## 3201      252  2.720  2.360 8.339e-02 2.729 2.400 7.246e-02
## 3202     1925  3.182  2.695 9.777e-08 3.047 2.874 9.938e-02
```

```
## 3203     3179  3.629  3.200 5.336e-06 3.447 3.444 7.414e-01
## 3204     3748  3.221  2.879 2.302e-06 3.071 3.149 8.554e-02
## 3205     1686  3.035  2.642 2.715e-06 2.954 2.745 5.908e-03
## 3206      592  3.735  3.273 3.126e-03 3.350 3.610 1.057e-01
## 3207     2539  2.759  2.489 5.594e-05 2.680 2.599 1.633e-01
## 3300     1039  1.654  1.475 1.347e-03 1.608 1.510 4.388e-02
## 3301     1795  2.256  1.923 3.947e-08 2.199 2.008 1.559e-03
## 3302      468  1.558  1.696 2.449e-01 1.459 1.765 1.838e-03
## 3303      781  1.680  1.659 7.807e-01 1.672 1.665 9.859e-01
## 3304     8051  2.283  2.081 3.648e-11 2.234 2.151 6.292e-03
## 3305     1995  1.893  1.771 3.396e-02 1.824 1.814 8.718e-01
## 3306     2628  3.230  2.826 1.189e-05 3.060 3.100 6.028e-01
## 3307      234  2.796  2.463 1.383e-01 2.790 2.492 1.603e-01
## 3308     2454  1.882  1.615 6.566e-09 1.823 1.658 3.028e-04
## 3309     1088  1.742  1.850 1.267e-01 1.732 1.869 7.963e-02
## 3310     1945  1.723  1.543 2.594e-05 1.681 1.600 3.566e-02
## 3311      344  2.796  2.403 4.182e-02 2.654 2.473 3.314e-01
## 3312     5837  1.857  1.569 1.296e-24 1.817 1.604 1.292e-13
## 3313      341  2.509  2.416 5.064e-01 2.424 2.444 8.957e-01
## 3314     1452  1.599  1.523 1.159e-01 1.567 1.560 9.490e-01
## 3315     1782  1.693  1.599 5.795e-02 1.678 1.617 2.554e-01
## 3316     2416  1.305  1.224 1.190e-03 1.311 1.218 1.847e-04
## 3317      374  2.053  1.943 4.058e-01 2.003 2.011 9.253e-01
## 3318      846  1.868  1.915 5.660e-01 1.790 2.091 3.822e-04
## 3319      561  3.036  2.884 4.516e-01 2.896 3.141 6.701e-02
## 3320     1529  1.258  1.330 1.204e-02 1.303 1.310 6.662e-01
## 3321      752  1.667  1.601 3.331e-01 1.648 1.614 9.205e-01
## 3322      565  1.577  1.709 8.117e-02 1.576 1.712 6.593e-02
## 3400     1137  4.433  4.073 1.001e-01 4.239 4.339 2.068e-01
## 3401       34  3.279  2.757 5.387e-01 3.021 3.053 8.206e-01
## 3402       50  4.199  4.245 8.274e-01 3.483 4.891 2.929e-02
## 3403       83  3.587  3.264 6.014e-01 3.000 4.076 2.510e-02
## 3404      104  3.633  3.064 3.547e-01 3.107 3.859 1.367e-01
## 3500      326  3.779  2.590 7.637e-06 3.534 2.733 6.203e-03
## 3501       32  3.944  4.495 5.666e-01 4.889 3.597 9.342e-02
## 3502       NA     NA     NA        NA    NA    NA        NA
## 3503       NA     NA     NA        NA    NA    NA        NA
## 3504      296  4.589  4.037 2.965e-01 4.503 4.086 2.820e-01
## 3505       29  5.294  2.333 6.962e-04 3.302 3.043 9.132e-01
## 3506       32  4.886  4.687 9.156e-01 6.480 4.476 4.468e-01
## 3600       89  2.549  2.322 5.038e-01 2.524 2.380 7.339e-01
## 3601      131  3.530  2.994 3.245e-01 3.218 3.763 7.819e-02
## 3602       47  3.528  2.535 1.243e-01 2.985 2.853 8.277e-01
## 3603       80  2.920  2.770 7.881e-01 2.829 2.921 9.372e-01
## 3604        5  4.583  2.759 7.609e-01 4.583 1.000 2.636e-01
## 3605      267  5.659  4.631 2.676e-02 5.216 5.043 7.682e-01
## 3606        8  2.913  2.711 1.000e+00 3.224 2.449 5.385e-01
## 3607      214  4.090  3.553 1.732e-01 4.156 3.602 1.851e-01
## 3608       NA     NA     NA        NA    NA    NA        NA
## 3609       98  2.867  3.593 1.519e-01 2.638 3.914 5.159e-03
```

```
## 3610       60  3.404  3.258 7.854e-01 3.329 3.305 8.491e-01
## 3611      251  4.036  3.501 2.791e-01 3.862 3.714 6.733e-01
## 3612     1495  3.719  4.045 7.387e-03 3.533 4.132 3.400e-06
## 3613       43  4.147  4.311 8.552e-01 4.606 4.244 7.494e-01
## 3614      510  3.443  4.009 7.596e-02 2.762 4.286 6.535e-08
## 3615       NA     NA     NA        NA    NA    NA        NA
## 3616      604  2.714  2.705 9.527e-01 2.543 3.019 6.057e-04
```

```
print(RegCoef)
```

```
##            FFA1       FLA1       2       3       4        5+       FFA2
## 1000 -4.349e-02 -3.418e-02 1.04239 1.13885 1.18687  1.300927 -0.0036252
## 1100 -1.195e-02 -3.052e-02 0.26502 0.33166 0.39299  0.491250  0.0211075
## 1101 -2.856e-02 -1.023e-01 0.21819 0.19081 0.20874  0.365867 -0.0089748
## 1102 -3.862e-04  9.924e-03 0.14261 0.14943 0.21399  0.283081  0.0091500
## 1103  3.831e-02  2.113e-02 0.08441 0.11269 0.12531  0.147807  0.0475309
## 1104  2.372e-02  6.402e-03 0.09535 0.12138 0.15419  0.244268  0.0277834
## 1105  3.335e-02  2.143e-02 0.10039 0.12801 0.15553  0.240334  0.0411374
## 1106  1.773e-02 -3.010e-02 0.36273 0.39597 0.46695  0.543874  0.0528881
## 1107 -1.925e-02 -4.074e-02 0.14474 0.17748 0.24130  0.319056 -0.0130868
## 1108  2.599e-02  2.808e-02 0.07396 0.10763 0.15445  0.206477  0.0309396
## 1109  4.032e-02 -1.053e-02 0.29160 0.36892 0.41182  0.561058  0.0917906
## 1110  5.958e-02  4.470e-02 0.20943 0.27448 0.36311  0.486947  0.0833115
## 1111 -7.284e-03  6.504e-03 0.07790 0.09301 0.12125  0.192302  0.0004209
## 1200  3.433e-02 -4.470e-03 0.25129 0.42422 0.51028  0.549492 -0.0529704
## 1201  2.275e-02  8.209e-03 0.22047 0.25454 0.28510  0.447314  0.0340327
## 1202  5.980e-02 -1.365e-02 0.37084 0.60993 0.65218  0.621464  0.0546829
## 1203  2.439e-02  3.167e-02 0.42830 0.45836 0.47059  0.515246  0.0577615
## 1204  1.072e-02  2.648e-02 0.27220 0.26990 0.31675  0.398072  0.0178227
## 1205  3.814e-02  1.759e-02 0.17490 1.07711 1.32852 -0.266183 -0.0603017
## 1206 -2.280e-01  8.817e-02 0.24100 0.61693 0.58254  0.513405 -0.1811519
## 1207  1.039e-01  4.523e-02 0.43220 0.53839 0.62345  0.618491  0.1955112
## 1208 -1.629e-01  2.175e-01 0.36506 0.86887 0.95105  0.941738 -0.2333299
## 1209 -1.256e-01 -7.170e-02 0.36645 0.49118 0.54230  0.296446 -0.1277546
## 1210 -6.731e-02  5.158e-02 0.39539 0.40056 0.73974  0.486664 -0.0706128
## 1211  6.248e-03  1.723e-03 0.21803 0.29807 0.42338  0.478905  0.0649834
## 1212  7.735e-02 -8.798e-02 0.39739 0.47793 0.60082  0.608294  0.0776861
## 1213  1.634e-03  1.403e-02 0.26340 0.60740 0.37638  0.562328 -0.0144605
## 1300 -5.022e-02 -5.264e-02 0.31136 0.36822 0.43759  0.617157 -0.0142213
## 1301  1.859e-01 -2.415e-01 0.33867 0.29974 0.16053  0.472302  0.2263407
## 1302 -6.612e-03  3.152e-02 0.15023 0.23160 0.27002  0.340175  0.0028779
## 1303 -1.192e-02 -1.551e-02 0.23725 0.26901 0.28970  0.359463 -0.0020765
## 1304 -2.510e-02 -2.884e-02 0.22522 0.25768 0.29155  0.330599 -0.0083331
## 1305 -1.843e-02 -1.408e-02 0.44287 0.48431 0.52592  0.624939  0.0071607
## 1306 -1.067e-02 -1.076e-02 0.31452 0.39454 0.42290  0.558419 -0.0007636
## 1307 -1.571e-02 -2.467e-02 0.27378 0.31506 0.34461  0.433770 -0.0076489
## 1308  1.876e-02  3.488e-03 0.22662 0.29414 0.33956  0.421863  0.0227794
## 1309 -1.318e-02 -8.993e-03 0.09687 0.12340 0.14747  0.250179 -0.0076228
## 1310  2.475e-02  9.498e-03 0.08260 0.12971 0.16682  0.256045  0.0329185
## 1311 -3.280e-02 -2.796e-02 0.15877 0.18351 0.20167  0.293142 -0.0204114
```

```
## 1312 -1.511e-02 -1.589e-02 0.18242 0.21755 0.24490  0.318513 -0.0090004
## 1313  9.197e-03 -9.704e-03 0.30906 0.36216 0.39509  0.459661  0.0143478
## 1314 -6.101e-03 -7.498e-03 0.11586 0.12769 0.16000  0.237116 -0.0017893
## 1315 -3.048e-02 -1.607e-02 0.11007 0.12546 0.16462  0.222996 -0.0205671
## 1400  3.313e-02  1.351e-02 0.27994 0.42137 0.48228  0.557557  0.0482861
## 1401  1.898e-02  6.533e-02 0.18515 0.22273 0.24870  0.258423  0.0258015
## 1402 -1.698e-02 -3.861e-02 0.22771 0.34531 0.43308 -0.007901 -0.0129580
## 1403  4.879e-02 -2.548e-02 0.31072 0.41941 0.44793  0.317353  0.0687933
## 1404  7.795e-02 -5.848e-03 0.27397 0.36619 0.43230  0.447169  0.0984007
## 1405  7.673e-02  4.583e-02 0.29565 0.41513 0.43535  0.319938  0.0955634
## 1406  4.586e-02 -7.576e-03 0.29928 0.33558 0.27372  0.192705  0.0586274
## 1407  5.928e-02 -5.434e-05 0.22913 0.32735 0.38670  0.368892  0.0717032
## 1408  1.384e-02 -1.883e-02 0.30116 0.38780 0.38969  0.276649  0.0207834
## 1409  6.543e-02  8.060e-03 0.12572 0.15304 0.24578  0.282595  0.0712283
## 1410  5.838e-02  1.923e-02 0.39261 0.55355 0.52426 -0.031708  0.0699260
## 1500 -5.363e-02 -7.513e-02 0.37534 0.40830 0.47465  0.532121 -0.0097513
## 1501  1.423e-01  1.427e-01 0.59716 0.69957 0.84948  0.893936  0.2532169
## 1502 -2.682e-02 -1.626e-02 0.41609 0.49030 0.55089  0.634740 -0.0050409
## 1503 -2.227e-03 -2.126e-02 0.12277 0.15715 0.15978  0.195540  0.0011405
## 1504  1.240e-02 -2.571e-02 0.33894 0.35687 0.40996  0.469078  0.0683254
## 1505 -1.721e-02 -2.737e-02 0.27635 0.32371 0.32452  0.355153 -0.0100656
## 1506 -2.907e-03 -7.530e-03 0.42552 0.46618 0.50777  0.529112  0.0337567
## 1507  1.564e-02 -7.925e-02 0.55526 0.60203 0.70495  0.849895  0.0640255
## 1508  5.096e-02  8.696e-03 0.34822 0.41854 0.43429  0.467693  0.0786305
## 1600 -1.729e-02 -5.427e-02 0.38661 0.39622 0.43007  0.493816  0.0030680
## 1601 -4.464e-02  3.526e-02 0.42904 0.41155 0.53625  0.514818  0.0184035
## 1602  2.309e-02 -2.102e-02 0.39188 0.43967 0.47313  0.521270  0.0515610
## 1603 -2.435e-02  1.529e-02 0.20236 0.22451 0.25368  0.272125 -0.0152870
## 1604  2.549e-02 -1.765e-03 0.16145 0.16304 0.19174  0.223765  0.0310826
## 1605  7.478e-03 -1.944e-02 0.15884 0.15868 0.17643  0.188550  0.0126955
## 1606 -1.387e-02 -1.137e-03 0.16809 0.19091 0.20986  0.230895  0.0003248
## 1607  2.667e-02  7.862e-03 0.22807 0.28573 0.27992  0.328572  0.0440253
## 1700  1.073e-02  2.065e-02 0.23293 0.35944 0.39620  0.519030  0.0251738
## 1701  6.934e-02 -1.713e-02 0.47506 0.36698 0.24171  0.323736  0.0565941
## 1702  2.746e-02  4.375e-02 0.19703 0.24462 0.23450  0.258180  0.0327878
## 1703 -3.378e-02 -2.156e-02 0.24532 0.33743 0.40672  0.504740 -0.0015753
## 1704 -5.157e-02 -5.123e-02 0.33573 0.43959 0.43113  0.499096 -0.0492162
## 1705  2.566e-03 -4.188e-04 0.28865 0.38780 0.46156  0.450282 -0.0029106
## 1706 -4.378e-02 -5.674e-02 0.27915 0.31171 0.31920  0.284536 -0.0399792
## 1707 -2.644e-02 -1.063e-01 0.29370 0.30528 0.25082  0.262176 -0.0087403
## 1708  5.307e-02 -1.226e-02 0.36508 0.45402 0.47759  0.613109  0.0648743
## 1709  8.116e-02  5.089e-02 0.17533 0.24479 0.27758  0.340736  0.0917560
## 1710 -3.812e-02 -3.487e-02 0.24586 0.31547 0.36592  0.372934 -0.0414815
## 1711 -9.427e-03 -1.175e-01 0.25901 0.29486 0.28927  0.283831  0.0102619
## 1712  2.358e-02 -2.465e-03 0.35524 0.43561 0.46676  0.516099  0.0368603
## 1800  6.564e-02  4.862e-02 0.08700 0.20362 0.14026  0.214230  0.0726642
## 1801 -5.090e-02  1.740e-01 0.05108 0.06229 0.05906 -0.032289 -0.0571264
## 1802 -5.353e-03 -4.040e-02 0.38069 0.45699 0.51059  0.332639  0.0228670
## 1803  9.821e-03 -2.003e-02 0.28203 0.35382 0.36110  0.244658  0.0259768
## 1804 -4.487e-02 -4.471e-02 0.25134 0.32736 0.46148  0.444601 -0.0033100
```

```
## 1900 -4.852e-02 -7.147e-02 0.35037 0.42431 0.50906  0.583749 -0.0417395
## 1901 -1.757e-03 -2.684e-02 0.17395 0.18928 0.22141  0.262903  0.0119378
## 1902  1.117e-02 -7.120e-03 0.13043 0.16510 0.18574  0.266658  0.0227339
## 1903 -3.753e-02  1.279e-02 0.23671 0.24153 0.37663  0.523901  0.0024110
## 1904  1.198e-02 -1.194e-02 0.11972 0.14325 0.17929  0.234557  0.0182355
## 1905  6.297e-02 -2.259e-02 0.08591 0.07525 0.12524  0.204613  0.0809234
## 1906 -6.579e-03 -1.578e-02 0.09476 0.10938 0.13445  0.173899  0.0027615
## 1907 -1.200e-02 -5.568e-03 0.26906 0.32883 0.37356  0.426309  0.0138781
## 1908 -1.153e-02 -2.799e-02 0.13744 0.14911 0.17325  0.233804  0.0007827
## 1909  2.118e-02 -2.364e-02 0.54351 0.49727 0.53584  0.489768  0.0792383
## 1910  7.317e-04 -2.174e-02 0.21437 0.24823 0.28414  0.337998  0.0208772
## 1911  9.755e-03 -4.376e-02 0.10272 0.17094 0.21155  0.272447  0.0213849
## 1912 -8.453e-03 -1.868e-02 0.15919 0.18241 0.19001  0.230213  0.0100973
## 1913  3.757e-02 -2.526e-02 0.03952 0.02067 0.01498  0.163997  0.0365283
## 2000  5.798e-02  8.118e-02 0.21055 0.33850 0.17692  0.888131  0.0555135
## 2001 -3.971e-02 -5.928e-03 0.22265 0.27342 0.34928  0.258224 -0.0363338
## 2002  5.188e-02  2.028e-02 0.21825 0.29187 0.35098  0.249433  0.0618474
## 2003  5.040e-02  3.897e-03 0.23293 0.31658 0.41339  0.389536  0.0603972
## 2100 -1.402e-02 -1.098e-02 0.22867 0.25802 0.26148  0.321900  0.0057764
## 2101 -9.199e-02 -4.613e-02 0.27995 0.45516 0.78679  0.763401  0.0086512
## 2102  1.332e-02 -1.789e-02 0.51569 0.61433 0.70261  0.713390  0.0690061
## 2103 -1.950e-03  3.678e-02 0.38040 0.43675 0.51937  0.537292  0.0241958
## 2104 -1.122e-01 -1.792e-01 0.14549 0.21084 0.30617  0.381004 -0.0846377
## 2105 -1.502e-02 -8.600e-02 0.53148 0.57736 0.67609  0.696974  0.0121314
## 2200  9.555e-02  2.447e-02 0.35562 0.50819 0.55198  0.756896  0.1474909
## 2201 -2.360e-02 -8.344e-02 0.34773 0.55250 0.60268  0.715422 -0.0082124
## 2202 -1.168e-01 -1.422e-01 0.33420 0.39804 0.37068  0.335446 -0.1136495
## 2203  2.290e-02 -6.578e-02 0.40047 0.43140 0.51915  0.375110  0.0519731
## 2204  1.584e-02 -1.692e-02 0.43646 0.50004 0.53820  0.608756  0.0396301
## 2205  2.108e-02 -3.182e-02 0.35350 0.37471 0.36379  0.360628  0.0490330
## 2206 -4.728e-02 -8.271e-02 0.33402 0.39182 0.40109  0.476956 -0.0282653
## 2207  1.499e-02 -7.165e-02 0.29894 0.37772 0.38594  0.393090  0.0395923
## 2208 -4.118e-03 -3.059e-02 0.36245 0.41078 0.40932  0.426659  0.0205500
## 2209  2.176e-02 -6.686e-02 0.49447 0.57560 0.59146  0.569367  0.0667888
## 2210 -1.732e-02 -1.968e-02 0.42386 0.50487 0.54579  0.713003  0.0418998
## 2211  1.231e-02  1.519e-02 0.29025 0.34797 0.38565  0.501895  0.0633989
## 2212 -8.131e-03 -7.133e-02 0.31109 0.44134 0.45814  0.597832  0.0217581
## 2213  4.791e-02 -2.089e-02 0.24886 0.32047 0.37203  0.389339  0.0818148
## 2214  1.387e-01 -3.729e-02 0.70420 0.76167 0.85994  0.746562  0.1180746
## 2215  9.921e-02 -7.376e-02 0.49486 0.52872 0.55445  0.495798  0.1659286
## 2216  4.979e-02  4.472e-02 0.45247 0.83526 0.68091  0.613881 -0.0170701
## 2300  2.539e-02  7.112e-03 0.30285 0.33109 0.37236  0.424260  0.0537586
## 2301 -3.414e-02  5.574e-02 0.14145 0.23029 0.18973  0.234489 -0.0280623
## 2302 -3.116e-03  7.947e-04 0.17818 0.21630 0.23708  0.306828  0.0162298
## 2303  2.418e-02  9.131e-03 0.11801 0.14662 0.19098  0.266811  0.0309634
## 2304  1.619e-02  1.349e-02 0.22461 0.24885 0.28541  0.337108  0.0303693
## 2305  2.368e-02 -1.703e-02 0.50549 0.52888 0.58585  0.619822  0.0597560
## 2306  9.746e-03  1.378e-02 0.22707 0.28407 0.35135  0.420833  0.0299370
## 2307  6.642e-02  1.442e-02 0.21016 0.25122 0.27430  0.394347  0.0930706
## 2308 -1.566e-02  2.796e-03 0.20561 0.22919 0.24470  0.291104 -0.0095690
```

```
## 2309 -1.194e-02 -5.077e-03 0.14741 0.17075 0.18060  0.210188 -0.0073790
## 2310  1.932e-02 -3.342e-02 0.45786 0.56790 0.57335  0.649825  0.0590144
## 2311  5.498e-03 -2.859e-02 0.53083 0.63933 0.66624  0.728692  0.0354053
## 2312 -5.366e-03 -3.689e-02 0.39478 0.44706 0.48576  0.526754  0.0300021
## 2400 -4.792e-02 -2.806e-02 0.15066 0.20259 0.22925  0.295656 -0.0257385
## 2401  5.139e-02 -9.142e-02 0.53833 0.58385 0.59214  0.692849  0.2183608
## 2402 -1.701e-02 -1.643e-02 0.50477 0.57459 0.61708  0.737911  0.0128847
## 2403 -7.617e-03 -3.430e-02 0.20038 0.22294 0.25746  0.345641 -0.0043402
## 2404 -5.959e-03  3.465e-03 0.20199 0.23318 0.26746  0.345319 -0.0009100
## 2405  2.515e-02 -1.790e-02 0.20360 0.26014 0.33323  0.419331  0.0386575
## 2406 -8.157e-03 -3.867e-02 0.12214 0.14123 0.17040  0.241679 -0.0072084
## 2500  5.401e-03 -7.573e-03 0.26136 0.29961 0.34074  0.508280  0.0450099
## 2501  2.823e-02  1.418e-01 0.24818 0.30549 0.42449  0.400648  0.0573260
## 2502  3.248e-02  4.992e-02 0.22296 0.26471 0.30511  0.328436  0.0429373
## 2503  5.551e-02 -2.055e-04 0.21240 0.26732 0.30821  0.351100  0.0761190
## 2504 -1.417e-02  1.335e-02 0.15137 0.20171 0.21216  0.244397  0.0038233
## 2505  1.398e-02  9.888e-03 0.18664 0.20894 0.25012  0.292579  0.0304025
## 2506  2.406e-02  5.452e-04 0.64873 0.75019 0.81052  0.849294  0.1162225
## 2507  2.131e-02 -1.414e-02 0.37705 0.44111 0.47596  0.491331  0.0584580
## 2508 -5.746e-03 -5.949e-03 0.24424 0.28502 0.32240  0.357729  0.0159180
## 2600 -5.281e-02 -5.555e-02 0.23984 0.36522 0.40870  0.454657 -0.0336715
## 2601  1.492e-01  3.015e-02 0.32574 0.43305 0.91298  0.653862  0.2183358
## 2602  3.623e-02 -6.400e-02 0.14452 0.26459 0.12617  0.317964  0.0245255
## 2603 -3.777e-02 -1.139e-02 0.22842 0.33482 0.32591  0.532870 -0.0021226
## 2604 -5.859e-02 -3.478e-02 0.24168 0.30691 0.31107  0.263653 -0.0269004
## 2605 -6.416e-02  2.699e-02 0.18926 0.24791 0.31382  0.379937 -0.0278557
## 2606  5.917e-02 -1.672e-01 0.18381 0.23818 0.29454  0.340150  0.0742238
## 2607  7.786e-03 -4.793e-02 0.13302 0.17405 0.19129  0.386986  0.0182290
## 2608  2.378e-02 -8.726e-02 0.17884 0.25712 0.23581  0.416989  0.0310235
## 2609 -2.407e-02 -1.122e-01 0.14081 0.27986 0.30138  0.181488 -0.0089850
## 2610 -1.434e-02  1.303e-02 0.24352 0.35292 0.48376  0.441993  0.0283771
## 2611 -1.300e-02  1.695e-03 0.16077 0.21900 0.29712  0.369473  0.0117235
## 2612 -5.165e-02 -1.241e-01 0.21276 0.37645 0.23292  0.402092 -0.0134713
## 2613 -4.545e-02 -2.841e-02 0.22525 0.33703 0.35318  0.438550 -0.0010488
## 2614  2.913e-02 -2.881e-02 0.16825 0.24941 0.25945  0.338405  0.0375856
## 2700 -1.436e-03 -5.059e-02 0.63301 0.83577 1.03864  1.451899  0.0997112
## 2701  4.986e-02  1.087e-02 0.42856 0.52509 0.59425  0.670070  0.1043419
## 2702  4.897e-03 -2.455e-02 0.09688 0.12430 0.10900  0.280456  0.0066194
## 2703  3.311e-02  2.820e-02 0.39468 0.53279 0.62072  0.801739  0.0814576
## 2704  7.049e-03  1.693e-03 0.25215 0.44299 0.57200  0.778959  0.0698127
## 2705  3.630e-02  3.972e-02 0.22871 0.32766 0.42871  0.678494  0.0626348
## 2706  5.525e-03  9.705e-03 0.21058 0.33234 0.42604  0.632470  0.0003806
## 2707  4.869e-02 -1.220e-02 0.36868 0.55631 0.62419  0.774167  0.0654045
## 2708 -3.371e-02 -2.503e-02 0.28894 0.29740 0.37912  0.545471  0.0079075
## 2709 -6.475e-02 -1.613e-01 0.28647 0.29325 0.37024  0.375301 -0.0467683
## 2710  1.300e-02  2.745e-02 0.17250 0.19017 0.22839  0.258578  0.0221291
## 2711 -3.023e-02 -2.706e-02 0.42588 0.64527 0.77831  0.957127 -0.0116218
## 2712  1.482e-02 -6.123e-03 0.24680 0.36731 0.45351  0.582973  0.0449128
## 2713  1.239e-02 -1.292e-02 0.20361 0.28554 0.32868  0.401237  0.0485808
## 2714  4.124e-02  1.867e-03 0.50405 0.66961 0.77063  0.960125  0.0936601
```

```
## 2715 -3.769e-02 -4.307e-02 0.24313 0.33410 0.48161  0.683988 -0.0237283
## 2716 -5.497e-02 -5.770e-02 0.23051 0.23847 0.23911  0.346911 -0.0366057
## 2717  1.532e-02 -9.976e-03 0.33325 0.44279 0.54085  0.662878  0.0227301
## 2718  1.952e-02 -5.621e-02 0.31325 0.28502 0.40848  0.498369  0.0117119
## 2719 -5.690e-03 -2.911e-02 0.37659 0.51123 0.57424  0.665615  0.0570370
## 2720  1.080e-02 -2.530e-02 0.30152 0.33045 0.37548  0.542277  0.0466064
## 2721  2.068e-04  5.708e-03 0.14759 0.18255 0.36387  0.515638  0.0266409
## 2722  1.193e-02 -1.620e-02 0.09755 0.14790 0.18008  0.286977  0.0207840
## 2723 -8.337e-03 -1.693e-02 0.19186 0.24044 0.32069  0.449038  0.0122165
## 2724 -2.058e-03 -1.112e-02 0.38515 0.54273 0.64452  0.759032  0.0470198
## 2725 -5.967e-03 -1.114e-02 0.38271 0.43834 0.50300  0.609425  0.0155723
## 2726  2.394e-02  1.470e-02 0.40024 0.48475 0.60426  0.740697  0.0785577
## 2727 -3.552e-02 -6.845e-02 0.18335 0.23688 0.35288  0.513801 -0.0149389
## 2728  1.665e-02  1.476e-02 0.27188 0.35217 0.42183  0.554469  0.0301614
## 2729  2.925e-02  2.100e-02 0.32280 0.43095 0.48851  0.617408  0.0568534
## 2730 -1.111e-03 -1.102e-02 0.38602 0.49235 0.55726  0.704575  0.0163222
## 2731 -3.063e-02 -3.107e-03 0.34657 0.42296 0.53495  0.723225  0.0025025
## 2732 -1.067e-02 -6.107e-03 0.34128 0.46786 0.57544  0.700237 -0.0081121
## 2733  5.274e-02  7.122e-02 0.16767 0.23582 0.35379  0.541483  0.0676159
## 2734 -6.574e-03 -1.706e-02 0.11683 0.19323 0.20077  0.333128  0.0124074
## 2735  5.256e-02  2.181e-02 0.29466 0.40505 0.47814  0.620721  0.0749380
## 2736  6.992e-03 -3.262e-02 0.33457 0.42689 0.48921  0.577997  0.0215799
## 2737  4.780e-03  6.987e-03 0.08772 0.11034 0.13928  0.299650  0.0074371
## 2738  1.876e-02 -2.285e-03 0.35775 0.43569 0.49490  0.588000  0.0365762
## 2739  2.802e-02 -1.841e-02 0.32877 0.44444 0.50377  0.577929  0.0605778
## 2740  1.023e-02  5.287e-03 0.20920 0.34130 0.41401  0.653024  0.0287614
## 2741 -6.778e-03 -3.463e-02 0.25880 0.36299 0.49272  0.685111  0.0156721
## 2742  1.365e-02 -6.261e-02 0.33316 0.46404 0.55556  0.673772  0.0086908
## 2743 -3.261e-02 -1.463e-02 0.21090 0.27496 0.28128  0.416609 -0.0075774
## 2744         NA         NA      NA      NA      NA         NA         NA
## 2745  3.028e-02  2.495e-02 0.27351 0.33850 0.47388  0.595900  0.0753704
## 2746  3.941e-03  1.136e-02 0.27012 0.39154 0.49076  0.681506  0.0290228
## 2747 -1.583e-02 -4.020e-02 0.11702 0.18240 0.30180  0.412508 -0.0178757
## 2748 -4.467e-02 -1.942e-02 0.59209 0.71036 0.81567  0.983638 -0.0057943
## 2800  2.003e-03 -2.493e-02 0.28922 0.32664 0.36295  0.467794  0.0174943
## 2801  2.583e-02  3.474e-02 0.18026 0.20920 0.23310  0.408243  0.0498119
## 2802  2.545e-02  2.248e-02 0.08089 0.08633 0.11824  0.169655  0.0303540
## 2803  1.867e-02  8.441e-03 0.06120 0.10670 0.14254  0.222816  0.0173034
## 2804 -1.403e-02 -4.521e-03 0.12431 0.17584 0.20991  0.286781 -0.0083846
## 2805  8.963e-03  7.131e-03 0.18325 0.20844 0.22207  0.258437  0.0202058
## 2806  1.072e-02  2.198e-02 0.21290 0.25267 0.28106  0.384635  0.0164711
## 2807  2.081e-02 -4.695e-02 0.18279 0.17686 0.23697  0.228925  0.0208576
## 2808  2.600e-02  1.759e-02 0.17564 0.22338 0.26845  0.365505  0.0371660
## 2809 -1.824e-02 -3.139e-02 0.33723 0.37714 0.42835  0.527687  0.0130190
## 2900  3.558e-02 -3.019e-02 0.28069 0.43813 0.53388  0.687146  0.0516289
## 2901  3.685e-02  9.017e-03 0.22673 0.39931 0.42636  0.443268  0.0413790
## 2902 -1.868e-02 -2.213e-02 0.32534 0.70188 0.92658  1.153593 -0.0716915
## 2903  4.619e-02 -5.773e-02 0.28834 0.56666 0.98074  1.501561  0.2190358
## 2904 -1.045e-02 -3.672e-02 0.18382 0.40956 0.61913  1.082356  0.0300483
## 2905  7.839e-02 -3.516e-02 0.47798 0.59873 0.60819  0.727693  0.0978351
```

```
## 2906  2.543e-01  9.161e-03 0.46977 0.80317 0.98901  1.215145  0.5305004
## 2907 -2.719e-02  9.521e-03 0.44204 0.84520 1.05809  1.280284 -0.0888332
## 2908  9.665e-02  5.317e-02 0.03717 0.30274 0.59576  1.163806  0.1220745
## 2909  4.795e-02 -3.029e-02 0.31098 0.42499 0.50177  0.584875  0.0574641
## 2910  3.969e-02 -9.695e-03 0.14910 0.27145 0.41014  0.566101  0.0928996
## 2911  6.196e-02 -2.455e-02 0.41861 0.59623 0.70623  1.004054  0.0714989
## 2912 -9.041e-03 -1.206e-01 0.59156 0.77877 0.84055  0.914930  0.0229279
## 2913  5.236e-01  3.463e-01 0.14136 0.41485 0.49879  0.517114  0.6326977
## 2914  1.790e-01 -2.430e-02 0.22570 0.47502 0.59087  0.650717  0.2182592
## 2915         NA         NA      NA      NA      NA        NA  0.0000000
## 2916 -1.480e-02 -5.344e-02 0.37140 0.46592 0.54555  0.631344  0.0221104
## 2917  1.046e-01 -2.285e-02 0.18581 0.31908 0.43605  0.489821  0.1336144
## 2918         NA         NA      NA      NA      NA        NA         NA
## 2919  3.781e-01  1.388e-01 0.14640 0.36437 0.41178  0.553729  0.4715171
## 2920  2.628e-01  6.607e-02 0.59853 0.60513 0.72415  0.892226  0.2819279
## 2921  5.853e-02  3.053e-02 0.24023 0.40467 0.46291  0.551962  0.0536311
## 2922  1.905e-01  3.259e-02 0.03027 0.20690 0.25209  0.286590  0.2436314
## 2923  8.460e-02 -1.668e-01 0.21001 0.35058 0.32080  0.292816  0.1326296
## 3000  1.909e-02  6.547e-02 0.28209 0.28809 0.35612  0.417150  0.0321089
## 3001         NA         NA      NA      NA      NA        NA  0.0000000
## 3002  9.041e-03 -1.574e-02 0.13915 0.15810 0.18269  0.264051  0.0119015
## 3003  8.582e-03 -6.272e-02 0.60346 0.69277 0.72782  0.736010  0.0119835
## 3004  2.557e-02 -2.754e-04 0.26163 0.31815 0.35191  0.429754  0.0344758
## 3005  3.058e-02  3.534e-02 0.13446 0.16936 0.19041  0.235724  0.0378496
## 3100 -3.912e-02  1.123e-02 0.35396 0.44613 0.52092  0.612759  0.0249516
## 3101  2.258e-02 -5.338e-02 0.32282 0.41331 0.45710  0.476003  0.0585190
## 3102  5.062e-02  9.533e-02 0.15628 0.18046 0.24929  0.272619  0.0679818
## 3103 -9.369e-03 -1.369e-02 0.21550 0.25293 0.25049  0.294735  0.0192650
## 3104 -1.552e-02  7.541e-03 0.15493 0.16221 0.18365  0.257148  0.0077066
## 3105 -3.792e-02 -1.114e-02 0.27364 0.35484 0.37560  0.405731 -0.0043751
## 3106 -1.683e-02 -3.229e-02 0.36643 0.45395 0.43846  0.363847  0.0215047
## 3107 -2.809e-02 -1.929e-02 0.27586 0.37954 0.43072  0.540789  0.0222896
## 3108 -1.862e-02  6.364e-03 0.13330 0.28913 0.31108  0.443642  0.0251825
## 3109 -3.542e-02  4.352e-02 0.19540 0.31552 0.37040  0.417955  0.0056827
## 3110  3.085e-02  3.132e-02 0.21186 0.26667 0.30223  0.337636  0.0540474
## 3200  1.790e-02 -1.092e-02 0.26751 0.32989 0.36628  0.412340  0.0467668
## 3201 -6.428e-03 -3.178e-02 0.14100 0.16897 0.22388  0.204989  0.0128966
## 3202  3.652e-03 -1.649e-03 0.25898 0.33131 0.35575  0.393110  0.0243325
## 3203  2.196e-05 -2.164e-02 0.33621 0.39269 0.43933  0.507110  0.0250318
## 3204  1.556e-02 -7.212e-03 0.17714 0.21428 0.26405  0.331674  0.0280432
## 3205  2.311e-02  1.386e-02 0.15695 0.16726 0.16872  0.207591  0.0369155
## 3206  3.139e-02  2.777e-03 0.30240 0.36308 0.38014  0.421417  0.0626670
## 3207  1.670e-02  1.213e-03 0.21416 0.27348 0.31719  0.321295  0.0359183
## 3300  5.207e-02  2.108e-02 0.26770 0.38101 0.39146  0.407849  0.0568205
## 3301  1.751e-02 -1.993e-02 0.26744 0.31632 0.34724  0.349812  0.0472351
## 3302 -3.091e-02 -3.090e-03 0.22776 0.21089 0.28494  0.361288 -0.0279653
## 3303 -2.662e-02  5.645e-02 0.19912 0.20904 0.24677  0.208097 -0.0234674
## 3304  6.550e-02  2.543e-02 0.23057 0.29529 0.37261  0.437664  0.0858946
## 3305  6.922e-02  3.738e-02 0.14099 0.20918 0.21202  0.289341  0.0773095
## 3306  1.232e-02 -2.676e-02 0.25062 0.34186 0.37303  0.444853  0.0384931
```

```
## 3307  6.425e-03 -2.171e-02 0.25617 0.27123 0.28502  0.365439  0.0284154
## 3308  4.542e-02 -3.521e-02 0.27507 0.38824 0.42870  0.448412  0.0674710
## 3309 -6.581e-02 -6.699e-02 0.36027 0.48569 0.53924  0.591514 -0.0406355
## 3310  6.554e-03  2.405e-02 0.41826 0.45123 0.47524  0.504818  0.0445592
## 3311  2.068e-02  1.682e-02 0.24592 0.24935 0.31187  0.239439  0.0232774
## 3312  2.092e-02  1.527e-02 0.28751 0.33632 0.35875  0.370464  0.0432755
## 3313  1.091e-01  2.399e-02 0.23177 0.24590 0.18813  0.140095  0.1206449
## 3314  8.806e-03  3.593e-03 0.27319 0.32072 0.32872  0.310319  0.0263365
## 3315  1.834e-02  3.731e-02 0.30617 0.37404 0.42771  0.489450  0.0416785
## 3316  7.144e-02 -4.549e-03 0.26636 0.37934 0.38912  0.465864  0.0892988
## 3317  6.609e-02  1.102e-02 0.17725 0.25428 0.25611  0.348181  0.0874849
## 3318  1.931e-02  5.279e-02 0.25429 0.29901 0.36800  0.342019  0.0201218
## 3319  1.446e-02  8.731e-03 0.31101 0.41804 0.48423  0.511641  0.0460852
## 3320  2.746e-02  2.703e-02 0.38704 0.48925 0.67452  0.600807  0.0169694
## 3321 -3.762e-02  2.554e-02 0.21277 0.23284 0.27169  0.254213 -0.0284355
## 3322  3.807e-03  3.650e-04 0.16904 0.20969 0.15606  0.235445 -0.0039536
## 3400 -1.152e-02 -2.237e-02 0.39831 0.53639 0.59054  0.667531 -0.0023851
## 3401  1.607e-01  1.016e-01 0.15183 0.25175 0.01782  0.204363  0.1389573
## 3402  1.192e-01  1.119e-02 0.08838 0.12228 0.14153  0.270571  0.1486833
## 3403  4.909e-02  1.024e-01 0.02765 0.01558 0.03175  0.026582  0.0508202
## 3404  2.864e-03 -3.055e-02 0.43603 0.59326 0.72437  0.779876  0.0603892
## 3500  4.113e-02  5.297e-03 0.37058 0.51645 0.56500  0.645140  0.1264808
## 3501  3.276e-01 -1.825e-01 0.65172 0.44478 0.55515  0.803560  0.3257343
## 3502         NA         NA      NA      NA      NA        NA         NA
## 3503         NA         NA      NA      NA      NA        NA         NA
## 3504  6.191e-02  5.207e-02 0.09615 0.22761 0.28813  0.419083  0.1117701
## 3505  4.132e-02 -9.219e-02 0.47933 0.49931 0.54588  0.719717  0.0958463
## 3506  3.072e-03 -5.486e-03 0.11468 0.12327 0.15858  0.190901  0.0082080
## 3600  3.041e-02  1.891e-03 0.32386 0.42750 0.51707  0.636188  0.0615886
## 3601  5.392e-02 -4.447e-02 0.24970 0.37008 0.37227  0.445633  0.0848291
## 3602  8.611e-02 -7.068e-02 0.19074 0.26664 0.36674  0.562287  0.1325483
## 3603  1.533e-01  1.288e-02 0.15572 0.24727 0.25519  0.595445  0.1468734
## 3604 -4.920e-01  4.920e-01 0.00000 0.00000 0.49200  0.000000  0.0000000
## 3605  7.588e-02  8.755e-02 0.25981 0.28761 0.34004  0.521237  0.1288518
## 3606         NA         NA      NA      NA      NA        NA         NA
## 3607 -5.095e-02 -3.054e-02 0.13098 0.09685 0.16424  0.304770 -0.0443039
## 3608         NA         NA      NA      NA      NA        NA         NA
## 3609  5.039e-03 -9.492e-02 0.09352 0.14579 0.24183  0.202052  0.0058783
## 3610 -5.646e-02  7.152e-02 0.11025 0.24240 0.38892  0.505218 -0.0161225
## 3611  9.712e-02 -1.263e-02 0.49012 0.64638 0.77557  1.041655  0.1935466
## 3612  3.783e-04 -2.877e-02 0.38249 0.51890 0.62609  0.745167  0.0024763
## 3613  4.638e-02 -9.359e-02 0.17137 0.19761 0.30332  0.393224  0.0847595
## 3614 -8.549e-02 -1.313e-01 0.28886 0.34821 0.43378  0.545251 -0.0725975
## 3615         NA         NA      NA      NA      NA        NA         NA
## 3616  4.493e-02  4.880e-02 0.30546 0.36592 0.40720  0.499397  0.0618739
##            FLA2       FFA3       FLA4
## 1000 -4.750e-02 -0.0064050 -0.0478566
## 1100 -3.239e-02  0.0164656 -0.0285193
## 1101 -1.174e-01 -0.0460669 -0.1207178
## 1102  8.441e-03  0.0105155  0.0102576
```

```
## 1103  1.882e-02  0.0518544  0.0317626
## 1104  5.879e-03  0.0286276  0.0107783
## 1105  1.818e-02  0.0446058  0.0276641
## 1106 -4.656e-02  0.0412083 -0.0310622
## 1107 -4.383e-02 -0.0199398 -0.0463696
## 1108  2.431e-02  0.0346870  0.0301323
## 1109 -3.596e-03  0.0913632  0.0108341
## 1110  3.450e-02  0.0897970  0.0532888
## 1111  1.186e-02  0.0018975  0.0119277
## 1200  8.410e-02  0.0234880  0.0366366
## 1201 -1.162e-02  0.0286467  0.0046260
## 1202 -1.183e-02  0.0438786  0.0380355
## 1203  3.554e-02  0.0812011  0.0736845
## 1204  1.642e-03  0.0188946  0.0135120
## 1205  1.143e-01  0.0507061  0.0559615
## 1206  2.037e-02 -0.1630329 -0.1396410
## 1207  5.483e-02  0.2353017  0.1996082
## 1208  2.909e-01  0.0491473  0.0645550
## 1209 -1.213e-01 -0.2218837 -0.2200332
## 1210  7.208e-02 -0.0077560  0.0107212
## 1211 -4.857e-03  0.0608991  0.0509564
## 1212 -4.289e-02  0.0402418  0.0243379
## 1213  2.936e-02  0.0123432  0.0161682
## 1300 -6.352e-02 -0.0227596 -0.0660100
## 1301 -2.889e-01  0.1837852 -0.2422344
## 1302  1.914e-02  0.0062706  0.0197154
## 1303 -1.526e-02 -0.0032368 -0.0154963
## 1304 -2.831e-02 -0.0107561 -0.0294067
## 1305 -2.383e-02  0.0043337 -0.0227422
## 1306 -2.015e-02 -0.0031031 -0.0202667
## 1307 -2.904e-02 -0.0098990 -0.0298743
## 1308 -1.773e-03  0.0226337  0.0006729
## 1309 -1.458e-02 -0.0092956 -0.0156631
## 1310  2.378e-03  0.0332358  0.0077530
## 1311 -3.161e-02 -0.0242326 -0.0348728
## 1312 -1.846e-02 -0.0105966 -0.0195480
## 1313 -1.200e-02  0.0137538 -0.0110035
## 1314 -1.420e-02 -0.0031881 -0.0144486
## 1315 -1.371e-02 -0.0219372 -0.0166296
## 1400  2.475e-02  0.0618908  0.0515107
## 1401  6.808e-02  0.0644642  0.0829102
## 1402 -3.325e-02 -0.0277582 -0.0390468
## 1403 -1.680e-02  0.0603315  0.0182491
## 1404 -1.296e-02  0.0936886  0.0255377
## 1405  5.438e-02  0.1229908  0.1033615
## 1406 -3.735e-03  0.0568817  0.0240182
## 1407  8.546e-03  0.0761235  0.0459229
## 1408 -2.121e-02  0.0101970 -0.0106923
## 1409  9.288e-03  0.0758281  0.0450702
## 1410  1.069e-02  0.0759993  0.0511685
```

```
## 1500 -6.835e-02 -0.0190024 -0.0703407
## 1501  9.404e-02  0.2702466  0.1662246
## 1502 -3.467e-02 -0.0090159 -0.0355299
## 1503 -2.118e-02  0.0003250 -0.0210979
## 1504 -9.243e-04  0.0682130  0.0129453
## 1505 -2.495e-02 -0.0111169 -0.0257020
## 1506  3.033e-02  0.0361532  0.0345352
## 1507 -4.926e-02  0.0489049 -0.0284314
## 1508  2.758e-03  0.0788934  0.0129880
## 1600 -5.866e-02 -0.0022856 -0.0582148
## 1601  5.298e-02  0.0250064  0.0558450
## 1602 -1.958e-02  0.0497035 -0.0123333
## 1603  1.693e-02 -0.0139421  0.0151566
## 1604 -1.091e-03  0.0310139  0.0020890
## 1605 -1.553e-02  0.0121112 -0.0148624
## 1606  4.239e-03  0.0006237  0.0042748
## 1607  1.321e-02  0.0450010  0.0179191
## 1700  1.666e-02  0.0339188  0.0300032
## 1701  1.637e-02  0.0642350  0.0425323
## 1702  3.959e-02  0.0498648  0.0542738
## 1703 -9.657e-03 -0.0042704 -0.0101548
## 1704 -4.176e-02 -0.0619521 -0.0564201
## 1705 -1.673e-02 -0.0099217 -0.0179812
## 1706 -7.143e-02 -0.0646808 -0.0867533
## 1707 -1.006e-01 -0.0304133 -0.1027136
## 1708 -1.256e-02  0.0615476  0.0056597
## 1709  3.886e-02  0.1078323  0.0791171
## 1710 -5.923e-02 -0.0695318 -0.0799944
## 1711 -1.152e-01 -0.0122953 -0.1127716
## 1712 -5.253e-04  0.0366916  0.0117320
## 1800  5.571e-02  0.0924531  0.0817423
## 1801  1.810e-01 -0.0308117  0.1689565
## 1802 -6.190e-02 -0.0019662 -0.0515708
## 1803 -1.203e-02  0.0214736 -0.0020401
## 1804 -2.892e-02 -0.0153726 -0.0304140
## 1900 -9.551e-02 -0.0731308 -0.1109827
## 1901 -2.556e-02  0.0080619 -0.0231019
## 1902 -9.148e-04  0.0226024  0.0033540
## 1903  5.770e-02  0.0110363  0.0580265
## 1904 -1.765e-02  0.0138149 -0.0122085
## 1905 -2.036e-02  0.0744816  0.0119359
## 1906 -1.005e-02  0.0014289 -0.0095355
## 1907  9.988e-03  0.0157674  0.0130433
## 1908 -2.313e-02 -0.0025896 -0.0229726
## 1909  1.804e-02  0.0830390  0.0370756
## 1910 -1.660e-02  0.0183633 -0.0124856
## 1911 -4.424e-02  0.0131870 -0.0391136
## 1912 -1.137e-02  0.0083272 -0.0093181
## 1913 -2.559e-02  0.0286740 -0.0139153
## 2000  1.071e-01  0.1226457  0.1403224
```

```
## 2001 -1.363e-02 -0.0450778 -0.0370314
## 2002  2.681e-02  0.0758260  0.0593799
## 2003  1.507e-02  0.0673585  0.0429650
## 2100 -1.100e-02         NA -0.0099892
## 2101 -6.473e-02 -0.0199329 -0.0598975
## 2102 -1.483e-02  0.0664312  0.0013948
## 2103  3.415e-02  0.0291607  0.0386609
## 2104 -1.767e-01 -0.1211449 -0.1985903
## 2105 -1.195e-01 -0.0230910 -0.1149668
## 2200  2.054e-02  0.1549684  0.0820184
## 2201 -1.522e-01 -0.0924298 -0.1570214
## 2202 -1.572e-01 -0.1634494 -0.1970762
## 2203 -6.474e-02  0.0350469 -0.0487459
## 2204 -2.933e-02  0.0357981 -0.0218524
## 2205 -2.044e-02  0.0443949 -0.0075591
## 2206 -8.498e-02 -0.0507765 -0.0933146
## 2207 -6.361e-02  0.0241088 -0.0522056
## 2208 -2.626e-02  0.0151836 -0.0212793
## 2209 -7.903e-02  0.0421598 -0.0536058
## 2210  4.804e-03  0.0427613  0.0139962
## 2211  3.298e-02  0.0684753  0.0462054
## 2212 -4.632e-02  0.0089289 -0.0398573
## 2213 -6.697e-03  0.0794557  0.0248929
## 2214 -9.134e-02  0.0669272 -0.0238792
## 2215 -1.415e-02  0.1633694  0.0188026
## 2216  4.800e-02  0.0182130  0.0355660
## 2300  5.981e-04  0.0538893  0.0148179
## 2301  4.392e-02  0.0005529  0.0252014
## 2302  8.798e-03  0.0176187  0.0118955
## 2303  6.190e-03  0.0319723  0.0124800
## 2304  1.407e-02  0.0318899  0.0184562
## 2305 -1.482e-02  0.0574149 -0.0022272
## 2306  1.035e-02  0.0321880  0.0180828
## 2307  1.379e-02  0.0947069  0.0278328
## 2308 -1.220e-02 -0.0138305 -0.0158816
## 2309 -1.018e-02 -0.0091567 -0.0117460
## 2310 -3.055e-02  0.0542523 -0.0192111
## 2311 -4.747e-02  0.0259123 -0.0387413
## 2312 -2.913e-02  0.0259656 -0.0237967
## 2400 -2.836e-02 -0.0292836 -0.0324592
## 2401 -6.817e-02  0.2051165 -0.0147765
## 2402 -1.866e-02  0.0111145 -0.0168994
## 2403 -3.989e-02 -0.0073881 -0.0403416
## 2404 -1.822e-03 -0.0010402 -0.0019113
## 2405 -2.100e-02  0.0373959 -0.0176096
## 2406 -4.386e-02 -0.0095590 -0.0444095
## 2500  2.085e-03  0.0453014  0.0106377
## 2501  1.291e-01  0.1272515  0.1614847
## 2502  4.992e-02  0.0476453  0.0554448
## 2503  1.331e-02  0.0771122  0.0214740
```

```
## 2504  2.198e-02  0.0063572  0.0225833
## 2505  1.725e-02  0.0316719  0.0207031
## 2506  2.692e-02  0.1195933  0.0483949
## 2507 -4.828e-03  0.0578655  0.0059631
## 2508  2.378e-03  0.0161183  0.0043918
## 2600 -3.874e-02 -0.0571779 -0.0591721
## 2601  3.469e-02  0.2386833  0.1729304
## 2602 -4.301e-02 -0.0043295 -0.0275258
## 2603 -1.042e-02 -0.0081809 -0.0117277
## 2604 -3.058e-02 -0.0379743 -0.0415907
## 2605  3.900e-02 -0.0165662  0.0296925
## 2606 -1.685e-01 -0.0026501 -0.1351488
## 2607 -4.142e-02 -0.0021660 -0.0321757
## 2608 -8.695e-02 -0.0251805 -0.0662367
## 2609 -9.615e-02 -0.0711376 -0.1020981
## 2610  1.971e-02  0.0350924  0.0306193
## 2611  1.267e-02  0.0151443  0.0163544
## 2612 -1.406e-01 -0.0683666 -0.1473367
## 2613 -1.268e-02 -0.0049623 -0.0130559
## 2614 -9.456e-03  0.0337789  0.0054047
## 2700 -2.047e-01  0.0065455 -0.1552636
## 2701  1.430e-02  0.1073674  0.0376723
## 2702 -4.083e-02 -0.0007997 -0.0393804
## 2703  3.256e-02  0.0876550  0.0505646
## 2704  1.416e-02  0.0714473  0.0244802
## 2705  1.209e-02  0.0644156  0.0250882
## 2706 -5.528e-02 -0.0096005 -0.0551891
## 2707 -9.586e-02  0.0230938 -0.0658110
## 2708 -2.700e-02  0.0034592 -0.0254976
## 2709 -1.184e-01 -0.0953480 -0.1384908
## 2710  2.186e-02  0.0247248  0.0249518
## 2711 -7.217e-02 -0.0254474 -0.0747134
## 2712 -2.444e-02  0.0410875 -0.0158692
## 2713 -1.244e-02  0.0463067 -0.0029343
## 2714  1.532e-03  0.0941072  0.0287892
## 2715 -1.140e-01 -0.0458757 -0.1202914
## 2716 -6.087e-02 -0.0475008 -0.0684914
## 2717 -6.921e-02  0.0065674 -0.0638435
## 2718 -1.077e-01 -0.0318924 -0.1026777
## 2719 -3.976e-02         NA -0.0170243
## 2720 -3.204e-02  0.0430694 -0.0251652
## 2721  1.946e-03  0.0268039  0.0051702
## 2722 -2.271e-02  0.0183105 -0.0200238
## 2723 -2.172e-02  0.0099671 -0.0200753
## 2724 -3.455e-02  0.0421714 -0.0261517
## 2725 -1.467e-02  0.0142834 -0.0129518
## 2726  2.344e-02  0.0810348  0.0332489
## 2727 -9.123e-02 -0.0319697 -0.0948326
## 2728 -7.181e-03  0.0289677 -0.0009619
## 2729 -2.353e-03  0.0563090  0.0115460
```

```
## 2730 -3.542e-02  0.0109259 -0.0323124
## 2731 -6.815e-03  0.0016997 -0.0064465
## 2732 -3.065e-02 -0.0142164 -0.0324780
## 2733  6.799e-02  0.0830864  0.0859838
## 2734 -2.911e-02  0.0078937 -0.0268608
## 2735 -5.058e-03  0.0737779  0.0132962
## 2736 -5.553e-02  0.0138602 -0.0519308
## 2737 -7.373e-03  0.0066898 -0.0062805
## 2738 -1.928e-02  0.0331142 -0.0123766
## 2739 -3.788e-02  0.0498809 -0.0204003
## 2740 -1.782e-02  0.0258604 -0.0114233
## 2741 -6.612e-02  0.0058329 -0.0630319
## 2742 -1.181e-01 -0.0263670 -0.1154020
## 2743 -2.846e-02 -0.0131996 -0.0301664
## 2744         NA         NA         NA
## 2745  3.748e-02  0.0803374  0.0492461
## 2746 -7.350e-03  0.0279992 -0.0019897
## 2747 -6.476e-02 -0.0252798 -0.0676820
## 2748 -4.185e-02 -0.0119487 -0.0431967
## 2800 -3.215e-02  0.0136346 -0.0293587
## 2801  2.593e-02  0.0527803  0.0327296
## 2802  1.983e-02  0.0335445  0.0252997
## 2803 -6.009e-03  0.0167728 -0.0042219
## 2804 -1.035e-02 -0.0093342 -0.0113424
## 2805  2.757e-03  0.0208311  0.0079493
## 2806  1.612e-02  0.0189614  0.0189777
## 2807 -6.295e-02  0.0076546 -0.0584618
## 2808  2.361e-03  0.0375005  0.0089342
## 2809 -3.089e-02  0.0096934 -0.0291030
## 2900 -1.521e-01 -0.0346017 -0.1263617
## 2901 -5.818e-02  0.0176874 -0.0449951
## 2902 -4.194e-01 -0.3298286 -0.4594114
## 2903 -3.288e-01 -0.0573718 -0.1198341
## 2904 -1.754e-01 -0.0442278 -0.1621439
## 2905 -1.401e-01  0.0263330 -0.0944245
## 2906 -2.620e-01  0.3241613  0.0833861
## 2907 -2.652e-01 -0.2843807 -0.3286583
## 2908 -1.241e-01  0.0437144 -0.0607964
## 2909 -8.400e-02  0.0308038 -0.0672428
## 2910 -2.077e-02  0.0770347  0.0502147
## 2911 -1.517e-01 -0.0270831 -0.1071035
## 2912 -2.612e-01 -0.1483641 -0.2478435
## 2913  2.809e-01  0.8331011  0.6362321
## 2914 -1.361e-01  0.1220496  0.0063669
## 2915  0.000e+00  0.0000000  0.0000000
## 2916 -8.257e-02 -0.0006192 -0.0765316
## 2917 -1.223e-01  0.0776862 -0.0838945
## 2918         NA         NA         NA
## 2919  5.248e-02  0.5040682  0.2680140
## 2920  7.447e-02  0.3212280  0.2240292
```

```
## 2921 -4.088e-02  0.0349278 -0.0182027
## 2922 -3.781e-02  0.2247064  0.0363383
## 2923 -2.229e-01 -0.0064737 -0.1648722
## 3000  4.618e-02  0.0434491  0.0548402
## 3001  0.000e+00  0.0000000  0.0000000
## 3002 -1.226e-02  0.0115207 -0.0117535
## 3003 -8.882e-02  0.0050744 -0.0875180
## 3004 -1.504e-02  0.0330374 -0.0107542
## 3005  3.228e-02  0.0416042  0.0375216
## 3100  3.154e-02  0.0307767  0.0372975
## 3101 -4.297e-02  0.0485198 -0.0271768
## 3102  1.029e-01  0.0892861  0.1200295
## 3103 -1.107e-03  0.0190719  0.0031489
## 3104  1.777e-02  0.0100357  0.0191235
## 3105 -8.315e-03 -0.0056553 -0.0092416
## 3106 -3.779e-02  0.0105836 -0.0301949
## 3107  6.558e-03  0.0232406  0.0109221
## 3108  1.825e-02  0.0267845  0.0213016
## 3109  6.032e-02  0.0205563  0.0619591
## 3110  3.883e-02  0.0582721  0.0472021
## 3200  2.437e-03  0.0475877  0.0184696
## 3201 -1.489e-02  0.0065062 -0.0093684
## 3202  5.107e-03  0.0260437  0.0133756
## 3203 -2.366e-02  0.0192511 -0.0174189
## 3204 -1.001e-02  0.0250452 -0.0018359
## 3205  1.970e-02  0.0422125  0.0300576
## 3206  7.293e-03  0.0641017  0.0207191
## 3207  6.594e-03  0.0384319  0.0203151
## 3300  5.211e-02  0.0930751  0.0911647
## 3301 -1.354e-02  0.0394068  0.0137976
## 3302 -1.752e-02 -0.0395639 -0.0362377
## 3303  5.596e-02  0.0142252  0.0401037
## 3304  2.877e-02  0.1016951  0.0759301
## 3305  3.113e-02  0.0990769  0.0859136
## 3306 -4.391e-02  0.0216093 -0.0292971
## 3307 -1.423e-02  0.0243388 -0.0055579
## 3308 -2.348e-02  0.0519365  0.0212872
## 3309 -9.449e-02 -0.1102875 -0.1243569
## 3310  2.900e-02  0.0635001  0.0581047
## 3311  2.149e-02  0.0365515  0.0356427
## 3312  2.941e-02  0.0625969  0.0579380
## 3313  2.873e-02  0.1304677  0.0713317
## 3314 -8.718e-03  0.0201771  0.0099273
## 3315  3.443e-02  0.0640531  0.0616186
## 3316 -1.080e-03  0.0883746  0.0754248
## 3317  1.279e-02  0.0946626  0.0616612
## 3318  1.747e-02  0.0309278  0.0295012
## 3319  6.315e-03  0.0484568  0.0232274
## 3320  4.253e-02  0.0513626  0.0561416
## 3321  3.298e-02 -0.0067339  0.0142197
```

```
## 3322  6.774e-05 -0.0039081 -0.0026122
## 3400 -5.632e-02 -0.0097029 -0.0566679
## 3401  1.010e-01  0.1855275  0.1718681
## 3402 -1.742e-02  0.1414944  0.0500321
## 3403  1.007e-01  0.0730803  0.1141308
## 3404 -1.237e-01  0.0221669 -0.1042249
## 3500  1.528e-02  0.1308826  0.0583935
## 3501 -1.944e-01  0.3513204 -0.2750498
## 3502         NA         NA         NA
## 3503         NA         NA         NA
## 3504  7.712e-02  0.1264745  0.1042136
## 3505 -7.594e-02  0.0838940 -0.0566598
## 3506 -1.765e-03  0.0081157 -0.0012642
## 3600 -1.986e-02  0.0545236  0.0022613
## 3601 -9.210e-02  0.0439974         NA
## 3602 -2.764e-02  0.1234377  0.0151477
## 3603  4.973e-03  0.1496737  0.0869258
## 3604  0.000e+00  2.0850000 -1.6510000
## 3605  7.786e-02  0.1493600  0.1139180
## 3606         NA         NA         NA
## 3607 -2.679e-02 -0.0473132 -0.0327818
## 3608         NA         NA         NA
## 3609 -1.108e-01 -0.0397533 -0.1086810
## 3610  1.148e-01  0.0117998  0.1106110
## 3611  1.328e-02  0.1976966  0.0753841
## 3612 -6.778e-02 -0.0168917 -0.0670178
## 3613 -4.982e-02  0.0768190 -0.0358527
## 3614 -1.770e-01 -0.1053897 -0.1944226
## 3615         NA         NA         NA
## 3616  3.123e-02  0.0723050  0.0510749
```

```r
print(RegP)
```

```
##          FFA1p     FLA1p         2p         3p         4p        5+p
## 1000 1.398e-09 1.671e-04  0.000e+00  0.000e+00  0.000e+00  0.000e+00
## 1100 8.951e-02 1.720e-04  5.338e-92 3.383e-135 1.052e-175 1.495e-314
## 1101 9.071e-02 2.912e-08  1.641e-27  3.779e-19  3.258e-15  2.085e-43
## 1102 9.686e-01 3.622e-01  8.681e-21  6.689e-23  2.174e-40  1.850e-70
## 1103 6.694e-11 1.061e-03  4.417e-24  5.265e-37  7.968e-37  8.327e-50
## 1104 2.250e-06 2.562e-01  9.875e-31  5.593e-46  4.364e-63 8.870e-162
## 1105 1.634e-19 1.487e-07  2.098e-70 7.511e-104 5.611e-122 3.724e-286
## 1106 4.042e-02 1.608e-03 1.228e-108 1.038e-127 1.024e-163 5.671e-254
## 1107 6.787e-02 6.339e-04  4.853e-23  2.143e-31  9.112e-51  5.209e-91
## 1108 6.445e-02 7.591e-02  1.245e-03  4.933e-06  5.349e-10  1.255e-15
## 1109 3.403e-08 2.150e-01 7.907e-125 1.407e-194 2.526e-214  0.000e+00
## 1110 1.244e-17 5.286e-09  2.147e-90 1.084e-143 1.102e-203  0.000e+00
## 1111 4.671e-01 5.875e-01  4.304e-07  3.090e-09  3.263e-13  1.344e-32
## 1200 6.814e-01 9.572e-01  1.964e-03  3.367e-04  2.432e-02  3.633e-02
## 1201 8.343e-04 2.398e-01 5.292e-158 1.205e-174 7.178e-146  0.000e+00
## 1202 2.302e-02 6.044e-01  7.860e-81  1.272e-81  1.573e-34  4.054e-20
```

```
## 1203 3.326e-02 5.584e-03 8.162e-301 1.272e-213 1.782e-124 6.106e-109
## 1204 7.111e-01 3.720e-01  1.323e-21  6.133e-14  2.704e-12  1.017e-19
## 1205 8.310e-01 9.212e-01  2.243e-01  2.976e-19  1.088e-12  2.043e-01
## 1206 1.880e-01 6.112e-01  9.160e-02  1.008e-02  2.867e-03  8.998e-03
## 1207 1.606e-06 3.629e-02  1.553e-86  1.527e-92  1.375e-68  8.020e-81
## 1208 2.203e-02 2.246e-03  2.730e-12  2.431e-15  2.356e-06  3.322e-02
## 1209 8.540e-02 3.347e-01  9.312e-08  5.975e-09  5.757e-05  1.311e-01
## 1210 2.897e-01 4.164e-01  1.994e-17  2.198e-05  9.282e-17  3.151e-03
## 1211 7.776e-01 9.382e-01  2.626e-24  2.728e-23  6.465e-31  1.176e-56
## 1212 6.204e-02 3.383e-02  5.400e-29  2.046e-17  1.363e-10  3.528e-16
## 1213 9.756e-01 7.920e-01  6.295e-09  1.670e-16  2.406e-03  6.421e-04
## 1300 1.714e-11 3.397e-10  1.375e-98 1.195e-132 7.494e-178  0.000e+00
## 1301 1.520e-01 9.968e-02  1.566e-02  1.242e-01  5.509e-01  3.172e-03
## 1302 6.109e-01 2.296e-02  4.107e-07  2.666e-15  2.012e-19  5.993e-36
## 1303 8.693e-07 2.104e-07 9.499e-127 1.555e-163 2.813e-188 7.552e-300
## 1304 6.448e-05 2.383e-04  3.736e-44  5.334e-58  9.563e-72  2.341e-97
## 1305 2.456e-02 1.436e-01 2.521e-120 2.465e-143 2.064e-163 1.597e-257
## 1306 1.213e-02 2.769e-02  5.053e-63 3.838e-106 1.289e-125 2.168e-238
## 1307 5.991e-08 7.438e-13  2.204e-91 1.786e-122 1.005e-146 1.680e-239
## 1308 2.906e-03 6.328e-01  8.769e-26  1.468e-44  1.798e-59  1.315e-98
## 1309 1.519e-02 1.344e-01  1.122e-11  5.431e-18  4.234e-24  6.872e-71
## 1310 1.493e-05 1.326e-01  6.599e-07  1.103e-15  5.751e-25  7.113e-62
## 1311 1.824e-16 8.404e-10  2.539e-59  6.598e-78  1.161e-89 4.836e-214
## 1312 4.693e-10 2.724e-08  1.380e-93 3.198e-134 4.734e-169 1.605e-300
## 1313 1.158e-01 1.638e-01  1.784e-40  1.267e-56  1.020e-68 1.659e-100
## 1314 1.169e-01 1.103e-01  7.680e-27  9.675e-33  3.597e-49 3.485e-113
## 1315 2.599e-03 1.783e-01  1.993e-06  5.710e-08  2.051e-12  2.636e-23
## 1400 8.435e-02 4.850e-01  2.297e-64  3.469e-90  3.480e-32  1.759e-28
## 1401 4.571e-01 1.126e-02  5.108e-15  1.067e-15  8.066e-08  3.568e-03
## 1402 3.732e-01 4.150e-02  1.521e-45  1.969e-77  2.673e-29  9.324e-01
## 1403 1.591e-03 9.889e-02 1.564e-107 5.751e-137  1.290e-55  3.387e-14
## 1404 8.215e-03 8.507e-01  2.479e-20  6.408e-30  3.436e-20  1.342e-09
## 1405 1.864e-05 1.193e-02  2.006e-75 2.167e-101  6.171e-40  1.450e-13
## 1406 1.798e-03 6.065e-01  1.279e-94  2.909e-85  5.862e-20  2.333e-04
## 1407 1.039e-05 9.968e-01  3.088e-64  5.054e-91  6.237e-51  2.367e-26
## 1408 2.741e-01 1.406e-01 8.599e-149 3.693e-170  3.076e-65  5.009e-21
## 1409 5.914e-03 7.349e-01  2.936e-07  7.937e-08  4.797e-08  8.435e-05
## 1410 1.034e-01 6.005e-01  3.928e-34  1.936e-43  2.942e-16  7.501e-01
## 1500 1.875e-05 1.758e-06  1.850e-99 1.734e-109 2.483e-121 2.324e-162
## 1501 4.264e-02 1.469e-01  2.455e-11  1.114e-13  1.825e-20  3.763e-12
## 1502 1.675e-02 2.385e-01  6.081e-62  7.851e-90 8.062e-109 5.701e-162
## 1503 7.369e-01 1.969e-02  3.050e-07  4.286e-11  3.105e-11  1.927e-16
## 1504 6.959e-01 4.449e-01  8.539e-09  1.590e-11  5.628e-15  9.072e-21
## 1505 2.182e-02 6.665e-03  2.547e-17  2.722e-23  4.969e-23  1.064e-27
## 1506 9.351e-01 8.759e-01  2.315e-14  1.070e-16  5.587e-17  1.695e-18
## 1507 7.332e-01 1.168e-01  1.678e-48  4.997e-46  1.286e-31  2.690e-37
## 1508 1.066e-01 8.185e-01  1.486e-11  4.663e-16  2.644e-15  9.356e-20
## 1600 1.545e-03 2.380e-14 1.510e-230 5.428e-246 9.279e-280  0.000e+00
## 1601 2.583e-01 3.916e-01  1.529e-12  9.329e-12  2.949e-18  1.437e-16
## 1602 3.572e-03 3.139e-02 3.443e-117 1.858e-147 8.931e-161 6.089e-212
```

```
## 1603 5.395e-02 3.147e-01  1.793e-07  7.908e-09  1.113e-10  3.473e-12
## 1604 1.902e-03 8.687e-01  1.449e-09  6.908e-10  6.793e-13  2.598e-17
## 1605 1.414e-01 1.213e-03  5.725e-34  1.075e-33  4.728e-39  4.869e-48
## 1606 6.078e-03 8.595e-01  2.201e-68  1.191e-86  9.629e-96 6.952e-119
## 1607 6.893e-03 5.151e-01  1.463e-22  2.226e-34  3.854e-32  9.870e-45
## 1700 7.082e-01 4.647e-01  3.076e-26  1.773e-36  1.007e-25  1.661e-44
## 1701 4.740e-01 8.503e-01  2.128e-08  2.090e-04  4.604e-02  2.187e-02
## 1702 1.642e-01 2.840e-02  3.351e-30  3.489e-31  1.763e-17  2.442e-15
## 1703 6.359e-02 2.639e-01  4.599e-44  7.330e-66  3.434e-68 1.640e-108
## 1704 1.095e-01 1.054e-01  4.067e-38  2.146e-54  3.972e-37  4.355e-45
## 1705 9.082e-01 9.849e-01  1.119e-50  1.518e-66  1.743e-56  1.986e-46
## 1706 9.258e-05 1.721e-06 2.980e-135 4.131e-136  1.080e-99  3.531e-91
## 1707 3.929e-01 1.532e-03  7.243e-23  1.531e-19  7.013e-10  1.218e-09
## 1708 6.740e-02 6.749e-01  2.680e-49  2.090e-62  4.468e-49  1.497e-72
## 1709 4.991e-04 3.552e-02  1.277e-11  5.301e-16  2.347e-15  3.465e-20
## 1710 2.980e-02 5.621e-02  2.901e-51  4.284e-64  7.686e-47  2.310e-38
## 1711 7.654e-01 6.447e-04  5.151e-22  6.161e-23  4.849e-13  1.681e-09
## 1712 1.424e-01 8.803e-01 1.476e-142 1.576e-170 4.519e-134 3.521e-146
## 1800 4.128e-02 1.341e-01  7.717e-03  3.031e-08  5.796e-03  4.409e-04
## 1801 6.674e-01 1.016e-01  6.779e-01  6.330e-01  7.760e-01  8.531e-01
## 1802 8.474e-01 1.678e-01  7.465e-55  4.030e-61  1.640e-36  6.736e-09
## 1803 6.013e-01 2.974e-01  1.043e-68  2.910e-83  2.020e-38  3.891e-09
## 1804 5.374e-02 6.623e-02  2.149e-49  4.024e-46  3.667e-28  6.789e-20
## 1900 6.267e-03 2.308e-04  4.277e-70  5.206e-91  3.223e-98 3.806e-144
## 1901 8.558e-01 1.967e-02  1.637e-34  8.882e-38  1.932e-43  6.417e-66
## 1902 1.593e-01 4.413e-01  9.640e-37  1.053e-51  1.293e-52 2.332e-121
## 1903 2.180e-01 6.814e-01  1.829e-12  4.113e-12  3.565e-20  7.554e-43
## 1904 1.906e-01 2.450e-01  9.118e-24  3.458e-31  1.172e-39  1.092e-72
## 1905 2.863e-01 7.491e-01  2.743e-01  3.403e-01  1.481e-01  2.008e-02
## 1906 3.614e-01 6.832e-02  7.192e-18  4.064e-22  1.079e-27  1.077e-47
## 1907 4.166e-01 7.353e-01  9.282e-39  3.770e-53  3.199e-59  2.631e-81
## 1908 1.402e-01 3.047e-03  6.692e-34  7.158e-37  3.300e-41  1.589e-79
## 1909 6.330e-01 5.991e-01  5.837e-46  8.702e-34  3.022e-27  2.188e-17
## 1910 9.298e-01 2.814e-02  1.328e-56  8.128e-70  4.760e-79 6.742e-122
## 1911 3.457e-01 2.758e-04  1.278e-14  5.174e-32  9.211e-41  1.089e-75
## 1912 2.075e-01 1.617e-02  8.928e-72  6.751e-85  2.722e-75 4.339e-127
## 1913 3.543e-01 5.431e-01  4.144e-01  7.095e-01  7.908e-01  4.311e-03
## 2000 1.708e-01 4.852e-02  2.058e-12  3.468e-14  8.007e-02  1.024e-08
## 2001 2.739e-01 8.713e-01  1.220e-14  1.264e-13  5.112e-07  5.223e-02
## 2002 2.914e-08 3.103e-02 1.056e-220 2.207e-220  7.521e-72  4.700e-15
## 2003 8.601e-04 7.977e-01  5.603e-94 2.130e-110  1.173e-38  1.683e-07
## 2100 5.124e-01 6.820e-01  8.682e-08  9.225e-10  5.514e-10  1.244e-15
## 2101 3.914e-01 6.616e-01  1.674e-03  1.451e-05  1.444e-12  1.157e-16
## 2102 6.219e-01 5.820e-01  7.860e-71  3.080e-93  2.457e-99 1.276e-107
## 2103 9.476e-01 2.567e-01  3.875e-27  5.946e-34  3.896e-41  4.928e-43
## 2104 2.956e-03 3.038e-05  3.335e-05  3.002e-08  1.186e-13  1.163e-23
## 2105 4.683e-01 3.553e-04  1.311e-94 4.100e-110 8.823e-130 1.292e-146
## 2200 4.338e-06 2.739e-01  3.306e-87 2.368e-124  8.637e-99 4.541e-208
## 2201 5.273e-01 3.147e-02  4.779e-26  1.552e-48  1.063e-44  2.656e-78
## 2202 1.898e-05 5.613e-07  3.286e-64  1.155e-70  5.253e-37  1.964e-27
```

```
## 2203 7.118e-01 3.471e-01  5.637e-18  3.153e-17  2.846e-15  3.051e-09
## 2204 1.253e-01 1.822e-01  1.115e-88 7.511e-124 5.769e-139 1.921e-207
## 2205 2.936e-01 1.468e-01  2.562e-76  1.225e-74  4.698e-48  4.990e-38
## 2206 3.837e-01 1.389e-01  2.381e-20  1.870e-21  3.636e-12  9.283e-13
## 2207 5.948e-01 2.288e-02  1.537e-48  9.980e-59  1.584e-39  1.258e-35
## 2208 7.312e-01 1.902e-02 4.006e-241 1.793e-266 6.319e-212 1.411e-267
## 2209 2.588e-01 2.306e-03 1.803e-173 2.284e-191 4.088e-125  4.917e-92
## 2210 2.363e-01 2.299e-01 1.757e-219 8.381e-272 1.499e-231  0.000e+00
## 2211 3.873e-01 3.668e-01  4.292e-72  8.941e-93  5.201e-88 8.941e-151
## 2212 8.562e-01 1.292e-01  3.956e-19  3.307e-31  1.279e-22  4.924e-33
## 2213 2.328e-02 3.450e-01  3.663e-32  1.387e-42  1.009e-40  4.909e-44
## 2214 1.494e-02 4.955e-01  1.461e-38  4.528e-33  1.632e-24  4.411e-11
## 2215 4.973e-03 4.107e-02  4.610e-71  7.890e-73  1.913e-48  3.199e-26
## 2216 6.001e-01 6.397e-01  8.462e-11  4.379e-19  2.223e-06  1.744e-03
## 2300 3.058e-03 4.624e-01 3.408e-119 3.962e-132 1.631e-143 1.146e-208
## 2301 1.735e-01 2.918e-02  3.573e-11  6.471e-17  1.035e-05  3.954e-05
## 2302 8.757e-01 9.714e-01  8.252e-12  2.140e-16  2.613e-15  1.045e-26
## 2303 6.328e-07 9.568e-02  8.561e-50  1.474e-72 6.027e-104 1.014e-209
## 2304 1.978e-02 9.980e-02  4.392e-51  3.535e-62  2.939e-77 2.269e-113
## 2305 1.060e-01 3.032e-01 3.706e-118 3.681e-125 5.303e-133 2.563e-162
## 2306 5.428e-01 4.315e-01  6.765e-22  4.764e-34  6.931e-42  1.037e-61
## 2307 6.321e-14 1.517e-01  1.675e-30  9.974e-44  9.362e-50 6.052e-116
## 2308 1.374e-01 8.046e-01  1.376e-63  4.905e-70  2.705e-62  2.592e-93
## 2309 2.283e-01 6.439e-01  1.387e-18  3.398e-24  1.231e-23  7.714e-31
## 2310 8.935e-02 7.987e-03 3.851e-122 3.937e-187 4.636e-181 3.641e-260
## 2311 7.216e-01 9.659e-02 2.290e-119 3.224e-181 4.872e-171 2.362e-244
## 2312 5.629e-01 7.161e-04 9.349e-175 7.293e-216 1.974e-222 4.354e-284
## 2400 2.917e-05 2.960e-02  2.898e-10  4.101e-17  6.542e-21  1.358e-39
## 2401 5.595e-01 2.987e-01  8.298e-05  1.625e-02  6.476e-06  1.700e-06
## 2402 1.002e-01 1.838e-01  4.078e-81 5.620e-106 1.106e-115 8.561e-171
## 2403 2.687e-02 2.250e-17  2.374e-42  2.684e-53  2.134e-70 5.722e-132
## 2404 1.962e-01 5.231e-01  2.487e-28  1.141e-37  3.019e-48  8.798e-83
## 2405 2.474e-03 6.839e-02  9.995e-12  1.072e-18  1.248e-29  1.185e-48
## 2406 1.612e-01 2.317e-08  1.120e-06  1.233e-08  6.920e-12  2.176e-23
## 2500 6.191e-01 5.564e-01  1.250e-70  6.986e-88  8.266e-99 9.025e-230
## 2501 6.426e-01 2.477e-02  4.227e-05  5.468e-07  6.682e-08  1.052e-05
## 2502 3.165e-03 7.049e-05  6.141e-13  8.589e-18  6.683e-23  5.124e-28
## 2503 1.876e-05 9.896e-01  2.913e-10  2.030e-15  1.356e-19  6.179e-25
## 2504 1.923e-01 2.875e-01  2.480e-22  7.395e-37  3.049e-38  4.324e-55
## 2505 5.036e-02 2.642e-01  1.151e-36  9.049e-45  4.294e-60  3.393e-83
## 2506 1.518e-01 9.780e-01 2.374e-135 2.186e-179 1.715e-222 1.612e-267
## 2507 6.808e-02 3.405e-01  3.833e-78  1.310e-99 3.221e-108 3.132e-119
## 2508 5.434e-01 6.112e-01  3.476e-38  2.564e-50  3.416e-60  5.721e-75
## 2600 1.525e-02 1.100e-02  5.805e-69  1.344e-61  4.562e-19  9.524e-12
## 2601 2.975e-02 6.632e-01  3.025e-12  6.477e-09  4.276e-13  5.167e-07
## 2602 2.299e-01 2.898e-02  2.553e-15  1.288e-16  5.854e-02  5.415e-04
## 2603 2.337e-01 7.252e-01  2.226e-30  3.914e-23  1.441e-05  1.148e-07
## 2604 3.315e-06 8.920e-03 6.816e-133 7.575e-136  2.521e-70  1.435e-51
## 2605 6.242e-03 2.858e-01  5.707e-20  4.780e-24  1.179e-24  1.654e-32
## 2606 3.368e-01 6.475e-03  8.522e-08  7.061e-07  2.061e-05  8.750e-04
```

```
## 2607 8.134e-01 1.558e-01  2.819e-09  7.716e-07  5.488e-05  2.851e-05
## 2608 4.749e-01 1.012e-02  6.860e-17  4.204e-12  2.717e-04  6.800e-06
## 2609 7.789e-01 2.093e-01  2.148e-03  2.931e-04  2.610e-02  3.165e-01
## 2610 6.264e-01 6.729e-01  2.388e-40  4.821e-50  6.300e-32  4.050e-19
## 2611 4.448e-01 9.266e-01  3.762e-22  2.723e-32  5.688e-37  2.508e-61
## 2612 2.549e-01 1.283e-02  1.688e-12  7.908e-20  6.107e-03  1.036e-04
## 2613 8.376e-04 5.390e-02  5.055e-79 4.958e-122  2.052e-67  2.655e-90
## 2614 2.404e-01 2.368e-01  3.619e-21  2.819e-26  2.184e-16  3.007e-21
## 2700 8.490e-01 8.512e-11  0.000e+00  0.000e+00  0.000e+00  0.000e+00
## 2701 6.274e-13 1.234e-01 2.866e-150 9.355e-256  0.000e+00  0.000e+00
## 2702 7.198e-01 1.037e-01  1.038e-04  7.288e-07  2.938e-05  5.121e-32
## 2703 3.872e-03 2.375e-02  6.651e-66 9.530e-129 8.319e-173  0.000e+00
## 2704 6.649e-01 9.237e-01  6.880e-08  8.552e-25  3.300e-43  2.769e-98
## 2705 4.534e-11 2.954e-09  1.132e-87 3.279e-197  0.000e+00  0.000e+00
## 2706 6.329e-01 4.698e-01  1.866e-15  2.207e-42  2.942e-71 7.276e-193
## 2707 2.225e-02 5.797e-01  5.165e-29  1.522e-61  1.521e-70 2.485e-173
## 2708 8.937e-04 2.227e-02  4.965e-41  7.955e-47  4.443e-72 4.347e-167
## 2709 4.439e-01 7.720e-02  1.959e-03  1.515e-02  4.836e-03  2.084e-03
## 2710 4.515e-01 1.329e-01  8.493e-04  1.850e-04  9.445e-06  2.267e-07
## 2711 1.361e-02 4.102e-02  4.224e-95 2.025e-216 9.556e-304  0.000e+00
## 2712 2.682e-02 4.085e-01  1.459e-37  1.582e-92 9.262e-145 3.714e-279
## 2713 7.289e-02 6.428e-02  1.336e-24  5.112e-53  2.456e-71 1.381e-123
## 2714 2.092e-01 9.552e-01  3.770e-23  1.400e-39  2.551e-47 2.595e-104
## 2715 6.536e-04 1.048e-03  7.199e-29  1.618e-59 3.545e-126  0.000e+00
## 2716 3.198e-12 5.773e-12  7.348e-26  8.218e-28  3.431e-28  3.096e-65
## 2717 7.681e-02 2.617e-01  6.878e-84 2.942e-148 1.656e-212  0.000e+00
## 2718 3.558e-01 9.663e-03  6.227e-19  6.717e-16  1.539e-29  8.805e-56
## 2719 5.744e-01 4.521e-03 2.503e-134 1.274e-245 6.726e-295  0.000e+00
## 2720 1.775e-01 6.992e-03  1.989e-53  2.911e-70  6.583e-94 3.264e-250
## 2721 9.884e-01 7.410e-01  7.430e-05  4.514e-07  3.055e-26  3.103e-61
## 2722 3.866e-01 2.787e-01  4.522e-03  1.669e-05  9.955e-08  2.020e-18
## 2723 1.550e-01 1.160e-02  3.513e-24  9.507e-42  1.711e-74 1.847e-162
## 2724 8.191e-01 2.739e-01  7.236e-53 5.020e-119 1.123e-177 7.033e-302
## 2725 2.676e-01 6.420e-02  9.916e-83 2.446e-116 6.207e-156 4.833e-250
## 2726 6.469e-03 1.187e-01  1.883e-60  2.209e-96 8.775e-160 4.358e-292
## 2727 4.995e-03 1.941e-06  1.842e-14  1.466e-23  1.858e-52 1.695e-148
## 2728 2.653e-04 4.377e-03 1.404e-117 1.091e-213 2.902e-306  0.000e+00
## 2729 2.082e-04 9.430e-03  6.849e-68 1.399e-140 5.339e-181  0.000e+00
## 2730 7.866e-01 1.651e-02 1.869e-166 1.001e-305  0.000e+00  0.000e+00
## 2731 7.473e-04 7.652e-01  2.929e-62  5.367e-99 1.923e-155  0.000e+00
## 2732 1.436e-01 4.286e-01 1.303e-163  0.000e+00  0.000e+00  0.000e+00
## 2733 2.729e-07 1.293e-10  1.652e-17  9.561e-35  1.174e-74 1.529e-187
## 2734 4.171e-01 5.283e-02  3.973e-12  1.340e-32  8.964e-35 1.173e-113
## 2735 7.389e-22 1.260e-04 1.190e-138 1.115e-288  0.000e+00  0.000e+00
## 2736 3.313e-01 5.129e-05  9.646e-76 6.155e-132 3.987e-174 1.507e-297
## 2737 4.084e-01 3.215e-01  2.324e-07  2.610e-11  5.123e-17  4.426e-79
## 2738 1.306e-06 5.669e-01  0.000e+00  0.000e+00  0.000e+00  0.000e+00
## 2739 1.446e-09 8.365e-05 8.901e-291  0.000e+00  0.000e+00  0.000e+00
## 2740 2.277e-01 6.070e-01  1.420e-23  3.144e-70 3.358e-106 1.779e-322
## 2741 3.233e-01 1.519e-05  4.096e-78 4.008e-169 8.646e-316  0.000e+00
```

```
## 2742 1.777e-01 1.775e-09  1.056e-87 1.420e-174 5.172e-225  0.000e+00
## 2743 1.235e-02 2.903e-01  4.762e-11  5.484e-21  2.071e-21  2.633e-52
## 2744        NA        NA         NA         NA         NA         NA
## 2745 3.282e-02 1.002e-01  2.537e-12  1.914e-19  1.064e-37  8.474e-69
## 2746 4.488e-01 6.541e-02 5.472e-163  0.000e+00  0.000e+00  0.000e+00
## 2747 1.410e-01 1.812e-03  1.086e-03  4.304e-08  4.971e-22  7.531e-48
## 2748 3.637e-05 1.528e-01 2.901e-110 2.645e-187 9.969e-253  0.000e+00
## 2800 6.347e-01 3.298e-07 5.104e-134 3.904e-169 1.306e-200  0.000e+00
## 2801 2.277e-01 1.424e-01  8.864e-03  2.322e-03  7.983e-04  4.773e-10
## 2802 4.571e-05 6.719e-04  2.012e-08  3.771e-09  5.275e-15  7.890e-31
## 2803 2.776e-02 3.554e-01  2.489e-02  4.962e-05  7.976e-08  5.830e-19
## 2804 2.002e-02 5.076e-01  7.074e-07  1.621e-12  4.041e-17  1.215e-31
## 2805 3.309e-01 4.667e-01  1.802e-20  9.195e-25  1.597e-25  2.125e-37
## 2806 3.249e-01 5.828e-02  6.309e-10  2.296e-13  5.492e-16  2.424e-29
## 2807 1.859e-01 3.096e-03  5.637e-06  1.303e-05  6.038e-09  5.195e-09
## 2808 3.617e-05 1.249e-02  3.218e-22  1.869e-36  1.165e-52 1.033e-105
## 2809 6.771e-02 6.346e-03  7.556e-50  3.779e-62  1.197e-75 1.495e-127
## 2900 5.363e-03 1.314e-02  2.256e-91 1.051e-175 8.801e-224  0.000e+00
## 2901 2.552e-01 7.503e-01  1.167e-10  5.446e-27  6.149e-24  1.844e-34
## 2902 2.323e-01 1.754e-01  1.859e-54 3.677e-143 1.936e-252  0.000e+00
## 2903 4.825e-01 3.767e-01  5.163e-08  4.569e-08  1.928e-09  8.806e-78
## 2904 8.301e-01 4.646e-01  1.712e-03  8.560e-09  7.270e-14  5.575e-48
## 2905 1.075e-02 2.291e-01  2.477e-47  3.059e-53  4.286e-44  3.807e-78
## 2906 8.649e-06 8.636e-01  1.400e-32  5.463e-47  1.420e-51  1.428e-62
## 2907 4.398e-01 7.896e-01  1.786e-37  1.628e-57 3.553e-104 8.441e-235
## 2908 3.161e-02 2.296e-01  3.011e-01  1.198e-07  2.418e-14  5.841e-55
## 2909 7.711e-03 7.786e-02  3.047e-30  4.025e-52  7.989e-63 1.292e-108
## 2910 3.204e-01 8.061e-01  9.712e-06  1.523e-08  2.094e-08  9.962e-18
## 2911 1.419e-02 3.274e-01  3.300e-66  1.564e-91  6.965e-91 1.828e-211
## 2912 7.674e-01 4.140e-05  2.734e-80  1.168e-92  2.050e-93  6.769e-80
## 2913 6.149e-25 8.804e-14  8.779e-04  1.521e-17  7.790e-22  4.799e-26
## 2914 1.281e-04 5.978e-01  1.682e-08  1.069e-19  1.388e-16  6.908e-25
## 2915        NA        NA         NA         NA         NA         NA
## 2916 1.439e-01 1.176e-07  5.710e-52  1.425e-87 3.493e-123 1.030e-195
## 2917 1.214e-02 4.856e-01  2.347e-05  1.855e-12  1.082e-23  3.048e-43
## 2918        NA        NA         NA         NA         NA         NA
## 2919 4.173e-15 1.127e-03  5.094e-05  8.750e-16  5.977e-14  2.393e-32
## 2920 4.797e-07 2.085e-01  5.929e-21  1.880e-18  1.649e-21  2.456e-33
## 2921 4.807e-03 1.245e-01  7.350e-20  6.618e-46  9.386e-51  7.238e-98
## 2922 2.185e-02 5.890e-01  6.908e-01  6.073e-03  2.056e-03  1.625e-03
## 2923 3.281e-01 1.570e-02  2.881e-05  8.347e-09  1.597e-05  1.017e-04
## 3000 3.362e-01 1.666e-03  6.365e-16  5.994e-18  8.465e-23  2.188e-40
## 3001        NA        NA         NA         NA         NA         NA
## 3002 1.666e-01 3.797e-02  1.217e-10  3.231e-13  7.372e-17  4.107e-36
## 3003 3.769e-01 1.332e-08 5.487e-134 3.709e-203 4.682e-232 1.405e-299
## 3004 1.211e-07 9.606e-01  5.556e-82 3.135e-125 1.789e-152 2.304e-250
## 3005 1.723e-07 3.624e-08  2.851e-16  6.097e-26  2.493e-32  4.616e-53
## 3100 7.489e-04 3.857e-01 3.620e-261  0.000e+00  0.000e+00  0.000e+00
## 3101 2.420e-01 8.903e-03  1.035e-74 1.147e-109 5.100e-107 9.518e-140
## 3102 5.242e-03 3.991e-06  5.294e-11  4.490e-13  3.360e-18  2.865e-21
```

```
## 3103 3.051e-01 1.892e-01  8.599e-77  7.165e-95  2.623e-74 1.271e-118
## 3104 6.606e-02 4.338e-01  6.725e-58  3.641e-57  3.238e-59 3.223e-127
## 3105 4.098e-02 6.231e-01  9.428e-36  2.479e-56  4.397e-54  1.032e-69
## 3106 2.695e-01 5.496e-02 1.633e-189 3.505e-235 7.051e-134 9.605e-107
## 3107 6.074e-02 2.617e-01  4.742e-71 1.445e-120 1.078e-129 7.491e-244
## 3108 2.333e-01 7.347e-01  3.218e-04  2.716e-15  1.466e-18  1.046e-44
## 3109 1.428e-01 8.223e-02  1.030e-30  3.885e-53  5.127e-34  4.453e-39
## 3110 2.747e-02 6.790e-02  1.562e-15  5.242e-23  2.435e-27  6.948e-34
## 3200 1.054e-02 1.216e-01 6.949e-170 1.341e-219 1.985e-216 3.119e-300
## 3201 7.530e-01 1.211e-01  1.944e-09  2.688e-10  1.038e-10  1.412e-09
## 3202 6.144e-01 8.205e-01 5.027e-138 4.538e-197 1.873e-174 3.148e-244
## 3203 9.967e-01 6.081e-05 8.867e-293  0.000e+00  0.000e+00  0.000e+00
## 3204 3.656e-03 1.689e-01 5.955e-107 4.127e-141 2.674e-173 6.487e-308
## 3205 1.031e-03 5.421e-02  6.891e-49  8.310e-50  7.011e-40  5.003e-62
## 3206 3.876e-03 8.082e-01  7.564e-48  2.815e-67  2.766e-70  1.313e-96
## 3207 1.593e-02 8.607e-01 2.421e-124 7.386e-166 3.570e-158 5.871e-170
## 3300 4.746e-03 2.482e-01  5.824e-66  4.770e-62  1.170e-29  8.544e-26
## 3301 7.647e-02 4.265e-02 1.838e-146 2.369e-138 6.717e-100  3.621e-93
## 3302 2.805e-01 9.159e-01  4.706e-17  1.910e-09  3.821e-08  2.002e-13
## 3303 1.460e-01 2.118e-03  6.801e-39  2.834e-20  1.934e-09  9.167e-06
## 3304 2.920e-35 1.371e-06  0.000e+00  0.000e+00  0.000e+00  0.000e+00
## 3305 6.538e-08 3.898e-03  9.312e-38  2.199e-42  1.905e-22  1.336e-31
## 3306 1.389e-01 1.210e-03  4.415e-97 6.560e-172 2.516e-172  0.000e+00
## 3307 7.800e-01 3.554e-01  1.055e-18  1.105e-18  7.539e-18  3.133e-27
## 3308 3.645e-05 1.467e-03 1.653e-149 4.216e-176 7.713e-133 8.446e-116
## 3309 3.501e-05 2.401e-05 1.948e-158 2.891e-134  5.987e-63  2.011e-79
## 3310 5.466e-01 2.636e-02 1.100e-315 5.143e-241 2.588e-147 1.467e-126
## 3311 6.155e-01 6.755e-01  1.529e-10  1.422e-07  1.056e-08  3.792e-07
## 3312 2.305e-03 2.624e-02  0.000e+00  0.000e+00 2.181e-192 4.354e-165
## 3313 3.606e-04 4.442e-01  6.800e-14  2.627e-13  6.844e-05  2.755e-02
## 3314 5.157e-01 7.907e-01 6.477e-100  1.779e-79  7.857e-42  4.969e-29
## 3315 1.697e-01 5.197e-03 1.252e-117  7.337e-96  1.206e-63  1.135e-94
## 3316 1.974e-04 8.126e-01  2.669e-53  2.706e-51  9.300e-29  8.990e-31
## 3317 2.713e-03 6.147e-01  1.665e-14  1.083e-21  2.571e-12  4.514e-20
## 3318 2.560e-01 1.679e-03  2.578e-58  4.719e-52  5.696e-48  3.556e-32
## 3319 3.530e-01 5.706e-01  2.083e-53  2.945e-79  2.136e-84  1.764e-94
## 3320 2.652e-01 2.700e-01 9.583e-111  1.523e-52  3.037e-24  3.081e-07
## 3321 4.784e-02 1.790e-01  3.012e-40  4.683e-21  1.184e-10  4.295e-06
## 3322 8.590e-01 9.864e-01  2.587e-20  1.390e-15  7.759e-04  5.663e-06
## 3400 1.721e-01 1.393e-02  4.641e-77 2.089e-160 2.265e-197 4.387e-281
## 3401 9.342e-02 3.510e-01  1.450e-01  8.212e-03  8.858e-01  1.633e-01
## 3402 9.637e-03 8.236e-01  2.060e-01  5.822e-02  2.740e-02  1.631e-05
## 3403 5.983e-02 3.745e-04  6.455e-01  7.857e-01  5.756e-01  6.318e-01
## 3404 9.137e-01 2.673e-01  2.063e-18  2.769e-39  8.393e-59  1.290e-72
## 3500 6.561e-03 7.496e-01  6.385e-64 1.614e-135 1.432e-161 2.127e-230
## 3501 7.584e-02 4.163e-01  1.756e-01  3.015e-01  2.663e-01  6.155e-02
## 3502        NA        NA         NA         NA         NA         NA
## 3503        NA        NA         NA         NA         NA         NA
## 3504 8.236e-03 5.522e-02  9.183e-03  5.654e-11  5.030e-17  7.845e-35
## 3505 4.243e-01 8.666e-02  1.417e-11  3.151e-13  2.361e-15  2.598e-19
```

```
## 3506 9.231e-01 8.778e-01  7.766e-02  7.585e-02  1.542e-02  1.593e-03
## 3600 8.382e-02 9.153e-01  1.067e-38  6.698e-63  1.706e-71 1.204e-129
## 3601 1.061e-01 1.614e-01  1.360e-09  8.156e-18  7.730e-16  8.483e-26
## 3602 1.411e-01 2.139e-01  5.644e-03  2.955e-04  7.074e-05  1.501e-13
## 3603 6.598e-03 8.200e-01  1.146e-02  9.803e-04  2.404e-03  1.575e-10
## 3604        NA        NA         NA         NA         NA         NA
## 3605 6.508e-02 3.562e-02  7.785e-04  2.227e-04  9.795e-06  1.161e-13
## 3606        NA        NA         NA         NA         NA         NA
## 3607 3.011e-02 2.393e-01  8.518e-03  3.830e-02  4.669e-04  1.065e-11
## 3608        NA        NA         NA         NA         NA         NA
## 3609 8.839e-01 3.783e-03  2.305e-02  7.779e-04  1.129e-06  7.124e-05
## 3610 1.635e-01 9.190e-02  7.589e-02  7.696e-05  1.522e-09  2.389e-13
## 3611 1.116e-02 7.373e-01  2.702e-23  1.766e-35  5.086e-48  1.397e-84
## 3612 9.676e-01 3.069e-03 5.705e-114 2.121e-217 3.005e-309  0.000e+00
## 3613 5.301e-01 2.016e-01  2.033e-02  1.508e-02  3.126e-04  1.058e-05
## 3614 4.669e-08 3.264e-13  2.528e-23  9.126e-38  1.740e-59 1.785e-107
## 3615        NA        NA         NA         NA         NA         NA
## 3616 5.049e-03 1.751e-03  3.503e-42  7.785e-52  2.636e-52  4.366e-82
##          FFA2p      FLA2p      FFA3p      FLA4p
## 1000 6.418e-01  1.747e-06  4.148e-01  1.742e-06
## 1100 3.438e-03  1.354e-04  2.302e-02  7.811e-04
## 1101 5.975e-01  4.106e-10  5.429e-03  2.480e-11
## 1102 3.586e-01  4.457e-01  2.884e-01  3.510e-01
## 1103 6.243e-16  3.609e-03  2.106e-19  5.257e-07
## 1104 4.442e-08  3.049e-01  1.231e-08  5.730e-02
## 1105 1.698e-28  9.775e-06  5.273e-34  9.536e-12
## 1106 4.943e-09  4.404e-06  4.360e-06  1.881e-03
## 1107 2.267e-01  3.711e-04  6.646e-02  1.652e-04
## 1108 2.979e-02  1.287e-01  1.437e-02  5.840e-02
## 1109 7.270e-33  6.884e-01  3.831e-32  2.272e-01
## 1110 1.437e-30  1.668e-05  7.370e-36  1.684e-11
## 1111 9.667e-01  3.265e-01  8.490e-01  3.185e-01
## 1200 5.407e-01  3.287e-01  5.242e-01  3.182e-01
## 1201 1.676e-06  1.103e-01  1.426e-05  4.925e-01
## 1202 6.960e-02  6.948e-01  1.547e-04  1.034e-03
## 1203 5.967e-06  5.320e-03  2.813e-15  7.602e-13
## 1204 5.484e-01  9.568e-01  4.253e-01  5.761e-01
## 1205 7.763e-01  5.888e-01  3.888e-01  3.406e-01
## 1206 3.522e-01  9.154e-01  3.154e-02  6.217e-02
## 1207 1.562e-14  3.371e-02  1.188e-39  1.163e-28
## 1208 2.156e-03  1.296e-04  2.221e-03  5.891e-05
## 1209 1.100e-01  1.306e-01  1.853e-04  2.230e-04
## 1210 3.075e-01  2.964e-01  8.169e-01  7.488e-01
## 1211 5.668e-03  8.386e-01  1.226e-05  2.969e-04
## 1212 6.337e-02  3.027e-01  4.466e-02  2.209e-01
## 1213 7.996e-01  6.058e-01  5.954e-01  4.863e-01
## 1300 6.635e-02  9.030e-13  3.474e-03  1.402e-13
## 1301 7.543e-02  4.299e-02  1.604e-01  9.096e-02
## 1302 8.290e-01  1.775e-01  6.362e-01  1.626e-01
## 1303 3.971e-01  4.506e-07  1.861e-01  2.844e-07
```

```
## 1304 1.871e-01  3.798e-04  8.866e-02  2.223e-04
## 1305 4.033e-01  1.993e-02  6.131e-01  2.617e-02
## 1306 8.611e-01  6.149e-05  4.749e-01  5.131e-05
## 1307 9.587e-03  8.798e-17  7.846e-04  9.971e-18
## 1308 4.009e-04  8.142e-01  4.414e-04  9.290e-01
## 1309 1.668e-01  1.706e-02  8.957e-02  9.859e-03
## 1310 1.562e-08  7.100e-01  8.675e-09  2.223e-01
## 1311 4.148e-07  7.292e-12  1.614e-09  3.287e-14
## 1312 2.612e-04  2.082e-10  1.650e-05  1.533e-11
## 1313 1.529e-02  9.200e-02  2.027e-02  1.231e-01
## 1314 6.496e-01  2.921e-03  4.154e-01  2.312e-03
## 1315 4.367e-02  2.540e-01  3.026e-02  1.637e-01
## 1400 1.539e-02  2.170e-01  2.372e-04  2.320e-03
## 1401 3.212e-01  9.458e-03  2.404e-03  1.111e-04
## 1402 5.044e-01  8.519e-02  1.213e-01  2.836e-02
## 1403 1.338e-05  2.877e-01  1.145e-05  1.822e-01
## 1404 1.100e-03  6.827e-01  1.192e-03  3.965e-01
## 1405 1.837e-07  3.479e-03  3.305e-14  3.213e-10
## 1406 9.925e-05  8.041e-01  3.457e-05  7.945e-02
## 1407 2.305e-07  5.394e-01  5.275e-10  1.864e-04
## 1408 1.083e-01  1.040e-01  3.714e-01  3.511e-01
## 1409 3.171e-03  7.004e-01  4.463e-04  3.688e-02
## 1410 7.233e-02  7.874e-01  2.522e-02  1.359e-01
## 1500 4.495e-01  3.071e-05  1.423e-01  1.735e-05
## 1501 3.302e-04  3.494e-01  1.097e-04  9.312e-02
## 1502 6.650e-01  1.826e-02  4.399e-01  1.571e-02
## 1503 8.639e-01  1.972e-02  9.610e-01  2.023e-02
## 1504 3.769e-02  9.792e-01  3.876e-02  7.157e-01
## 1505 1.785e-01  1.303e-02  1.376e-01  1.060e-02
## 1506 3.511e-01  5.323e-01  3.216e-01  4.810e-01
## 1507 1.789e-01  3.403e-01  3.001e-01  5.695e-01
## 1508 1.613e-02  9.437e-01  1.595e-02  7.361e-01
## 1600 5.855e-01  8.991e-15  6.877e-01  2.066e-14
## 1601 6.487e-01  2.228e-01  5.349e-01  1.977e-01
## 1602 2.288e-10  5.238e-02  9.917e-10  2.202e-01
## 1603 2.294e-01  2.700e-01  2.729e-01  3.228e-01
## 1604 1.579e-04  9.184e-01  1.595e-04  8.447e-01
## 1605 1.257e-02  9.885e-03  1.727e-02  1.356e-02
## 1606 9.487e-01  5.088e-01  9.016e-01  5.046e-01
## 1607 9.076e-06  2.752e-01  5.599e-06  1.388e-01
## 1700 3.832e-01  5.563e-01  1.642e-01  2.092e-01
## 1701 5.680e-01  8.585e-01  4.624e-01  6.001e-01
## 1702 9.654e-02  4.736e-02  4.876e-03  2.459e-03
## 1703 9.333e-01  6.310e-01  8.144e-01  6.011e-01
## 1704 1.417e-01  2.013e-01  5.745e-02  7.657e-02
## 1705 8.967e-01  4.532e-01  6.269e-01  3.752e-01
## 1706 4.317e-04  3.377e-09  4.120e-09  9.897e-14
## 1707 7.797e-01  3.128e-03  3.265e-01  2.116e-03
## 1708 2.869e-02  6.756e-01  3.457e-02  8.469e-01
## 1709 1.085e-04  1.142e-01  1.587e-06  6.680e-04
```

```
## 1710 1.936e-02  1.329e-03  1.180e-05  1.217e-06
## 1711 7.483e-01  1.007e-03  7.042e-01  1.274e-03
## 1712 2.474e-02  9.749e-01  1.991e-02  4.628e-01
## 1800 2.537e-02  9.102e-02  2.420e-03  8.269e-03
## 1801 5.484e-01  7.766e-02  7.483e-01  1.019e-01
## 1802 4.378e-01  4.900e-02  9.441e-01  7.939e-02
## 1803 1.788e-01  5.418e-01  2.450e-01  9.134e-01
## 1804 8.886e-01  2.414e-01  4.722e-01  1.719e-01
## 1900 2.436e-02  2.944e-06  5.790e-05  2.241e-08
## 1901 2.192e-01  2.780e-02  4.069e-01  4.607e-02
## 1902 4.618e-03  9.227e-01  4.498e-03  7.188e-01
## 1903 9.398e-01  7.602e-02  7.289e-01  7.482e-02
## 1904 4.818e-02  8.889e-02  1.331e-01  2.357e-01
## 1905 1.808e-01  7.723e-01  1.662e-01  8.496e-01
## 1906 7.022e-01  2.479e-01  8.419e-01  2.684e-01
## 1907 3.603e-01  5.570e-01  2.984e-01  4.419e-01
## 1908 9.203e-01  1.498e-02  7.385e-01  1.457e-02
## 1909 8.025e-02  6.989e-01  6.169e-02  4.150e-01
## 1910 1.338e-02  1.016e-01  2.912e-02  2.152e-01
## 1911 4.242e-02  3.108e-04  2.043e-01  1.144e-03
## 1912 1.330e-01  1.468e-01  2.131e-01  2.313e-01
## 1913 3.805e-01  5.429e-01  4.602e-01  7.224e-01
## 2000 1.874e-01  8.860e-03  2.178e-04  1.322e-05
## 2001 3.336e-01  7.197e-01  1.307e-01  2.199e-01
## 2002 7.068e-11  4.884e-03  1.886e-20  4.418e-13
## 2003 8.211e-05  3.278e-01  9.589e-07  1.816e-03
## 2100 7.878e-01  6.766e-01         NA  7.055e-01
## 2101 9.406e-01  5.684e-01  8.408e-01  5.314e-01
## 2102 1.574e-02  6.725e-01  2.185e-02  9.681e-01
## 2103 4.433e-01  3.310e-01  3.597e-01  2.759e-01
## 2104 2.918e-02  6.650e-05  1.913e-03  6.177e-06
## 2105 5.794e-01  4.319e-06  2.911e-01  5.179e-06
## 2200 9.386e-12  3.777e-01  1.584e-14  1.309e-04
## 2201 8.366e-01  2.532e-04  4.663e-03  2.487e-06
## 2202 4.924e-05  9.317e-08  3.129e-09  7.850e-12
## 2203 4.191e-01  3.802e-01  5.801e-01  4.908e-01
## 2204 1.985e-04  2.757e-02  8.293e-04  1.010e-01
## 2205 1.501e-02  3.513e-01  2.670e-02  7.276e-01
## 2206 6.112e-01  1.374e-01  3.525e-01  9.195e-02
## 2207 1.620e-01  4.430e-02  3.881e-01  9.159e-02
## 2208 9.133e-02  4.874e-02  2.102e-01  1.072e-01
## 2209 9.315e-04  6.952e-04  3.627e-02  1.747e-02
## 2210 6.325e-03  7.811e-01  5.493e-03  4.172e-01
## 2211 1.159e-05  5.739e-02  1.967e-06  7.349e-03
## 2212 6.343e-01  3.395e-01  8.413e-01  3.942e-01
## 2213 1.606e-04  7.712e-01  1.078e-04  2.482e-01
## 2214 5.500e-02  1.284e-01  2.128e-01  6.409e-01
## 2215 2.592e-06  6.922e-01  3.835e-06  5.942e-01
## 2216 8.612e-01  6.260e-01  7.756e-01  5.825e-01
## 2300 8.928e-10  9.522e-01  8.903e-10  1.359e-01
```

```
## 2301 2.773e-01  9.526e-02  9.778e-01  2.151e-01
## 2302 4.216e-01  6.966e-01  3.758e-01  5.921e-01
## 2303 2.712e-10  2.653e-01  4.786e-11  2.340e-02
## 2304 1.470e-05  9.013e-02  4.887e-06  2.536e-02
## 2305 1.055e-04  4.065e-01  2.318e-04  9.007e-01
## 2306 7.155e-02  5.731e-01  5.144e-02  3.207e-01
## 2307 9.349e-25  1.823e-01  1.738e-25  7.339e-03
## 2308 3.735e-01  2.909e-01  1.863e-01  1.570e-01
## 2309 4.593e-01  3.600e-01  3.565e-01  2.885e-01
## 2310 1.658e-06  2.881e-02  1.452e-05  1.705e-01
## 2311 3.796e-02  1.533e-02  1.360e-01  4.744e-02
## 2312 2.153e-03  1.362e-02  8.852e-03  4.468e-02
## 2400 2.510e-02  3.034e-02  1.057e-02  1.295e-02
## 2401 1.518e-02  4.700e-01  1.885e-02  8.705e-01
## 2402 2.384e-01  1.651e-01  3.098e-01  2.085e-01
## 2403 2.146e-01  3.685e-22  3.420e-02  9.133e-23
## 2404 8.463e-01  7.394e-01  8.242e-01  7.265e-01
## 2405 6.649e-06  3.780e-02  1.260e-05  8.074e-02
## 2406 2.200e-01  3.085e-10  1.040e-01  1.782e-10
## 2500 4.779e-05  8.749e-01  3.657e-05  4.170e-01
## 2501 3.525e-01  4.415e-02  1.593e-02  3.273e-03
## 2502 9.853e-05  7.968e-05  1.677e-05  1.299e-05
## 2503 5.317e-09  4.029e-01  3.820e-09  1.815e-01
## 2504 7.251e-01  8.131e-02  5.572e-01  7.229e-02
## 2505 2.176e-05  5.260e-02  9.627e-06  1.999e-02
## 2506 1.195e-09  2.673e-01  9.979e-10  4.726e-02
## 2507 9.031e-07  7.515e-01  1.402e-06  6.960e-01
## 2508 9.567e-02  8.422e-01  9.274e-02  7.140e-01
## 2600 1.274e-01  8.004e-02  1.239e-03  8.465e-04
## 2601 3.083e-03  6.511e-01  8.826e-05  5.347e-03
## 2602 4.207e-01  1.450e-01  8.531e-01  2.220e-01
## 2603 9.476e-01  7.518e-01  7.567e-01  6.631e-01
## 2604 3.458e-02  2.336e-02  1.396e-03  9.553e-04
## 2605 2.403e-01  1.320e-01  4.606e-01  2.285e-01
## 2606 2.261e-01  5.581e-03  9.617e-01  1.283e-02
## 2607 5.845e-01  2.274e-01  9.403e-01  2.761e-01
## 2608 3.594e-01  1.175e-02  3.348e-01  1.174e-02
## 2609 9.155e-01  2.740e-01  2.565e-01  1.124e-01
## 2610 3.618e-01  5.498e-01  2.275e-01  3.186e-01
## 2611 5.004e-01  5.017e-01  3.738e-01  3.755e-01
## 2612 7.691e-01  5.780e-03  1.097e-01  1.547e-03
## 2613 9.396e-01  4.008e-01  7.084e-01  3.661e-01
## 2614 1.364e-01  7.022e-01  1.526e-01  8.146e-01
## 2700 2.751e-32 4.314e-118  3.872e-01  2.162e-88
## 2701 3.005e-44  5.934e-02  9.414e-46  8.952e-07
## 2702 6.336e-01  7.175e-03  9.531e-01  7.980e-03
## 2703 1.248e-10  2.041e-02  5.441e-12  3.151e-04
## 2704 1.099e-04  4.901e-01  8.326e-05  2.344e-01
## 2705 1.053e-25  9.618e-02  2.636e-27  4.895e-04
## 2706 9.752e-01  1.118e-04  4.269e-01  8.824e-05
```

```
## 2707 7.977e-03  1.496e-04  3.209e-01  5.238e-03
## 2708 4.479e-01  1.833e-02  7.389e-01  2.515e-02
## 2709 5.873e-01  2.005e-01  2.209e-01  9.508e-02
## 2710 2.042e-01  2.355e-01  1.567e-01  1.765e-01
## 2711 3.938e-01  1.060e-06  6.376e-02  4.527e-07
## 2712 1.836e-10  1.904e-03  5.127e-09  4.312e-02
## 2713 6.104e-12  8.186e-02  3.911e-11  6.792e-01
## 2714 7.619e-03  9.653e-01  5.566e-03  3.963e-01
## 2715 4.993e-02  8.017e-15  1.723e-04  2.170e-16
## 2716 4.467e-06  9.811e-13  2.076e-09  6.899e-16
## 2717 1.678e-02  3.929e-13  4.915e-01  2.475e-11
## 2718 5.944e-01  1.804e-06  1.260e-01  1.268e-06
## 2719 1.278e-07  3.076e-04         NA  1.044e-01
## 2720 3.130e-08  1.186e-03  3.201e-07  1.070e-02
## 2721 7.320e-02  9.155e-01  7.208e-02  7.779e-01
## 2722 1.362e-01  1.320e-01  1.891e-01  1.847e-01
## 2723 4.422e-02  1.799e-03  9.960e-02  3.785e-03
## 2724 1.244e-06  1.920e-03  1.484e-05  1.897e-02
## 2725 5.668e-03  1.880e-02  1.142e-02  3.856e-02
## 2726 1.175e-16  1.934e-02  1.771e-17  9.596e-04
## 2727 2.668e-01  7.563e-09  1.857e-02  1.966e-09
## 2728 2.226e-10  1.810e-01  9.637e-10  8.570e-01
## 2729 8.397e-12  7.822e-01  7.889e-12  1.700e-01
## 2730 1.709e-04  5.328e-13  1.179e-02  4.108e-11
## 2731 7.966e-01  5.471e-01  8.613e-01  5.691e-01
## 2732 3.012e-01  2.570e-04  7.096e-02  1.083e-04
## 2733 2.822e-10  3.620e-09  2.029e-15  2.170e-14
## 2734 1.334e-01  1.255e-03  3.381e-01  2.789e-03
## 2735 1.720e-38  3.993e-01  1.393e-38  2.406e-02
## 2736 4.166e-03  8.297e-11  6.708e-02  1.232e-09
## 2737 2.096e-01  3.090e-01  2.560e-01  3.831e-01
## 2738 6.751e-19  4.842e-06  1.222e-15  3.433e-03
## 2739 4.873e-35  2.213e-14  9.834e-25  3.205e-05
## 2740 1.395e-03  9.970e-02  3.674e-03  2.846e-01
## 2741 3.580e-02  1.165e-13  4.374e-01  1.363e-12
## 2742 4.252e-01  2.600e-26  1.362e-02  1.680e-26
## 2743 5.684e-01  4.555e-02  3.142e-01  3.176e-02
## 2744        NA         NA         NA         NA
## 2745 7.654e-07  2.111e-02  1.361e-07  2.401e-03
## 2746 2.142e-07  2.708e-01  5.950e-07  7.651e-01
## 2747 1.087e-01  1.512e-06  2.312e-02  4.621e-07
## 2748 6.225e-01  6.392e-03  3.123e-01  4.978e-03
## 2800 5.492e-05  1.470e-10  1.636e-03  4.406e-09
## 2801 2.568e-02  2.867e-01  1.842e-02  1.805e-01
## 2802 1.376e-06  2.801e-03  7.530e-08  1.222e-04
## 2803 4.498e-02  5.127e-01  5.142e-02  6.450e-01
## 2804 1.719e-01  1.332e-01  1.275e-01  9.925e-02
## 2805 2.959e-02  7.795e-01  2.328e-02  4.147e-01
## 2806 1.407e-01  1.699e-01  8.913e-02  1.052e-01
## 2807 1.845e-01  7.982e-05  6.258e-01  2.347e-04
```

```
## 2808 7.433e-09  7.409e-01  4.618e-09  2.086e-01
## 2809 2.092e-01  1.014e-02  3.520e-01  1.550e-02
## 2900 2.433e-04  2.429e-32  6.047e-03  2.684e-27
## 2901 2.294e-01  4.820e-02  5.878e-01  1.119e-01
## 2902 1.169e-03  3.738e-91  6.312e-74 7.690e-151
## 2903 4.996e-02  1.909e-03  1.863e-01  3.197e-02
## 2904 5.800e-01  1.614e-03  3.616e-01  1.027e-03
## 2905 3.626e-03  6.306e-06  3.843e-01  1.076e-03
## 2906 2.147e-16  1.171e-05  1.167e-18  3.320e-02
## 2907 3.283e-02  6.800e-11  1.232e-25  4.154e-35
## 2908 4.662e-02  2.754e-02  3.937e-01  2.012e-01
## 2909 3.352e-03  4.173e-06  1.117e-01  2.223e-04
## 2910 2.130e-02  6.091e-01  2.887e-03  5.383e-02
## 2911 9.477e-03  1.614e-08  1.894e-01  1.358e-07
## 2912 5.412e-01  7.560e-14  7.787e-08  1.428e-19
## 2913 6.779e-29  3.777e-08  2.018e-86  2.251e-50
## 2914 2.398e-05  6.710e-03  5.876e-04  8.592e-01
## 2915        NA         NA         NA         NA
## 2916 4.073e-02  4.084e-14  9.542e-01  2.006e-12
## 2917 2.911e-03  2.714e-04  6.387e-02  9.426e-03
## 2918        NA         NA         NA         NA
## 2919 6.012e-21  2.373e-01  1.048e-33  1.261e-11
## 2920 5.719e-07  1.851e-01  7.417e-11  6.088e-06
## 2921 1.744e-02  5.325e-02  1.032e-01  3.694e-01
## 2922 4.894e-03  5.204e-01  6.264e-03  5.345e-01
## 2923 1.262e-01  1.426e-03  9.286e-01  5.685e-03
## 3000 1.151e-01  3.103e-02  3.068e-02  9.262e-03
## 3001        NA         NA         NA         NA
## 3002 7.180e-02  1.123e-01  8.174e-02  1.285e-01
## 3003 2.599e-01  4.176e-12  6.409e-01  1.386e-11
## 3004 2.413e-12  8.317e-03  1.969e-11  5.921e-02
## 3005 1.247e-10  5.582e-07  1.259e-12  5.040e-09
## 3100 4.087e-02  2.440e-02  1.129e-02  7.419e-03
## 3101 3.617e-03  4.432e-02  1.509e-02  1.925e-01
## 3102 1.839e-04  6.801e-07  5.997e-07  3.611e-09
## 3103 3.593e-02  9.165e-01  3.732e-02  7.643e-01
## 3104 3.612e-01  6.686e-02  2.296e-01  4.639e-02
## 3105 8.172e-01  7.197e-01  7.641e-01  6.884e-01
## 3106 1.737e-01  3.340e-02  4.974e-01  8.160e-02
## 3107 1.455e-01  7.147e-01  1.284e-01  5.410e-01
## 3108 1.140e-01  3.481e-01  9.232e-02  2.725e-01
## 3109 8.208e-01  2.137e-02  3.912e-01  1.402e-02
## 3110 1.329e-04  2.571e-02  3.845e-05  6.683e-03
## 3200 7.964e-11  7.377e-01  1.300e-11  9.368e-03
## 3201 5.289e-01  4.700e-01  7.314e-01  6.228e-01
## 3202 1.153e-03  4.969e-01  3.423e-04  6.696e-02
## 3203 8.319e-06  3.079e-05  5.790e-04  2.046e-03
## 3204 2.681e-07  6.149e-02  2.311e-06  7.247e-01
## 3205 1.697e-07  6.464e-03  9.567e-10  2.146e-05
## 3206 2.145e-08  5.347e-01  1.306e-08  8.013e-02
```

```
## 3207 3.821e-07  3.509e-01  1.353e-08  2.676e-03
## 3300 3.220e-03  6.273e-03  1.505e-11  2.276e-11
## 3301 3.887e-06  1.848e-01  7.302e-06  1.158e-01
## 3302 3.378e-01  5.562e-01  8.729e-02  1.244e-01
## 3303 2.087e-01  2.784e-03  3.063e-01  4.098e-03
## 3304 1.154e-55  1.333e-07 1.530e-104  3.447e-59
## 3305 2.853e-09  1.786e-02  2.442e-26  7.261e-20
## 3306 8.871e-06  3.302e-07  9.612e-03  4.078e-04
## 3307 2.236e-01  5.509e-01  2.876e-01  8.120e-01
## 3308 5.312e-09  4.396e-02  5.303e-09  1.713e-02
## 3309 1.235e-02  4.885e-09  2.275e-24  6.977e-31
## 3310 2.010e-04  1.538e-02  4.435e-11  1.576e-09
## 3311 5.823e-01  6.024e-01  2.794e-01  2.794e-01
## 3312 1.324e-09  3.966e-05  3.507e-29  4.066e-25
## 3313 9.507e-05  3.657e-01  7.590e-06  1.729e-02
## 3314 6.409e-02  5.402e-01  5.336e-02  3.421e-01
## 3315 3.617e-03  1.648e-02  1.350e-08  4.955e-08
## 3316 1.721e-05  9.586e-01  9.027e-15  3.681e-11
## 3317 1.035e-04  5.685e-01  7.686e-07  1.248e-03
## 3318 2.575e-01  3.155e-01  4.059e-02  4.621e-02
## 3319 5.403e-03  6.985e-01  2.126e-03  1.351e-01
## 3320 5.175e-01  1.033e-01  4.286e-04  1.133e-04
## 3321 1.400e-01  8.668e-02  6.467e-01  3.338e-01
## 3322 8.572e-01  9.975e-01  8.106e-01  8.728e-01
## 3400 7.891e-01  7.789e-09  2.800e-01  7.335e-09
## 3401 1.229e-01  2.764e-01  1.589e-02  2.828e-02
## 3402 1.273e-03  7.178e-01  1.145e-03  2.759e-01
## 3403 4.938e-02  4.297e-04  4.730e-03  6.320e-05
## 3404 4.684e-02  8.947e-05  4.678e-01  9.554e-04
## 3500 8.330e-15  4.015e-01  3.792e-16  9.197e-04
## 3501 1.042e-01  4.857e-01  6.505e-02  3.486e-01
## 3502        NA         NA         NA         NA
## 3503        NA         NA         NA         NA
## 3504 2.483e-06  5.590e-03  4.401e-08  1.329e-04
## 3505 1.023e-01  2.086e-01  1.649e-01  3.489e-01
## 3506 7.981e-01  9.602e-01  8.004e-01  9.715e-01
## 3600 9.614e-04  2.888e-01  2.654e-03  9.011e-01
## 3601 1.452e-02  4.804e-03  1.573e-01         NA
## 3602 3.404e-02  6.415e-01  4.282e-02  7.928e-01
## 3603 1.245e-02  9.324e-01  3.836e-03  9.353e-02
## 3604        NA         NA         NA         NA
## 3605 2.288e-03  7.413e-02  3.467e-04  8.667e-03
## 3606        NA         NA         NA         NA
## 3607 6.238e-02  3.105e-01  4.746e-02  2.165e-01
## 3608        NA         NA         NA         NA
## 3609 8.663e-01  7.867e-04  2.219e-01  3.984e-04
## 3610 6.965e-01  9.115e-03  7.709e-01  1.041e-02
## 3611 2.447e-06  7.467e-01  6.148e-07  5.771e-02
## 3612 8.058e-01  1.427e-10  9.158e-02  1.336e-10
## 3613 2.437e-01  4.977e-01  2.805e-01  6.164e-01
```

```
## 3614 1.076e-05  3.619e-20  4.218e-10  1.079e-23
## 3615        NA         NA         NA         NA
## 3616 1.935e-04  5.234e-02  7.588e-06  1.129e-03
```

```r
print(RegStar)
```

```
##      FFA1p FLA1p 2p    3p    4p    5+p   FFA2p FLA2p FFA3p FLA4p
## 1000 "***" "***" "***" "***" "***" "***" ""    "***" ""    "***"
## 1100 ""    "***" "***" "***" "***" "***" "**"  "***" "*"   "***"
## 1101 ""    "***" "***" "***" "***" "***" ""    "***" "**"  "***"
## 1102 ""    ""    "***" "***" "***" "***" ""    ""    ""    ""
## 1103 "***" "**"  "***" "***" "***" "***" "***" "**"  "***" "***"
## 1104 "***" ""    "***" "***" "***" "***" "***" ""    "***" ""
## 1105 "***" "***" "***" "***" "***" "***" "***" "***" "***" "***"
## 1106 "*"   "**"  "***" "***" "***" "***" "***" "***" "***" "**"
## 1107 ""    "***" "***" "***" "***" "***" ""    "***" ""    "***"
## 1108 ""    ""    "**"  "***" "***" "***" "*"   ""    "*"   ""
## 1109 "***" ""    "***" "***" "***" "***" "***" ""    "***" ""
## 1110 "***" "***" "***" "***" "***" "***" "***" "***" "***" "***"
## 1111 ""    ""    "***" "***" "***" "***" ""    ""    ""    ""
## 1200 ""    ""    "**"  "***" "*"   "*"   ""    ""    ""    ""
## 1201 "***" ""    "***" "***" "***" "***" "***" ""    "***" ""
## 1202 "*"   ""    "***" "***" "***" "***" ""    ""    "***" "**"
## 1203 "*"   "**"  "***" "***" "***" "***" "***" "**"  "***" "***"
## 1204 ""    ""    "***" "***" "***" "***" ""    ""    ""    ""
## 1205 ""    ""    ""    "***" "***" ""    ""    ""    ""    ""
## 1206 ""    ""    ""    "*"   "**"  "**"  ""    ""    "*"   ""
## 1207 "***" "*"   "***" "***" "***" "***" "***" "*"   "***" "***"
## 1208 "*"   "**"  "***" "***" "***" "*"   "**"  "***" "**"  "***"
## 1209 ""    ""    "***" "***" "***" ""    ""    ""    "***" "***"
## 1210 ""    ""    "***" "***" "***" "**"  ""    ""    ""    ""
## 1211 ""    ""    "***" "***" "***" "***" "**"  ""    "***" "***"
## 1212 ""    "*"   "***" "***" "***" "***" ""    ""    "*"   ""
## 1213 ""    ""    "***" "***" "**"  "***" ""    ""    ""    ""
## 1300 "***" "***" "***" "***" "***" "***" ""    "***" "**"  "***"
## 1301 ""    ""    "*"   ""    ""    "**"  ""    "*"   ""    ""
## 1302 ""    "*"   "***" "***" "***" "***" ""    ""    ""    ""
## 1303 "***" "***" "***" "***" "***" "***" ""    "***" ""    "***"
## 1304 "***" "***" "***" "***" "***" "***" ""    "***" ""    "***"
## 1305 "*"   ""    "***" "***" "***" "***" ""    "*"   ""    "*"
## 1306 "*"   "*"   "***" "***" "***" "***" ""    "***" ""    "***"
## 1307 "***" "***" "***" "***" "***" "***" "**"  "***" "***" "***"
## 1308 "**"  ""    "***" "***" "***" "***" "***" ""    "***" ""
## 1309 "*"   ""    "***" "***" "***" "***" ""    "*"   ""    "**"
## 1310 "***" ""    "***" "***" "***" "***" "***" ""    "***" ""
## 1311 "***" "***" "***" "***" "***" "***" "***" "***" "***" "***"
## 1312 "***" "***" "***" "***" "***" "***" "***" "***" "***" "***"
## 1313 ""    ""    "***" "***" "***" "***" "*"   ""    "*"   ""
## 1314 ""    ""    "***" "***" "***" "***" ""    "**"  ""    "**"
## 1315 "**"  ""    "***" "***" "***" "***" "*"   ""    "*"   ""
```

```
## 1400 ""    ""    "***" "***" "***" "***" "*"   ""    "***" "**"
## 1401 ""    "*"   "***" "***" "***" "**"  ""    "**"  "**"  "***"
## 1402 ""    "*"   "***" "***" "***" ""    ""    ""    ""    "*"
## 1403 "**"  ""    "***" "***" "***" "***" "***" ""    "***" ""
## 1404 "**"  ""    "***" "***" "***" "***" "**"  ""    "**"  ""
## 1405 "***" "*"   "***" "***" "***" "***" "***" "**"  "***" "***"
## 1406 "**"  ""    "***" "***" "***" "***" "***" ""    "***" ""
## 1407 "***" ""    "***" "***" "***" "***" "***" ""    "***" "***"
## 1408 ""    ""    "***" "***" "***" "***" ""    ""    ""    ""
## 1409 "**"  ""    "***" "***" "***" "***" "**"  ""    "***" "*"
## 1410 ""    ""    "***" "***" "***" ""    ""    ""    "*"   ""
## 1500 "***" "***" "***" "***" "***" "***" ""    "***" ""    "***"
## 1501 "*"   ""    "***" "***" "***" "***" "***" ""    "***" ""
## 1502 "*"   ""    "***" "***" "***" "***" ""    "*"   ""    "*"
## 1503 ""    "*"   "***" "***" "***" "***" ""    "*"   ""    "*"
## 1504 ""    ""    "***" "***" "***" "***" "*"   ""    "*"   ""
## 1505 "*"   "**"  "***" "***" "***" "***" ""    "*"   ""    "*"
## 1506 ""    ""    "***" "***" "***" "***" ""    ""    ""    ""
## 1507 ""    ""    "***" "***" "***" "***" ""    ""    ""    ""
## 1508 ""    ""    "***" "***" "***" "***" "*"   ""    "*"   ""
## 1600 "**"  "***" "***" "***" "***" "***" ""    "***" ""    "***"
## 1601 ""    ""    "***" "***" "***" "***" ""    ""    ""    ""
## 1602 "**"  "*"   "***" "***" "***" "***" "***" ""    "***" ""
## 1603 ""    ""    "***" "***" "***" "***" ""    ""    ""    ""
## 1604 "**"  ""    "***" "***" "***" "***" "***" ""    "***" ""
## 1605 ""    "**"  "***" "***" "***" "***" "*"   "**"  "*"   "*"
## 1606 "**"  ""    "***" "***" "***" "***" ""    ""    ""    ""
## 1607 "**"  ""    "***" "***" "***" "***" "***" ""    "***" ""
## 1700 ""    ""    "***" "***" "***" "***" ""    ""    ""    ""
## 1701 ""    ""    "***" "***" "*"   "*"   ""    ""    ""    ""
## 1702 ""    "*"   "***" "***" "***" "***" ""    "*"   "**"  "**"
## 1703 ""    ""    "***" "***" "***" "***" ""    ""    ""    ""
## 1704 ""    ""    "***" "***" "***" "***" ""    ""    ""    ""
## 1705 ""    ""    "***" "***" "***" "***" ""    ""    ""    ""
## 1706 "***" "***" "***" "***" "***" "***" "***" "***" "***" "***"
## 1707 ""    "**"  "***" "***" "***" "***" ""    "**"  ""    "**"
## 1708 ""    ""    "***" "***" "***" "***" "*"   ""    "*"   ""
## 1709 "***" "*"   "***" "***" "***" "***" "***" ""    "***" "***"
## 1710 "*"   ""    "***" "***" "***" "***" "*"   "**"  "***" "***"
## 1711 ""    "***" "***" "***" "***" "***" ""    "**"  ""    "**"
## 1712 ""    ""    "***" "***" "***" "***" "*"   ""    "*"   ""
## 1800 "*"   ""    "**"  "***" "**"  "***" "*"   ""    "**"  "**"
## 1801 ""    ""    ""    ""    ""    ""    ""    ""    ""    ""
## 1802 ""    ""    "***" "***" "***" "***" ""    "*"   ""    ""
## 1803 ""    ""    "***" "***" "***" "***" ""    ""    ""    ""
## 1804 ""    ""    "***" "***" "***" "***" ""    ""    ""    ""
## 1900 "**"  "***" "***" "***" "***" "***" "*"   "***" "***" "***"
## 1901 ""    "*"   "***" "***" "***" "***" ""    "*"   ""    "*"
## 1902 ""    ""    "***" "***" "***" "***" "**"  ""    "**"  ""
## 1903 ""    ""    "***" "***" "***" "***" ""    ""    ""    ""
```

```
## 1904 ""    ""    "***" "***" "***" "***" "*"   ""    ""    ""
## 1905 ""    ""    ""    ""    ""    "*"   ""    ""    ""    ""
## 1906 ""    ""    "***" "***" "***" "***" ""    ""    ""    ""
## 1907 ""    ""    "***" "***" "***" "***" ""    ""    ""    ""
## 1908 ""    "**"  "***" "***" "***" "***" ""    "*"   ""    "*"
## 1909 ""    ""    "***" "***" "***" "***" ""    ""    ""    ""
## 1910 ""    "*"   "***" "***" "***" "***" "*"   ""    "*"   ""
## 1911 ""    "***" "***" "***" "***" "***" "*"   "***" ""    "**"
## 1912 ""    "*"   "***" "***" "***" "***" ""    ""    ""    ""
## 1913 ""    ""    ""    ""    ""    "**"  ""    ""    ""    ""
## 2000 ""    "*"   "***" "***" ""    "***" ""    "**"  "***" "***"
## 2001 ""    ""    "***" "***" "***" ""    ""    ""    ""    ""
## 2002 "***" "*"   "***" "***" "***" "***" "***" "**"  "***" "***"
## 2003 "***" ""    "***" "***" "***" "***" "***" ""    "***" "**"
## 2100 ""    ""    "***" "***" "***" "***" ""    ""    ""    ""
## 2101 ""    ""    "**"  "***" "***" "***" ""    ""    ""    ""
## 2102 ""    ""    "***" "***" "***" "***" "*"   ""    "*"   ""
## 2103 ""    ""    "***" "***" "***" "***" ""    ""    ""    ""
## 2104 "**"  "***" "***" "***" "***" "***" "*"   "***" "**"  "***"
## 2105 ""    "***" "***" "***" "***" "***" ""    "***" ""    "***"
## 2200 "***" ""    "***" "***" "***" "***" "***" ""    "***" "***"
## 2201 ""    "*"   "***" "***" "***" "***" ""    "***" "**"  "***"
## 2202 "***" "***" "***" "***" "***" "***" "***" "***" "***" "***"
## 2203 ""    ""    "***" "***" "***" "***" ""    ""    ""    ""
## 2204 ""    ""    "***" "***" "***" "***" "***" "*"   "***" ""
## 2205 ""    ""    "***" "***" "***" "***" "*"   ""    "*"   ""
## 2206 ""    ""    "***" "***" "***" "***" ""    ""    ""    ""
## 2207 ""    "*"   "***" "***" "***" "***" ""    "*"   ""    ""
## 2208 ""    "*"   "***" "***" "***" "***" ""    "*"   ""    ""
## 2209 ""    "**"  "***" "***" "***" "***" "***" "***" "*"   "*"
## 2210 ""    ""    "***" "***" "***" "***" "**"  ""    "**"  ""
## 2211 ""    ""    "***" "***" "***" "***" "***" ""    "***" "**"
## 2212 ""    ""    "***" "***" "***" "***" ""    ""    ""    ""
## 2213 "*"   ""    "***" "***" "***" "***" "***" ""    "***" ""
## 2214 "*"   ""    "***" "***" "***" "***" ""    ""    ""    ""
## 2215 "**"  "*"   "***" "***" "***" "***" "***" ""    "***" ""
## 2216 ""    ""    "***" "***" "***" "**"  ""    ""    ""    ""
## 2300 "**"  ""    "***" "***" "***" "***" "***" ""    "***" ""
## 2301 ""    "*"   "***" "***" "***" "***" ""    ""    ""    ""
## 2302 ""    ""    "***" "***" "***" "***" ""    ""    ""    ""
## 2303 "***" ""    "***" "***" "***" "***" "***" ""    "***" "*"
## 2304 "*"   ""    "***" "***" "***" "***" "***" ""    "***" "*"
## 2305 ""    ""    "***" "***" "***" "***" "***" ""    "***" ""
## 2306 ""    ""    "***" "***" "***" "***" ""    ""    ""    ""
## 2307 "***" ""    "***" "***" "***" "***" "***" ""    "***" "**"
## 2308 ""    ""    "***" "***" "***" "***" ""    ""    ""    ""
## 2309 ""    ""    "***" "***" "***" "***" ""    ""    ""    ""
## 2310 ""    "**"  "***" "***" "***" "***" "***" "*"   "***" ""
## 2311 ""    ""    "***" "***" "***" "***" "*"   "*"   ""    "*"
## 2312 ""    "***" "***" "***" "***" "***" "**"  "*"   "**"  "*"
```

```
## 2400 "***" "*"   "***" "***" "***" "***" "*"   "*"   "*"   "*"
## 2401 ""    ""    "***" "*"   "***" "***" "*"   ""    "*"   ""
## 2402 ""    ""    "***" "***" "***" "***" ""    ""    ""    ""
## 2403 "*"   "***" "***" "***" "***" "***" ""    "***" "*"   "***"
## 2404 ""    ""    "***" "***" "***" "***" ""    ""    ""    ""
## 2405 "**"  ""    "***" "***" "***" "***" "***" "*"   "***" ""
## 2406 ""    "***" "***" "***" "***" "***" ""    "***" ""    "***"
## 2500 ""    ""    "***" "***" "***" "***" "***" ""    "***" ""
## 2501 ""    "*"   "***" "***" "***" "***" ""    "*"   "*"   "**"
## 2502 "**"  "***" "***" "***" "***" "***" "***" "***" "***" "***"
## 2503 "***" ""    "***" "***" "***" "***" "***" ""    "***" ""
## 2504 ""    ""    "***" "***" "***" "***" ""    ""    ""    ""
## 2505 ""    ""    "***" "***" "***" "***" "***" ""    "***" "*"
## 2506 ""    ""    "***" "***" "***" "***" "***" ""    "***" "*"
## 2507 ""    ""    "***" "***" "***" "***" "***" ""    "***" ""
## 2508 ""    ""    "***" "***" "***" "***" ""    ""    ""    ""
## 2600 "*"   "*"   "***" "***" "***" "***" ""    ""    "**"  "***"
## 2601 "*"   ""    "***" "***" "***" "***" "**"  ""    "***" "**"
## 2602 ""    "*"   "***" "***" ""    "***" ""    ""    ""    ""
## 2603 ""    ""    "***" "***" "***" "***" ""    ""    ""    ""
## 2604 "***" "**"  "***" "***" "***" "***" "*"   "*"   "**"  "***"
## 2605 "**"  ""    "***" "***" "***" "***" ""    ""    ""    ""
## 2606 ""    "**"  "***" "***" "***" "***" ""    "**"  ""    "*"
## 2607 ""    ""    "***" "***" "***" "***" ""    ""    ""    ""
## 2608 ""    "*"   "***" "***" "***" "***" ""    "*"   ""    "*"
## 2609 ""    ""    "**"  "***" "*"   ""    ""    ""    ""    ""
## 2610 ""    ""    "***" "***" "***" "***" ""    ""    ""    ""
## 2611 ""    ""    "***" "***" "***" "***" ""    ""    ""    ""
## 2612 ""    "*"   "***" "***" "**"  "***" ""    "**"  ""    "**"
## 2613 "***" ""    "***" "***" "***" "***" ""    ""    ""    ""
## 2614 ""    ""    "***" "***" "***" "***" ""    ""    ""    ""
## 2700 ""    "***" "***" "***" "***" "***" "***" "***" ""    "***"
## 2701 "***" ""    "***" "***" "***" "***" "***" ""    "***" "***"
## 2702 ""    ""    "***" "***" "***" "***" ""    "**"  ""    "**"
## 2703 "**"  "*"   "***" "***" "***" "***" "***" "*"   "***" "***"
## 2704 ""    ""    "***" "***" "***" "***" "***" ""    "***" ""
## 2705 "***" "***" "***" "***" "***" "***" "***" ""    "***" "***"
## 2706 ""    ""    "***" "***" "***" "***" ""    "***" ""    "***"
## 2707 "*"   ""    "***" "***" "***" "***" "**"  "***" ""    "**"
## 2708 "***" "*"   "***" "***" "***" "***" ""    "*"   ""    "*"
## 2709 ""    ""    "**"  "*"   "**"  "**"  ""    ""    ""    ""
## 2710 ""    ""    "***" "***" "***" "***" ""    ""    ""    ""
## 2711 "*"   "*"   "***" "***" "***" "***" ""    "***" ""    "***"
## 2712 "*"   ""    "***" "***" "***" "***" "***" "**"  "***" "*"
## 2713 ""    ""    "***" "***" "***" "***" "***" ""    "***" ""
## 2714 ""    ""    "***" "***" "***" "***" "**"  ""    "**"  ""
## 2715 "***" "**"  "***" "***" "***" "***" "*"   "***" "***" "***"
## 2716 "***" "***" "***" "***" "***" "***" "***" "***" "***" "***"
## 2717 ""    ""    "***" "***" "***" "***" "*"   "***" ""    "***"
## 2718 ""    "**"  "***" "***" "***" "***" ""    "***" ""    "***"
```

```
## 2719 ""    "**"  "***" "***" "***" "***" "***" "***" ""    ""
## 2720 ""    "**"  "***" "***" "***" "***" "***" "**"  "***" "*"
## 2721 ""    ""    "***" "***" "***" "***" ""    ""    ""    ""
## 2722 ""    ""    "**"  "***" "***" "***" ""    ""    ""    ""
## 2723 ""    "*"   "***" "***" "***" "***" "*"   "**"  ""    "**"
## 2724 ""    ""    "***" "***" "***" "***" "***" "**"  "***" "*"
## 2725 ""    ""    "***" "***" "***" "***" "**"  "*"   "*"   "*"
## 2726 "**"  ""    "***" "***" "***" "***" "***" "*"   "***" "***"
## 2727 "**"  "***" "***" "***" "***" "***" ""    "***" "*"   "***"
## 2728 "***" "**"  "***" "***" "***" "***" "***" ""    "***" ""
## 2729 "***" "**"  "***" "***" "***" "***" "***" ""    "***" ""
## 2730 ""    "*"   "***" "***" "***" "***" "***" "***" "*"   "***"
## 2731 "***" ""    "***" "***" "***" "***" ""    ""    ""    ""
## 2732 ""    ""    "***" "***" "***" "***" ""    "***" ""    "***"
## 2733 "***" "***" "***" "***" "***" "***" "***" "***" "***" "***"
## 2734 ""    ""    "***" "***" "***" "***" ""    "**"  ""    "**"
## 2735 "***" "***" "***" "***" "***" "***" "***" ""    "***" "*"
## 2736 ""    "***" "***" "***" "***" "***" "**"  "***" ""    "***"
## 2737 ""    ""    "***" "***" "***" "***" ""    ""    ""    ""
## 2738 "***" ""    "***" "***" "***" "***" "***" "***" "***" "**"
## 2739 "***" "***" "***" "***" "***" "***" "***" "***" "***" "***"
## 2740 ""    ""    "***" "***" "***" "***" "**"  ""    "**"  ""
## 2741 ""    "***" "***" "***" "***" "***" "*"   "***" ""    "***"
## 2742 ""    "***" "***" "***" "***" "***" ""    "***" "*"   "***"
## 2743 "*"   ""    "***" "***" "***" "***" ""    "*"   ""    "*"
## 2744 ""    ""    ""    ""    ""    ""    ""    ""    ""    ""
## 2745 "*"   ""    "***" "***" "***" "***" "***" "*"   "***" "**"
## 2746 ""    ""    "***" "***" "***" "***" "***" ""    "***" ""
## 2747 ""    "**"  "**"  "***" "***" "***" ""    "***" "*"   "***"
## 2748 "***" ""    "***" "***" "***" "***" ""    "**"  ""    "**"
## 2800 ""    "***" "***" "***" "***" "***" "***" "***" "**"  "***"
## 2801 ""    ""    "**"  "**"  "***" "***" "*"   ""    "*"   ""
## 2802 "***" "***" "***" "***" "***" "***" "***" "**"  "***" "***"
## 2803 "*"   ""    "*"   "***" "***" "***" "*"   ""    ""    ""
## 2804 "*"   ""    "***" "***" "***" "***" ""    ""    ""    ""
## 2805 ""    ""    "***" "***" "***" "***" "*"   ""    "*"   ""
## 2806 ""    ""    "***" "***" "***" "***" ""    ""    ""    ""
## 2807 ""    "**"  "***" "***" "***" "***" ""    "***" ""    "***"
## 2808 "***" "*"   "***" "***" "***" "***" "***" ""    "***" ""
## 2809 ""    "**"  "***" "***" "***" "***" ""    "*"   ""    "*"
## 2900 "**"  "*"   "***" "***" "***" "***" "***" "***" "**"  "***"
## 2901 ""    ""    "***" "***" "***" "***" ""    "*"   ""    ""
## 2902 ""    ""    "***" "***" "***" "***" "**"  "***" "***" "***"
## 2903 ""    ""    "***" "***" "***" "***" "*"   "**"  ""    "*"
## 2904 ""    ""    "**"  "***" "***" "***" ""    "**"  ""    "**"
## 2905 "*"   ""    "***" "***" "***" "***" "**"  "***" ""    "**"
## 2906 "***" ""    "***" "***" "***" "***" "***" "***" "***" "*"
## 2907 ""    ""    "***" "***" "***" "***" "*"   "***" "***" "***"
## 2908 "*"   ""    ""    "***" "***" "***" "*"   "*"   ""    ""
## 2909 "**"  ""    "***" "***" "***" "***" "**"  "***" ""    "***"
```

```
## 2910 ""    ""    "***" "***" "***" "***" "*"   ""    "**"  ""
## 2911 "*"   ""    "***" "***" "***" "***" "**"  "***" ""    "***"
## 2912 ""    "***" "***" "***" "***" "***" ""    "***" "***" "***"
## 2913 "***" "***" "***" "***" "***" "***" "***" "***" "***" "***"
## 2914 "***" ""    "***" "***" "***" "***" "***" "**"  "***" ""
## 2915 ""    ""    ""    ""    ""    ""    ""    ""    ""    ""
## 2916 ""    "***" "***" "***" "***" "***" "*"   "***" ""    "***"
## 2917 "*"   ""    "***" "***" "***" "***" "**"  "***" ""    "**"
## 2918 ""    ""    ""    ""    ""    ""    ""    ""    ""    ""
## 2919 "***" "**"  "***" "***" "***" "***" "***" ""    "***" "***"
## 2920 "***" ""    "***" "***" "***" "***" "***" ""    "***" "***"
## 2921 "**"  ""    "***" "***" "***" "***" "*"   ""    ""    ""
## 2922 "*"   ""    ""    "**"  "**"  "**"  "**"  ""    "**"  ""
## 2923 ""    "*"   "***" "***" "***" "***" ""    "**"  ""    "**"
## 3000 ""    "**"  "***" "***" "***" "***" ""    "*"   "*"   "**"
## 3001 ""    ""    ""    ""    ""    ""    ""    ""    ""    ""
## 3002 ""    "*"   "***" "***" "***" "***" ""    ""    ""    ""
## 3003 ""    "***" "***" "***" "***" "***" ""    "***" ""    "***"
## 3004 "***" ""    "***" "***" "***" "***" "***" "**"  "***" ""
## 3005 "***" "***" "***" "***" "***" "***" "***" "***" "***" "***"
## 3100 "***" ""    "***" "***" "***" "***" "*"   "*"   "*"   "**"
## 3101 ""    "**"  "***" "***" "***" "***" "**"  "*"   "*"   ""
## 3102 "**"  "***" "***" "***" "***" "***" "***" "***" "***" "***"
## 3103 ""    ""    "***" "***" "***" "***" "*"   ""    "*"   ""
## 3104 ""    ""    "***" "***" "***" "***" ""    ""    ""    "*"
## 3105 "*"   ""    "***" "***" "***" "***" ""    ""    ""    ""
## 3106 ""    ""    "***" "***" "***" "***" ""    "*"   ""    ""
## 3107 ""    ""    "***" "***" "***" "***" ""    ""    ""    ""
## 3108 ""    ""    "***" "***" "***" "***" ""    ""    ""    ""
## 3109 ""    ""    "***" "***" "***" "***" ""    "*"   ""    "*"
## 3110 "*"   ""    "***" "***" "***" "***" "***" "*"   "***" "**"
## 3200 "*"   ""    "***" "***" "***" "***" "***" ""    "***" "**"
## 3201 ""    ""    "***" "***" "***" "***" ""    ""    ""    ""
## 3202 ""    ""    "***" "***" "***" "***" "**"  ""    "***" ""
## 3203 ""    "***" "***" "***" "***" "***" "***" "***" "***" "**"
## 3204 "**"  ""    "***" "***" "***" "***" "***" ""    "***" ""
## 3205 "**"  ""    "***" "***" "***" "***" "***" "**"  "***" "***"
## 3206 "**"  ""    "***" "***" "***" "***" "***" ""    "***" ""
## 3207 "*"   ""    "***" "***" "***" "***" "***" ""    "***" "**"
## 3300 "**"  ""    "***" "***" "***" "***" "**"  "**"  "***" "***"
## 3301 ""    "*"   "***" "***" "***" "***" "***" ""    "***" ""
## 3302 ""    ""    "***" "***" "***" "***" ""    ""    ""    ""
## 3303 ""    "**"  "***" "***" "***" "***" ""    "**"  ""    "**"
## 3304 "***" "***" "***" "***" "***" "***" "***" "***" "***" "***"
## 3305 "***" "**"  "***" "***" "***" "***" "***" "*"   "***" "***"
## 3306 ""    "**"  "***" "***" "***" "***" "***" "***" "**"  "***"
## 3307 ""    ""    "***" "***" "***" "***" ""    ""    ""    ""
## 3308 "***" "**"  "***" "***" "***" "***" "***" "*"   "***" "*"
## 3309 "***" "***" "***" "***" "***" "***" "*"   "***" "***" "***"
## 3310 ""    "*"   "***" "***" "***" "***" "***" "*"   "***" "***"
```

```
## 3311 ""    ""    "***" "***" "***" "***" ""    ""    ""    ""
## 3312 "**"  "*"   "***" "***" "***" "***" "***" "***" "***" "***"
## 3313 "***" ""    "***" "***" "***" "*"   "***" ""    "***" "*"
## 3314 ""    ""    "***" "***" "***" "***" ""    ""    ""    ""
## 3315 ""    "**"  "***" "***" "***" "***" "**"  "*"   "***" "***"
## 3316 "***" ""    "***" "***" "***" "***" "***" ""    "***" "***"
## 3317 "**"  ""    "***" "***" "***" "***" "***" ""    "***" "**"
## 3318 ""    "**"  "***" "***" "***" "***" ""    ""    "*"   "*"
## 3319 ""    ""    "***" "***" "***" "***" "**"  ""    "**"  ""
## 3320 ""    ""    "***" "***" "***" "***" ""    ""    "***" "***"
## 3321 "*"   ""    "***" "***" "***" "***" ""    ""    ""    ""
## 3322 ""    ""    "***" "***" "***" "***" ""    ""    ""    ""
## 3400 ""    "*"   "***" "***" "***" "***" ""    "***" ""    "***"
## 3401 ""    ""    ""    "**"  ""    ""    ""    ""    "*"   "*"
## 3402 "**"  ""    ""    ""    "*"   "***" "**"  ""    "**"  ""
## 3403 ""    "***" ""    ""    ""    ""    "*"   "***" "**"  "***"
## 3404 ""    ""    "***" "***" "***" "***" "*"   "***" ""    "***"
## 3500 "**"  ""    "***" "***" "***" "***" "***" ""    "***" "***"
## 3501 ""    ""    ""    ""    ""    ""    ""    ""    ""    ""
## 3502 ""    ""    ""    ""    ""    ""    ""    ""    ""    ""
## 3503 ""    ""    ""    ""    ""    ""    ""    ""    ""    ""
## 3504 "**"  ""    "**"  "***" "***" "***" "***" "**"  "***" "***"
## 3505 ""    ""    "***" "***" "***" "***" ""    ""    ""    ""
## 3506 ""    ""    ""    ""    "*"   "**"  ""    ""    ""    ""
## 3600 ""    ""    "***" "***" "***" "***" "***" ""    "**"  ""
## 3601 ""    ""    "***" "***" "***" "***" "*"   "**"  ""    ""
## 3602 ""    ""    "**"  "***" "***" "***" "*"   ""    "*"   ""
## 3603 "**"  ""    "*"   "***" "**"  "***" "*"   ""    "**"  ""
## 3604 ""    ""    ""    ""    ""    ""    ""    ""    ""    ""
## 3605 ""    "*"   "***" "***" "***" "***" "**"  ""    "***" "**"
## 3606 ""    ""    ""    ""    ""    ""    ""    ""    ""    ""
## 3607 "*"   ""    "**"  "*"   "***" "***" ""    ""    "*"   ""
## 3608 ""    ""    ""    ""    ""    ""    ""    ""    ""    ""
## 3609 ""    "**"  "*"   "***" "***" "***" ""    "***" ""    "***"
## 3610 ""    ""    ""    "***" "***" "***" ""    "**"  ""    "*"
## 3611 "*"   ""    "***" "***" "***" "***" "***" ""    "***" ""
## 3612 ""    "**"  "***" "***" "***" "***" ""    "***" ""    "***"
## 3613 ""    ""    "*"   "*"   "***" "***" ""    ""    ""    ""
## 3614 "***" "***" "***" "***" "***" "***" "***" "***" "***" "***"
## 3615 ""    ""    ""    ""    ""    ""    ""    ""    ""    ""
## 3616 "**"  "**"  "***" "***" "***" "***" "***" ""    "***" "**"
```

```r
print("Gender-based average team sizes in 2018")
```

```
## [1] "Gender-based average team sizes in 2018"
```

```r
print(GenderAnalysed)
```

```
##      CitationSet Female1st FemaleLast
## 1000        2403       813        606
## 1100        2852      1225        841
```

```
## 1101          264         85          77
## 1102         1156        373         273
## 1103         1729        857         543
## 1104         1944        750         518
## 1105         4781       1907        1345
## 1106         1082        556         384
## 1107          792        257         183
## 1108          260         95          79
## 1109          948        374         238
## 1110         1261        457         392
## 1111          664        205         150
## 1200          229        114         119
## 1201         2187       1095         995
## 1202         2028        809         796
## 1203         1689        979         945
## 1204          438        205         176
## 1205           91         36          37
## 1206           91         62          50
## 1207          737        281         264
## 1208         1439        700         702
## 1209          106         68          62
## 1210          307        145         137
## 1211         2003        550         547
## 1212         1275        414         419
## 1213          955        529         520
## 1300         4679       1937        1317
## 1301          115         39          34
## 1302          296        134         110
## 1303         3550       1273         825
## 1304         1209        346         245
## 1305         1392        448         319
## 1306         2638       1189         878
## 1307         2514       1092         697
## 1308          727        320         222
## 1309          841        434         317
## 1310          962        546         396
## 1311         3201       1433        1019
## 1312         4360       1828        1220
## 1313         1383        557         342
## 1314         2012        904         574
## 1315          445        144          94
## 1400          597        181         185
## 1401          260         89          99
## 1402          588        143         138
## 1403         1108        399         362
## 1404          180         49          41
## 1405          716        200         170
## 1406          707        304         275
## 1407          777        325         283
## 1408         1580        542         499
```

```
## 1409          336        135         126
## 1410          211         63          66
## 1500         1240        299         233
## 1501           79         21          18
## 1502          986        272         194
## 1503         1080        291         188
## 1504           52         27          19
## 1505          462        114          71
## 1506           52          7           5
## 1507          321         60          39
## 1508          249         58          38
## 1600         3590       1030         711
## 1601          178         63          39
## 1602          853        286         183
## 1603          275         81          53
## 1604          777        213         161
## 1605         1343        378         231
## 1606         2335        566         371
## 1607          768        247         180
## 1700          545        108         109
## 1701           63         13          15
## 1702          102         23          18
## 1703          465         91          93
## 1704          225         38          27
## 1705          269         57          66
## 1706         1106        276         271
## 1707          177         32          22
## 1708          197         21          19
## 1709          261        106         107
## 1710          630        163         171
## 1711          292         47          26
## 1712          648         94          97
## 1800          163         46          37
## 1801           15          5           4
## 1802          200         52          47
## 1803          555        129         117
## 1804          545        123         106
## 1900          678        201         152
## 1901          433        128          82
## 1902         1209        381         255
## 1903          164         34          32
## 1904          841        288         233
## 1905           60         12          10
## 1906          571        166         103
## 1907          606        164         121
## 1908          613        178         111
## 1909          339         64          56
## 1910          644        259         186
## 1911          429        152         100
## 1912         1020        242         183
```

```
## 1913          96         29          19
## 2000         398         94          95
## 2001         293         93          81
## 2002        2981        731         707
## 2003        1042        199         195
## 2100         545        132         100
## 2101         119         25          28
## 2102         703        112          98
## 2103         502         95          80
## 2104         390         68          60
## 2105        1033        284         222
## 2200         431         99          85
## 2201         282         70          50
## 2202         515         63          55
## 2203         167         28          23
## 2204        1565        494         347
## 2205        1120        189         190
## 2206         134         12          14
## 2207         250         26          28
## 2208         891        118         105
## 2209         834        164         129
## 2210        2030        298         253
## 2211        1385        214         196
## 2212         284         76          48
## 2213         618        155         148
## 2214          74         30          25
## 2215         611         94          81
## 2216         160         57          57
## 2300        1094        438         347
## 2301         405        168         125
## 2302         240         73          55
## 2303        3100       1239         843
## 2304        1304        481         356
## 2305         712        238         186
## 2306         571        227         169
## 2307        1130        564         421
## 2308        1684        611         510
## 2309        1134        466         324
## 2310         980        355         260
## 2311         614        217         176
## 2312        1452        444         342
## 2400        1160        477         347
## 2401          30         17          14
## 2402         428        164         109
## 2403        2289       1164         740
## 2404        1715        866         573
## 2405         543        283         180
## 2406         922        470         292
## 2500        2653        537         420
## 2501         134         58          51
```

```
## 2502          736        254         185
## 2503          532        122          88
## 2504         1135        194         174
## 2505          721        181         130
## 2506          523        108          89
## 2507          709        155         125
## 2508         1048        241         175
## 2600         1095        150         151
## 2601          115         27          27
## 2602          462         59          53
## 2603          375         40          34
## 2604         1595        235         212
## 2605          467         67          60
## 2606          220         33          26
## 2607          304         57          55
## 2608          364         51          50
## 2609           87         13          11
## 2610          286         30          18
## 2611         1039        235         217
## 2612          133         19          14
## 2613         1198        246         202
## 2614          377         59          43
## 2700         2871       1324        1093
## 2701         1968       1118         855
## 2702          375        178         138
## 2703          995        429         321
## 2704          205        107          72
## 2705         3885       1096         789
## 2706          878        351         282
## 2707          306        158         143
## 2708         1258        666         531
## 2709            2          1           1
## 2710          104         62          52
## 2711         1365        540         447
## 2712         1381        778         560
## 2713         1523        939         705
## 2714          345        170         144
## 2715         1270        457         258
## 2716         1030        587         425
## 2717         1029        614         498
## 2718          870        436         357
## 2719         2703       1425        1207
## 2720         1225        573         375
## 2721          489        159         102
## 2722          449        193         162
## 2723         1787        873         587
## 2724          947        462         351
## 2725         2482       1265         935
## 2726         1067        496         352
## 2727          457        195         117
```

```
## 2728        4895      1854       1259
## 2729        1876      1321       1030
## 2730        4514      2158       1652
## 2731        1254       488        365
## 2732        3858       946        647
## 2733        1436       516        387
## 2734        1531       748        595
## 2735        4837      2923       2230
## 2736        1981       941        721
## 2737        1393       534        377
## 2738        5813      3393       2716
## 2739        6127      3708       3105
## 2740        1628       570        426
## 2741        3482      1148        813
## 2742        1076       607        470
## 2743         450       302        204
## 2744           0         0          0
## 2745         291       161        129
## 2746        7242      2173       1366
## 2747         614       263        169
## 2748        1335       444        230
## 2800        2561      1098        749
## 2801         259       123         93
## 2802        1092       601        482
## 2803         631       335        241
## 2804        1116       497        348
## 2805        1145       580        465
## 2806         335       195        131
## 2807         209       140        103
## 2808        1945       849        593
## 2809         578       204        158
## 2900        1385      1114       1003
## 2901         296       246        226
## 2902         656       471        423
## 2903         241       182        167
## 2904         135        80         64
## 2905         222       176        161
## 2906         355       295        275
## 2907         329       198        184
## 2908         263       214        191
## 2909         411       283        243
## 2910         279       123        120
## 2911         549       399        374
## 2912         296       246        239
## 2913         209       179        156
## 2914         199       136        129
## 2915          97        72         62
## 2916        1021       709        538
## 2917         228       190        162
## 2918           0         0          0
```

```
## 2919          212        188         172
## 2920           32         25          24
## 2921          387        267         232
## 2922          102         98          91
## 2923          131        111         113
## 3000          324        163         151
## 3001           78         37          31
## 3002          840        296         172
## 3003          972        379         234
## 3004         2332       1023         702
## 3005         1436        751         544
## 3100         2284        441         341
## 3101          661         79          75
## 3102          323         74          65
## 3103          989        214         190
## 3104         2651        431         355
## 3105          751        149         105
## 3106          578         74          49
## 3107         1333        194         155
## 3108          343         89          65
## 3109          399         32          31
## 3110          306         70          66
## 3200         1983       1054         950
## 3201          252        125         105
## 3202         1925       1064         907
## 3203         3179       1866        1500
## 3204         3748       2495        2207
## 3205         1686        859         741
## 3206          592        305         237
## 3207         2539       1473        1370
## 3300         1039        458         445
## 3301         1795       1031         956
## 3302          468        213         193
## 3303          781        335         329
## 3304         8051       4753        4590
## 3305         1995        777         729
## 3306         2628       1678        1476
## 3307          234         83          69
## 3308         2454       1039         996
## 3309         1088        616         630
## 3310         1945       1131        1096
## 3311          344        120         117
## 3312         5837       2669        2571
## 3313          341         86          86
## 3314         1452        795         774
## 3315         1782        961         951
## 3316         2416       1227        1234
## 3317          374        217         201
## 3318          846        618         579
## 3319          561        386         346
```

```
## 3320        1529       456        481
## 3321         752       291        271
## 3322         565       211        215
## 3400        1137       718        475
## 3401          34        19         17
## 3402          50        36         22
## 3403          83        59         42
## 3404         104        76         51
## 3500         326       117        104
## 3501          32        17         16
## 3502           0         0          0
## 3503           1         1          0
## 3504         296        77         65
## 3505          29        10          6
## 3506          32         9          5
## 3600          89        52         45
## 3601         131       100         86
## 3602          47        19         16
## 3603          80        53         45
## 3604           5         2          4
## 3605         267       131        105
## 3606           8         4          4
## 3607         214        98         76
## 3608           0         0          0
## 3609          98        80         67
## 3610          60        23         21
## 3611         251       134        108
## 3612        1495       648        551
## 3613          43         9          4
## 3614         510       218        153
## 3615           0         0          0
## 3616         604       434        378
```

```
write.table(BasicStats, file=paste("D:\\Downloads\\basicStats", Country,
MaxCountriesAllowed, " ", FirstYearForCitationAnalysis, "-",
LastYearForCitationAnalysis, "var3.txt", sep=""))
write.table(BasicStats2, file=paste("D:\\Downloads\\basicStats2", Country,
MaxCountriesAllowed," ",FirstYearForCitationAnalysis,"-
",LastYearForCitationAnalysis, "var3.txt", sep=""))
write.table(GenderTeamSize, file=paste("D:\\Downloads\\GenderTeamSize",
Country, MaxCountriesAllowed, " ", FirstYearForCitationAnalysis, "-",
LastYearForCitationAnalysis, "var3.txt", sep=""))
write.table(RegCoef, file=paste("D:\\Downloads\\RegCoef", Country,
MaxCountriesAllowed, " ", FirstYearForCitationAnalysis, "-",
LastYearForCitationAnalysis, "var3.txt", sep=""))
write.table(RegP, file=paste("D:\\Downloads\\RegP", Country,
MaxCountriesAllowed, " ", FirstYearForCitationAnalysis, "-",
LastYearForCitationAnalysis, "var3.txt", sep=""))
write.table(RegStar, file=paste("D:\\Downloads\\RegStar", Country,
```

```
MaxCountriesAllowed," ", FirstYearForCitationAnalysis, "-",
LastYearForCitationAnalysis, "var3.txt", sep=""))
```